SMART METER BASED PROFILING FOR LOAD FORECASTING AND DEMAND SIDE MANAGEMENT IN SMART GRIDS

by

ZAFAR ALI KHAN

A thesis submitted to The University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Electronic, Electrical and Systems Engineering
College of Engineering and Physical Sciences
University of Birmingham
November 2019

DEDICATION

I dedicate this thesis to my parents

ACKNOWLEDGEMENTS

Undertaking this PhD has been a true life changing experience for me and it would not have been possible without support and guidance that I received from many people. Firstly, I would like to first express my profound gratitude to my supervisor Dr. Dilan Jayaweera. His invaluable advice, suggestions and guidance with meticulous feedback and critique made this project a success.

I owe a special thanks to my family, my wife Shareen and my children, Mansur and Haniya who have been extremely supportive of me throughout the process of PhD and made countless sacrifices to help me get to this point. I owe huge thanks to my sisters Salma and Hina, my brothers Taimoor, Babar and Arslan who supported and helped me throughout my life and during the entire process of PhD.

I would like to acknowledge and thank my colleagues Dr. Manuel Alvarez-Alvarado, Hasan Gunduz, Abdullah Altamimi, Daniel Donaldson, Dr. Waqar Aslam, Dr. Hasan Almahroug, Waleed Abri and Bader Alharbi for their support and encouragement throughout my PhD. Special thanks goes to all friends who encouraged and supported me during the process.

ABSTRACT

The smart grid incorporates an integrated system of smart meters and communication networks that enable two-way communication between utilities and consumers. The granular information from smart meters can be used to improve the load forecast and influence consumer's energy consumption patterns through demand side management (DSM). However, for localized studies of power system, using a large quantity of smart meter data having high level of noise preclude the use of computationally intensive techniques. Reduction of smart meter data to extract the load profiles and smoother load profiles at lower aggregation level (individual consumer or small groups of consumers) are highly desirable for use in linear techniques for power system studies. Therefore, this thesis addresses the challenges of smart meter data size, complexity, variability and volatility for efficient use in load forecasting and DSM.

This thesis presents a novel clustering-based approach for analysis of smart meter data, aimed at more accurate and detailed load profiling, reduced profile complexity and improved load forecast accuracy and DSM solutions. The approach uses an innovative clustering algorithm to reduce the data size by proposing new cluster validity indices. The extremely volatile profiles having high levels of noise and complexity are linearized using Taylor series linearization process to alleviate the non-linearity and complexity of profiles. Finally, particle swarm optimization is applied for energy optimization in linearized profiles. The approach is demonstrated on Irish smart meter dataset and simulated PV data, to achieve improved load forecast accuracy using artificial neural network and improved DSM solutions using linear optimization with reduced computational burden.

Investigations suggest that proposed clustering algorithm can produce clusters with high intra-cluster pattern similarity as a result of the introduction of new stopping criteria specifically tailored for load forecasting applications. A comparison between the proposed alternative profiles and raw profiles further suggests that the alternative profiles guide the underlying energy consumption with reduced complexity making them computationally efficient. Use of the alternative profiles suggests that the load forecasting accuracy can potentially be higher compared to raw profiles. The alternative profiles in combination with the novel cluster selection approach provide higher peak reduction by shifting the load from peak hours to off-peak hours and higher monetary benefits for the participating consumers. The proposed clustering algorithm and the alternative profiles represent an advancement of the conventional load profiling approach, benefiting system operators through more accurate forecasting and efficient DSM. The novel mathematical framework suggested in this thesis provides an advancement to the new knowledge in the area of smart metering and smart power grids.

List of Publications

Parts of this research outcomes were published as a book chapter, journal articles and conferences. They are listed below.

1. **Z. A. Khan** and D. Jayaweera, "Planning and Operational Challenges in a Smart Grid," in Smart Power Systems and Renewable Energy System Integration, D. Jayaweera, Ed., ed Cham: Springer International Publishing, 2016, pp. 153-177.
2. **Z. A. Khan**, D. Jayaweera, and H. Gunduz, "Smart meter data taxonomy for demand side management in smart grids," in 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2016, pp. 1-8.
3. **Z. A. Khan** and D. Jayaweera, "Approach for smart meter load profiling in Monte Carlo simulation applications," IET Generation, Transmission & Distribution, vol. 11, pp. 1856-1864, 2017.
4. **Z. A. Khan** and D. Jayaweera, "Approach for forecasting smart customer demand with significant energy demand variability," in 2018 1st International Conference on Power, Energy and Smart Grid (ICPESG), 2018, pp. 1-5.
5. **Z. A. Khan**, D. Jayaweera, and M. S. Alvarez-Alvarado, "A novel approach for load profiling in smart power grids using smart meter data," Electric Power Systems Research, vol. 165, pp. 191-198, 2018.
6. H. Gunduz, **Z. A. Khan**, A. Altamimi, and D. Jayaweera, "An Innovative Methodology for Load and Generation Modelling in a Reliability Assessment with PV and Smart Meter Readings," in 2018 IEEE Power & Energy Society General Meeting (PESGM), 2018, pp. 1-5.
7. **Z. A. Khan** and D. Jayaweera, "Efficient Management of Demand in a Power Distribution System with Smart Meter Data," in 2019 IEEE Milan PowerTech, pp. 1-6. IEEE, 2019.

8. **Z. A. Khan** and D. Jayaweera, "Smart Meter Data Based Load Forecasting and Demand Side Management with Embedded PV Systems in Distribution Networks, " submitted to IEEE ACCESS

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
ANN	Artificial Neural Networks
AP	Alternative Profile
ARIMA	Autoregressive Integrated Moving Average
CER	Commission for Energy Regulation
CPCC	Cophenetic Correlation Coefficient
CPP	Critical Peaking Price
DB	Davies-Bouldin
DNO	Distribution Network Operator
DSM	Demand Side Management
EBP	Error Backpropagation
HAN	Home Area Network
HEMS	Home Energy Management System
ICTs	Information and Communication Technologies
ISSDA	Irish Social Science Data Archive
LM	Levenberg-Marquardt
LTLF	Long Term Load Forecast
LV	Low Voltage
MAPE	Mean Absolute Percentage Error
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression

MTLF	Medium Term Load Forecast
NAN	Neighbour-hood Area Network
Ofgem	Office of Gas and Electricity Markets
PSO	Particle Swarm Optimization
PV	Photovoltaic
RES	Renewable energy sources
RP	Raw Profile
RTP	Real Time Pricing
STLF	Short Term Load Forecast
ToU	Time of Use
VEE	Validation, Estimation and Editing
VSTLF	Very Short-Term Load Forecast

CHAPTER 1

1 Introduction

1.1 Background and Motivation

Smart grid incorporates an integrated system of smart meters, communication networks and data management systems that enable two-way communication between utilities and consumers [1]. Apart from the bidirectional flow of information, smart grid characterizes bidirectional flow of power due to increasing penetration of intermittent renewable energy sources (RES) at distribution network level. The integration of RES is evolving the conventional passive distribution networks into active distribution networks with power flow and voltages determined by the generation as well as load [2]. The intermittent nature of RES augments uncertainties in the distribution network necessitating a more localized and active distribution network management with enhanced consumer engagement to ensure secure and reliable operation of distribution network as well as to optimize the network capacity.

An influx of new data stream with granular information of consumers load patterns has become available with large-scale deployment of smart meters at individual consumer level in low voltage (LV) network that provides new opportunities for distribution network management. The scale of deployment of smart meters can be conceived from the fact that the U.K. is aiming to install over 50 million smart meters by 2020 [3]. The network operator can use the monitoring capability of the smart meters for increased visibility in LV networks for many power system applications including load forecasting, demand side management (DSM).

Traditionally, the load is forecasted at a higher aggregation level and DSM is predominantly strategized at distribution network level [4]. With the active distribution network and increased penetration of RES, system studies including load forecast, DSM needs to be conducted at lower aggregation level and instead of centralized control, distributed control is anticipated [4, 5]. This necessitates granular information about the load patterns within the network that can be extracted from the voluminous smart meter data. The large volume of the data tends to be a major hurdle for data analytics, particularly for power system studies requiring efficient and swift processing of the data. The variety of consumer patterns embedded within this big data cannot be extracted without suitable computational techniques. Apart from data volume challenge, load forecast error at lower aggregation level is high due to many contributing factors such as consumers' energy consumption behaviour, high variations in load, variety and volatility of complex load profiles. High forecast error might affect desired outcomes of the DSM by leading to improper DSM planning. Moreover, direct deployment of DSM on a large number of smart meter consumers makes scalability of such approach a major challenge. For DSM optimization, the complex load profiles at lower aggregation levels prove to be computationally expensive with the need for non-linear techniques. Additionally, the complexity of data precludes direct use of computationally efficient linear techniques for power system applications including DSM.

The evolving nature of the smart grid has changed the operational regime necessitating power system applications at different aggregation levels for management of the active distribution network. Need for granular information about the consumer energy consumption has become imperative for such system studies including load forecast and DSM at lower aggregation level. However, complex load profiles with large volumes of data, high levels of noise, high volatility and variability pose significant challenges in efficient processing and effective utilization of smart meter data for power system applications. New load profiling approaches for data abstraction and alleviating the noise, variability, volatility and complexity of data are highly desirable. Moreover, improvement in the load forecast at lower aggregation level and efficient and effective DSM planning require new research that can support efficient management of the active distribution networks.

1.2 Research Challenges

The voluminous smart meter data of a large number of electricity consumers being generated and transmitted at varying resolution and frequencies can limit the visibility into the LV network due to its sheer size. Extraction of similar patterns for load profiling from this big data is challenging as with the increased data volume and data complexity, the computational burden becomes manifold. The state-of-art research proposes load profiling of smart meter data by application of data clustering to reduce the size of data [4, 6]. However, existing clustering approaches use cluster validity indices such as Davies-Bouldin indicator, Silhouette Index etc. that are not specifically designed for power system applications [4, 7]. Use of such cluster validity indices is a big hurdle in the extraction of appropriate clusters, which can facilitate improvement in load forecasting at lower aggregation level. Moreover, the existing approaches do not explore load patterns embedded within the lower aggregation level of the smart meter data and tend to create small number of clusters with a large number of consumers. The existing approaches can hinder identification of potential consumers for DSM participation by obscuring the visibility into patterns embedded at deeper levels. Therefore, a major challenge in load profile development is to manage the size of load profiles, extract similar patterns for load forecasting and explore the data at lower levels of aggregation to make it suitable for DSM.

The load profiles at lower aggregation level tend to be highly variable and volatile, and they carry high levels of noise that is often cancelled at the higher aggregation level. The volatility and variability can significantly impact load forecast performance and consequently higher forecast error is observed [8, 9]. Moreover, the high variability of the smart meter data makes it highly non-linear that can make the processing of such data computationally expensive. This stands particularly true for applications that require linear processing of non-linear data. Decrease in the accuracy of forecast due to the noise, volatility and variability can increase uncertainty and reduce the effectiveness of DSM planning. The smart grids are expected to benefit from the DSM with active participation of consumers and prosumers, which was limited in the pre-smart grid era. The active participation of the consumers and prosumers is characterized by their energy consumption patterns. These patterns are communicated to the system operator for selection for potential participating consumers and prosumers that can participate in DSM. The system operator forecasts the future energy demand of consumers and prosumers and sends pricing signals to them to which they can respond by varying their load. However, the above discussed smart meter data challenges for load forecasting potentially lead to inaccurate load forecast and thus create uncertainties in the DSM planning. Therefore, to improve load forecast and DSM planning, these challenges need to be handled using new approaches. Issue of noise and variability also persists in RES and needs to be tackled for studies incorporating both load and generation profiles. Moreover, managing demand for a large number of consumers at the individual level creates a major scalability problem. Identification of potential consumers for DSM while considering the uncertainty of forecast is another major challenge.

1.3 Research Aim and Objectives

This research aims to develop innovative approaches to refine smart meter data-based load profiles to improve the accuracy of load forecast and DSM solutions. In order to achieve the aim, following objectives are established:

- Innovate a clustering technique to reduce the size of smart meter data and extract groups of consumers with similar patterns for load forecasting by proposing a new clustering approach and stopping indices to automate the clustering process without any external input.
- Develop a mathematical framework to refine smart meter-based load profiles by reducing the data complexity, non-linearity, variability and volatility, while maintaining the accuracy of representation by optimizing the refined profiles.
- Investigate the efficacy of proposed clustering and load profiling approach for load forecasting application considering varying levels of RES penetration to validate the accuracy and efficacy of the proposed approach.
- Develop a novel cluster selection method for DSM that can identify fewer consumers to get the desired flexibility while considering the uncertainty of load forecast with and without RES at varying demand flexibility levels.
- Investigate the impacts of the forecast generated using a new clustering and profiling approach with varying load and demand flexibility of the participating consumers for use in DSM applications providing monetary benefits and peak reduction using load shifting.

1.4 Research Contributions

The main contributions of the research presented in this thesis are summarised as follows:

- An innovative data-driven clustering algorithm for clustering profiles of consumers using non-conventional application-based cluster validity indices is proposed. The algorithm automatically selects the appropriate level of aggregation based on the number of consumers and produces clusters with high intra-cluster pattern similarity. High intra-cluster pattern similarity can benefit in benefit in many power system applications including load forecasting and DSM. The research related to this contribution has resulted in following publications [10, 11].
- A new mathematical model for a new load profiling approach is proposed. The approach linearizes and optimizes the profiles to reduce non-linearity, noise, variability and volatility of profiles. The approach generates refined load profiles that have lower noise, variability and volatility with increased linearity. The approach simplifies the complexity of data while maintaining the precision of representation close to the original data as compared to conventional linear approximation techniques. The research related to this contribution has resulted in following publications [11-13].
- The proposed load profiling approach is extended to reduce volatility and variability of intermittent RES generation profiles. The alternate RES generation profiles show higher accuracy as compared to the load profiles due to lesser variability. Application of alternative load and generation profiles in power system reliability assessment shows high level of accuracy with reduced computation time as reported in the published paper [14].

- An alternative load forecasting approach is proposed to improve accuracy of load forecast at lower aggregation level of load with reduced training time for artificial neural networks. The research related to this contribution has resulted in following publications [12],[15, 16].
- A novel cluster selection index for DSM is proposed. The index benefits in selection of appropriate clusters thus reduces the number of consumers required to participate in the DSM program. Moreover, with detailed information about energy consumption behaviour of individual clusters and consumers, incentives for DSM can be tailored to needs of specific consumers. The research related to this contribution has resulted in publication the following [16].
- A novel DSM approach is proposed to address scalability problem and selection of appropriate consumers with improved economic and technical benefits. The research related to this contribution has resulted in the following publication [16].

1.5 Thesis Layout

This thesis contains six additional chapters. Chapter 2 consists of a literature review, chapter 3-6 include the main body of research for the thesis and the final chapter concludes with the main findings of the research.

Chapter 2 presents an overview of the literature review of the different steps involved in the implementation of demand side management, starting from smart meter data collection. This literature review explores consumer classification for load profile development based on data clustering, current trends in load forecasting, and finally, demand side management.

Various clustering, forecasting and DSM techniques, with the relevant literature in the context of smart grid, are discussed to identify the future direction of this research.

Chapter 3 presents the methodology adopted to develop an innovative clustering algorithm for load profile development. It describes the new indices adopted to facilitate the extension of the existing k-means clustering algorithm. The algorithm does not require user to define the number of clusters and rather evaluates the cluster numbers depending on the number of load patterns within the data. The stopping criteria for the algorithm are determined by two indices, which are designed considering their suitability for load forecasting application. The output clusters from the clustering algorithm are used to extract average load profiles, which are developed into alternative profiles in Chapter 4.

Chapter 4 is dedicated to the development of a novel load profiling approach. A novel mathematical model is introduced for the development of the alternative profiles, which are concatenation of linear profiles. The proposed profiles are validated for use in Mote-Carlo simulation application. These alternative profiles are used for load forecasting as detailed in Chapter 5.

Chapter 5 brings together the load profile development to the load forecasting problem. This chapter proposes an alternative approach to forecast load at lower aggregation level. The impact of the alternative approach on the accuracy of load forecast is assessed by comparing the proposed approach with conventional forecasting approach using artificial neural networks. The forecast generated using both approaches is used to plan DSM with the aim to reduce peak by selecting clusters that will cause minimum consumer disruption while achieving the best possible results as demonstrated in Chapter 6.

Chapter 6 describes the development process of a new cluster selection index and its application to DSM. It proposes a novel approach for DSM using alternate load forecasting approach and new cluster selection index. Different scenarios are simulated to evaluate the benefit of the proposed approach in terms of monetary savings and peak reduction.

Chapter 7 presents the conclusions drawn from the main findings of the research presented in Chapters 3 to Chapter 6. It also points out future research using the proposed approaches.

CHAPTER 2

2 Literature Review

This chapter presents a comprehensive literature review of consumer categorization for load profile development based on data clustering, current trends in load forecasting, and demand side management. Various clustering, forecasting and demand side management techniques, with the relevant literature in the context of smart grids are reviewed critically.

2.1 Smart Meter Data

Smart meters are capable to measure, record and send granular energy consumption data to the system operator. They are considered as the first step towards the development of smart grids that will enable utilities and consumers to gather and utilize the granular information about load in a more intelligent and cost-effective manner. With rapid development in information and communications technologies (ICTs), smart meters are expected to play a key role in future smart grids by sending and receiving real time digital information regarding electricity use, electricity cost and price that can be used to implement many smart grid initiatives including dynamic pricing, demand side management (DSM) and many other programs.

Smart grid initiatives taken by the government of the United Kingdom (UK) include shift from conventional fossil fuel generation to renewable energy sources (RES) and large-scale deployment of smart meters. The UK aims to install smart meter for each individual domestic and smaller-non-domestic consumers by the end of 2020 [3]. Figure 1 shows the rapid growth in the number of smart meters installed in the UK over the last six years [3].

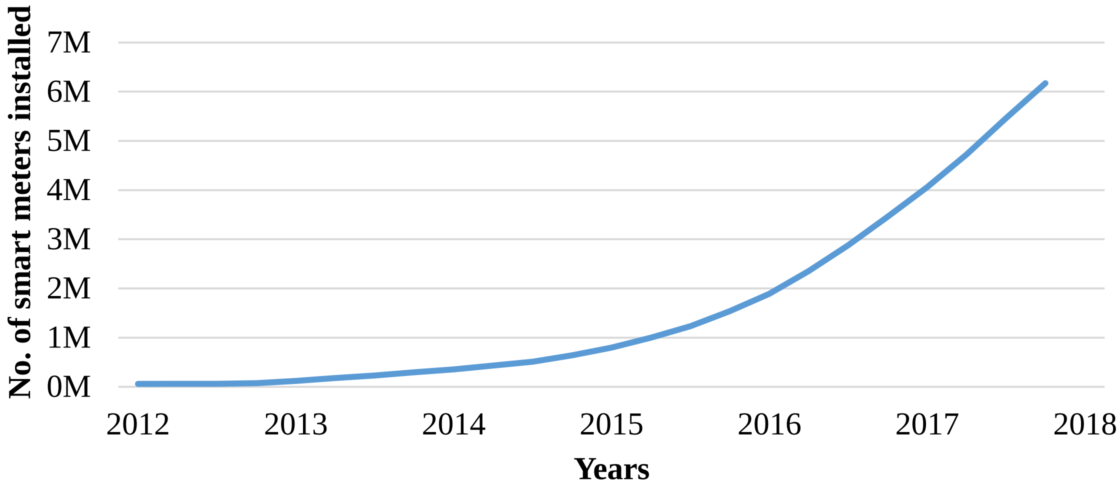


Figure 2.1 *Number of smart meters (Millions) installed in the UK from 2012 to 2018 [3]*

Smart meters are generating granular load data of millions of consumers in high volume and this data is also being delivered at high rates. Further to the data granularity, as can be seen from Figure 2.1, the rapid proliferation of smart meters is augmenting the complexity of the data challenges with every passing day. The smart meter data provides detailed energy consumption information at a consumer level, but without the transformation of this information into knowledge, its benefits cannot be exploited to their full potential as without such knowledge, implementation of smart grid programmes like DSM, distributed control etc. will not be possible.

Identification and extraction of useful knowledge from big data require the data to be properly managed for processing. The knowledge extraction process requires identification of the useful information in the smart meter data by iteratively processing the data, ensuring pre-processing of the data including data cleansing. Pre-processing of data includes re-arranging the data in a more appropriate form for faster processing and ensuring high data quality. Smart meter data provides utilities with an opportunity to improve their operational efficiency [17] by improving the forecasting of energy consumption [18, 19]. However, the improvements heavily rely on data quality, which is incomplete without data pre-processing and data cleansing.

2.1.1 Data Pre-Processing and Data Cleansing

Data acquisition is inherently prone to errors [20] and typically errors of around 5% or more can be observed in the data acquisition process [21]. Understanding the data is an important aspect in data pre-processing as it enables user to define the data quality features, which can lead to an effective data cleansing process. Having a robust understanding of data and quality factors will results in persistence in data cleansing process [20]. Visualization of data can help in understanding the data and can potentially lead to discovery of certain features of data, which can be helpful in setting the standards for data quality. Moreover, data visualization can detect inconsistencies in data.

Data cleansing is a data mining process that is used to identify and correct missing or incorrect data, reducing noise or outlier detection in the raw data [22]. It is an interactive approach, which is data driven as the standards for data quality/validity vary with data. A standard approach for data cleansing requires definition of error, identifying the errors in the data and finally correcting the errors as shown in Figure 2.2.

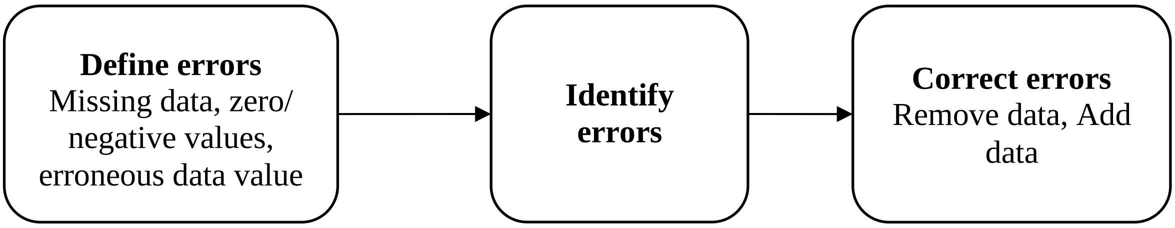


Figure 2.2 *Data cleansing process*

Smart meter data can inherently have different quality issues and inaccuracies. Depending on the application, poor quality and inaccurate smart meter data can lead to many planning and operational challenges in a power system. For example, smart meter data can lead to inaccurate forecast, prohibiting optimal asset capacity planning that can result in having assets that are under-rated leading to a failure of asset or over-rated resulting in underutilization of asset [11].

Research published on applications using smart meter data often does not incorporate the data pre-processing and data cleansing. This is probably because the data provided by utilities has already been through the stages of validation, estimation and editing (VEE). Moreover, the amount of data easily available to researchers is often limited. The real-world smart meter data is highly susceptible to noise, inconsistencies and missing values, which can significantly influence results of the data analysis [9].

Typical issues of smart meter data include data inconsistencies (e.g. inconsistent measuring units switching from W to kW) [23], duplicate data (non-identical or identical due to different data resolution) [24], zero and missing values (due to communication or other equipment failure or storm) [24] and outliers. Inconsistencies are dealt with by making the data consistent, duplicates are removed to keep only one data. Presence of zeros in smart meter data is usually not expected due to proliferation of electronic equipment. These equipment stay connected to the power source in standby mode and according to reference [24], 5-10% of the total residential electricity use is comprised of the standby power. Therefore, the group having highest number of zeros is usually considered to have incomplete data since bad data can potentially skew the load profile [25]. If the missing value or zero series is not too long, data can be estimated using different methods discussed in [26]. It is important to select the appropriate estimation method otherwise wrong estimation can lead to bad data quality. Smart meters carrying missing values for long duration are often removed from the data to ensure high data quality [25].

The goal of outlier detection is to find such exceptions in the datasets, which indicate incorrect values [20]. Three fundamental approaches to the problem of outlier detection include outlier detection with no prior knowledge, modelling both the normalities and abnormalities, and modelling the normalities only [27]. Based on the three approaches, there are many candidate methods to detect outlier including clustering [28, 29], pattern-based, statistical [30] or neural network based [27]. Modelling normalities and abnormalities involve high modelling intensity for a large number of smart meter consumers. As outliers are considered to be abnormal values or values that deviate markedly from other samples, once outliers are detected, they are removed from the dataset. Statistically, outliers are numerically distant from the rest of the values. Data clustering can detect the outliers by identifying the consumers who are distant from the centroids. Authors of [31, 32] used data clustering for outlier detection, which is one of commonly used approach for outlier detection.

2.2 Load Profiling Using Smart Meter Data

Smart meters record the electric energy consumptions periodically at each connection point across the power grid, e.g. residential and commercial consumers, industries, locomotives etc. The plot of the recorded smart meter load data shows the load curve. With the variations in these curves, the electricity supplier and network must respond to the consumer’s power demand. Therefore, different operation and planning activities such as load forecasting, DSM, etc. essentially require time series information of load preferably for each class of load [4].

Prior to smart grids, the load classification was generally limited to higher aggregation level of load. However, with advent of smart meters, the classification has become diverse in nature and each class can potentially be divided into sub-classes based on its energy consumption patterns. The classification is carried out by grouping consumers with similar load patterns to develop typical load profiles. A load profile represents variation in electrical load against time. According to [4, 33], *"Load profiling refers to the classification of load curves or consumers according to electricity consumption behaviors"*. Future smart grids tend to use detailed information of load at different hierarchical level of power grid and this is achievable by embedding behaviour of consumers in profiling at distribution level for applications such as load forecasting, DSM etc. [34].

2.2.1 Applications of Smart Meter-Based Load Profiling

Load profiles can be used for different applications in a power system depending on the hierarchical level of load profile. A brief discussion on application of load profiling in load forecasting and DSM is given below;

- **Load Forecasting**

Historically, power distribution network level load forecasting was performed using historic loads and trending method to see the future trends for prediction [35]. Use of clustering the profiles in forecasting using trending method is given in [36]. However, the accuracy of the trending method is significantly low due to unavailability of detailed consumption records. On the other hand, forecasting in the era of smart meter uses multivariate methods, which include many external factors apart from the load profiles [37]. Relationship between the load and other factors such as temperature, time of day, previous hour load etc. is formulated using different statistical or artificial intelligence techniques [38]. However, without the granular data of smart meters, mapping relationship between different factors and load was not possible with high precision but, the smart meter data provides an opportunity to improve forecast by understanding the energy consumption behaviour of consumers even at an individual level. A detailed discussion about load forecasting is given in the section 2.3.

- **Demand Side Management (DSM)**

Consumer participation in DSM program is decided based upon their load profile [39]. The utility offers incentives to consumer based on their load profiles by evaluating the benefit for utility as well as consumer. The benefits for utility come in the shape of deferral in infrastructure and savings by avoiding use of Peaker plants whereas, consumers benefit from savings in electricity bills. It is pertinent to mention that the load profiles used for DSM at consumer level are estimated profiles or forecasted profiles and without detailed knowledge of the consumer, accurate prediction of load and selection of appropriate DSM program is not feasible [4]. Smart meter data provides the necessary granular information about the load. Thus, DSM can benefit from smart meter data indirectly by improved certainty in forecast and selection of appropriate consumers from their load profiles developed using smart meter data. Detailed discussion regarding DSM is given in section 2.4.

2.2.2 Load Profiling Using Data Clustering

The process of consumer characterization in smart meters is priori unknown. A number of classes in smart meter data depends on the number of significant patterns within the smart meter data and thus unsupervised learning approaches suit the purpose of consumer classification for load profiling. A commonly used unsupervised machine learning technique that can determine the intrinsic groups in an unlabelled dataset is called data clustering. It does not require any prior knowledge and can group consumers based on similarity of their load curves [40]. It categorizes consumers based on their energy consumption behaviour.

Data clustering is broadly classified into two types i.e. hard and soft clustering. Hard clustering or crisp clustering refers to the fact that a data point can belong to only one cluster whereas, soft clustering allows probabilistic allocation of the data points to the clusters [41, 42]. Some other forms of classification of data clustering classify clustering methods as partitional clustering, hierarchical clustering, density-based clustering, mixture model clustering and spectral clustering [43]. However, they all fall into broader categories of hard and soft clustering. Figure 2.3 shows classification of clustering algorithms based on soft and hard clustering [43].

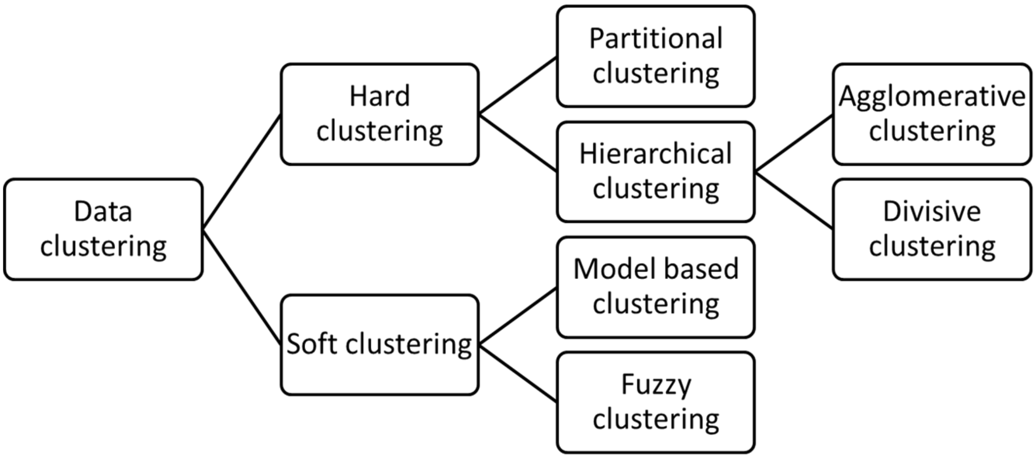


Figure 2.3 *Clustering method classification [43]*

The partitional and hierarchical clustering are most commonly used clustering methods in load profiling. The clustering algorithms tend to capture the energy consumption behaviour of the consumers for load profiling, which means capturing the similarities in the shapes is important. Selection of appropriate proximity/distance measure, which can quantify similarity or dissimilarity can impact the clustering solution [42]. Brief description of proximity measures is given below.

2.2.3 Data Clustering

A dataset 'D'[1] having 'm' objects[2] with each object having 'n' features[3] can be denoted by (2.1);

$$D = \{x_1, x_2, \dots, x_i, \dots, x_m\} \tag{2.1}$$

$$x_i = \{x_{i1}, x_{i2}, \dots, x_{ik}, \dots, x_{in}\} \tag{2.2}$$

where x_{ik} represents k_{th} component of the feature of object or smart meter x_i [43]. The entire dataset to be clustered is organized into a $m \times n$ matrix where each row represents object and each column denotes features. Organization of data in matric form enables the computation to be efficient for data cleansing and clustering both.

- **Proximity Measures**

The intended purpose of a clustering algorithm for load profiling is to group objects having similar shapes and this requires a notion of similarity or dissimilarity measure. Such a measure is usually termed as proximity measure or distance measure [28]. If the objects are closer, the distance between them will be lower, which indicates higher similarity and lower dissimilarity [44]. Mathematically distance between two data points x_i and x_j is denoted as $d(x_i, x_j)$.

[1] Data set 'D' refers to data containing records of all smart meters or patterns of energy consumption

[2] Object refers to each individual smart meter records

[3] Features are attributes or dimension of each energy pattern

Different distance measures are used to determine the dissimilarity in energy consumption patterns. Some of the commonly used distance measures are given in Table 2.1. For two smart meter consumers $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ and $x_j = (x_{j1}, x_{j2}, \dots, x_{jn})$, the well-known Euclidean distance matric can be calculated as below [23];

$$d(x_i, x_j) = (\sum_{k=1}^{n} (x_{ik} - x_{jk})^2)^{1/2} \qquad (2.3)$$

Table 2.1 *Distance measures for clustering [23]*

DISTANCE MEASURE	MATHEMATICAL FORMULATION
Euclidean	$d_{euc}(x_i, x_j) = \left(\sum_{k=1}^{n} (x_{ik} - x_{jk})^2\right)^{\frac{1}{2}}$
Squared Euclidean	$d_{seuc}(x_i, x_j) = \sum_{k=1}^{n} (x_{ik} - x_{jk})^2$
Manhattan	$d_{man}(x_i, x_j) = \sum_{k=0}^{n} \lvert x_{ik} - x_{jk} \rvert$
Mahalnobis	$d_{mahb}(x_i, x_j) = \left[(x_i - x_j) \sum^{-1} (x_i - x_j)^T\right]^{\frac{1}{2}}$
Cosine	$\mathrm{d}_{\cos}(x_i, x_j) = 1 - \frac{\sum_{k=1}^{N} x_{ik} x_{jk}}{\left(\sum(k = 1)^n x_{ik}^2 \sum_{k=1}^{n} x_{jk}^2\right)^{\frac{1}{2}}}$
Minkowski	$d_{mink}(x_i, x_j) = \left(\sum_{k=1}^{n} \lvert x_{ik} - x_{jk} \rvert^p\right)^{\frac{1}{p}}, p \geq 1$
Chebychev	$d_{chev} = \lim_{1 \leq k \leq n} \lvert x_{ik} - x_{jk} \rvert$
Canberra	$d_{can}(x_i, x_j) = \sum_{k=1}^{n} \frac{\lvert x_{ik} - x_{jk} \rvert}{\lvert x_{ik} \rvert + \lvert x_{jk} \rvert}$

where n is number of attributes; x_{ik} is k_{th} attribute of feature vector x_i; x_{jk} is k_{th}attribute of feature vector x_j;p is order of Minkowski distance.

2.2.4 Clustering Algorithms

A number of clustering algorithms are used in different applications, selection of algorithm depends on multiple factors including data to be clustered and definition of meaningful cluster [45, 46]. Some of the most commonly used hard clustering algorithms are discussed below;

2.2.4.1 Partitional Clustering

Partitional clustering divides the data into n number of partitions. The partitional methods are often used for clustering of patterns. The partitions resulting from the partitional clustering should have following characteristics [47];

- Each cluster should be assigned at least one pattern i.e. $C_i \neq \emptyset, \forall \text{ i} \in \{1,2,3, \dots, \text{k})$
- Two different clusters should have no identical pattern i.e.

  $C_i \cap C_j = \emptyset, \forall\, i \neq j, i, j \in \{1,2,3, \dots, \text{k})$

- Each pattern should be assigned to a cluster

Partitional techniques tend to optimize the criterion function locally or globally in order to minimize the error [48]. K-means clustering is one such partitional clustering algorithm, which optimizes the error function. It is one of the most widely used and applied clustering algorithm.

2.2.4.1.1 k-means Clustering Algorithm

k-means clustering is a simple and robust clustering algorithm. It assigns each pattern/data point to the nearest cluster centre, also called centroid. Clustering process starts by defining the number of centroids 'k' and these centroids are allocated initial positions randomly. Data points nearest to each centroid are allocated to the centroid and are called members of the cluster. To allocate precise position to the centroid, the positions of centroids are recomputed to minimize the distances between the cluster members and the centroids. After each iteration the distance from each cluster member to the centroids is measured and as each data point chooses to be member of the cluster with shortest distance to its centroid, recalculation of centroids position results in change in the number of cluster members. Centroids are recomputed for each iteration until cluster membership becomes consistent and no change in membership occurs for more iterations. The k-means clustering can be mathematically modelled as below [28, 49].

For a dataset 'D' having 'n' attributes, let $c_1, c_2, \dots, c_k$ be the k disjoint clusters of the dataset. The error function can be formulated as in (2.4) [28, 49].

$$E = \sum_{i=1}^{k} \sum_{x \in c_i} d(x, \mu(C_i)) \tag{2.4}$$

where $\mu(C_i)$ is the centroid of cluster C_i; $d\big(x, \mu(C_i)\big)$ shows the distance between data point x and the centroid of cluster i. The distance can be selected based upon the application.

Some of the typical issues faced by k-means clustering include convergence to local minima. This can be avoided by repeating the k-means for different starting positions of the centroids. Moreover, selection of k i.e. number of clusters is an important issue that is often contentious. One of the well-known attempt to resolve the selection of 'k' is x-means clustering [50] but the computation involved makes it complex and expensive method. As discussed earlier, selection of number of clusters is governed by nature of datasets and use case of the clusters. Thus, different domains require devising new domain-based methods for cluster selection. Power system application-based matrices for selection of number of clusters in k-means have not been explored. Research in load profiling of electricity consumers should focus on the application of forecast to develop matrices for determination of number of clusters.

2.2.4.1.2 Hierarchical Clustering

Hierarchical clustering is different from partitional clustering as it hierarchically divides dataset into clusters. It generates a hierarchical decomposition of a set of N elements represented by a dendrogram. The dendrogram shows the possible set of clusters, which can be selected by user based on the level of similarity required [42].

Hierarchical clustering can be agglomerative (bottom-up) or divisive (top-down). An agglomerative algorithm begins with each element of dataset initially being a separate cluster and merges them into successively larger clusters. Whereas, a divisive algorithm begins with the entire dataset being one cluster and successively divides them into smaller clusters [47]. The approach adopted in the divisive algorithm resembles the classification process in the human brain [51]. The computational complexity of divisive clustering algorithms is high as for a cluster having n data points, it require to compute $(2n - 1)$ possible divisions [52]. Thus, the computational complexity of divisive algorithm often makes the agglomerative technique preferable over divisive hierarchical algorithm.

The agglomerative algorithm starts merging 'n' clusters until all clusters are merged into a single cluster. The merging takes place based on the similarity calculated from the similarity matrix, and the clusters with the least distance between them are clustered first. Different similarity measures commonly used for hierarchical clustering include the following:

a. Single-link Clustering

This method considers the shortest distance between two closest members of the two clusters as the distance between these clusters. Single-link clustering is also known as connectedness, the minimum method or the nearest neighbour method [47].

b. Complete-link Clustering

Also known as the maximum clustering method or the furthest neighbour method, defines distance between two clusters as the distance between the member of the clusters, which are located at longest distance from any member of the other cluster [47].

c. Average-link Clustering

Average-link or minimum variance method considers the distance between two clusters as the average distance from any member of cluster to any member of the other cluster [47].

Key advantages of the hierarchical clustering are no prior requirement to specify number of clusters and independence from initial conditions (centroid position). However, they are oblivious of the global shape or size of clusters, which is a key drawback for them [47]. Moreover, hierarchical clustering is high on time complexity as compared to partitional clustering [47].

Application of any clustering algorithm needs to be verified for its efficacy in providing the required solution. A discussion on cluster validity measures is presented below.

2.2.5 Cluster Validity

The clustering process is unsupervised type of learning, which implies that no predefined classes exist [53]. Thus, cluster validity process is used to evaluate goodness of clustering solution. Cluster validation can benefit the clustering process in many ways, such as determining the clustering tendency of a set of data, comparing results of clustering to externally known results, evaluating fitting of the solution to data, determining a correct number of clusters etc. [54].

By definition, an ideal cluster is a set of data points that is compact and isolated [45]. This means that a clustering solution should generate clusters that are compact and separate from each other. For clustering energy consumption data, this can be considered as the consumers within one cluster should have similar energy consumption behaviour and consumers of different clusters should have different energy consumption behaviour. This can be termed as high intra-cluster pattern similarity and high inter-cluster pattern dissimilarity, and these can be used to formulate the cluster validation criteria. However, defining limits of acceptable pattern similarity depends on the application of clustered data.

Performance of clustering solution for 2D-data sets is usually evaluated visually to verify the validate the results [54]. Evaluation of clustering algorithm results is commonly carried out using cluster validity indices. In general, there are three criteria that are used to evaluate cluster validity which are external criteria, internal criteria and relative criteria [55]. The criteria and some of the indices developed from the criteria are discussed below;

2.2.5.1 External Criteria

In the external criterion, the evaluation of the clustering algorithm is based on the assumption of pre-specified structure. This is usually used to measure the extent to which the cluster labels match the known labels [55].

2.2.5.2 Internal Criteria

The internal criteria evaluates the goodness of clustering solution for the quantities, which involve data e.g. proximity matrix [54]. Most commonly used internal criteria is Cophenetic correlation coefficient (CPCC), which is used for hierarchical clustering structures [43]. For a proximity matrix $P = \{p_{ij}\}$ of dataset D, CPCC measures the degree of similarity between proximity matrix P and the cophenetic matrix $Q = \{q_{ij}\}$, whose elements record the proximity level [43]. Let μ_P and μ_q be the means of P and Q then CPCC can be given as below [43];

$$\mu_P = \frac{1}{M}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} P_{ij} \tag{2.5}$$

$$\mu_q = \frac{1}{M}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} q_{ij} \tag{2.6}$$

$$CPCC = \frac{\frac{1}{M}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} p_{ij}q_{ij} - \mu_P\mu_Q}{\sqrt{\left(\frac{1}{M}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} P_{ij}^2 - \mu_P^2\right)\left(\frac{1}{M}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} q_{ij}^2 - \mu_Q^2\right)}} \tag{2.7}$$

where $\mathrm{M} = \frac{\mathrm{N(N-1)}}{2}$ and the value of CPCC ranges between -1 and 1. Index values close to 1 indicate a high similarity between P and Q.

2.2.5.3 Relative criteria

Relative criteria compares clustering results of different algorithms or same algorithm with different input parameters for cluster validation [43]. For hierarchical clustering algorithms, number of clusters is decided by determining the cutting point of dendrogram whereas, for partitional clustering, the number of clusters is decided by user expertise or a priori information for application. Having too many clusters can lead to difficulty in interpretation of the clustering results and having too few clusters can result in loss of information. In the following section, some of the commonly used methods and relative criteria indices for hard clustering are discussed.

2.2.5.3.1 Data Visualization

Data visualization can be considered as the most direct method to estimate the number of clusters. Visual inspection can provide some useful insight on number of clusters but the complexity of data can pose challenges, which can be resolved using dimension reduction techniques as discussed in [43] or by using small experimental data for visualization to ascertain the efficacy of clustering [54].

2.2.5.3.2 Validation Indices and Stopping Rules

For partitional clustering, a sequence of clustering structures can be obtained by simulating different number of clusters (*k*) by running clustering algorithm many times. Different indices can be used to evaluate best value of *k*. For hierarchical clustering, stopping rules are used to determine the best level for cutting the dendrogram. These stopping rules are functions of certain factors such as the defined square error, the geometric or statistical properties of the data, number of objects in cluster, dissimilarity or similarity measure etc. [54].

A study by Milligan and Cooper [56] compared 30 indices and identified Caliński and Harabasz index [57] as one of the best performing indices. This index can be mathematically defined as [43];

$$CH(K) = \frac{Tr(\boldsymbol{S_B})}{K-1} \Big/ \frac{Tr(\boldsymbol{S_w})}{N-k} \qquad (2.8)$$

where N is the number of objects, Tr(**S_B**) and Tr(**S_W**) represent the inter and intra-cluster scatter mix respectively. Best value of *k* is one which maximizes the CH(K).

Another index called Davies-Bouldin (DB) index, attempts to maximize the inter-cluster distance while minimizing the intra-cluster distance. The index calculates cluster index for each cluster and tends to minimize the mean of the cluster index. The individual cluster index R_i is given as [43];

$$R_i = \max_{j \neq 1} \left(\frac{e_i + e_j}{D_{ij}} \right) \qquad (2.9)$$

where $D_{ij}$is the distance between centroids of cluster i and j respectively and $e_i and\ e_j$ are the average errors for cluster $i\ and\ j$ respectively. The DB index can be given as [43];

$$DB(K) = \frac{1}{K} \sum_{i=1}^{k} R_i \qquad (2.10)$$

The minimum value of DB(K) can be the potential optimum number of clusters in the data.

The Dunn index [58] defines the distance between two cluster $C_i and\ C_j$ as minimum distance between points x and y belonging to the cluster $C_i and\ C_j$ respectively (as in single-linkage algorithm).

$$\mathrm{D}\left(C_i, C_j\right) = \min_{x \in C_i, y \in C_j} D(x, y) \qquad (2.11)$$

The maximum distance between two members of same cluster is given as;

$$diam(C_i) = \max_{x,y \in C_i} D(x, y) \qquad (2.12)$$

Based on the above given intra and inter-cluster similarity/dissimilarity measures, the Dunn index can be formulated as below [43];

$$Du(K) = \min_{i=1,\dots,k}\left[\min_{i=1,\dots,k}\left[\frac{D(C_i, C_j)}{\max\limits_{l=1,\dots,k}(C_l)}\right]\right] \tag{2.13}$$

The higher value of $Du(K)$ points to compact and well-separated clusters, which can be used to estimate best value of K.

There are many other indices and stopping rules which are given in the literature, however, it is pertinent to mention that the indices are data dependent and the selection of the best number of cluster requires input from the expert in the area of application [41].

Although the cluster validity indices discussed above are important in ascertaining the validity of a cluster, however number of clusters depends on the application. The techniques discussed above are well established techniques and are widely used in power system applications, however, validation of the clusters needs a new approach. The future of the research tends to develop more automated, data driven algorithms which can derive clusters without intervention of any expert and are validated based on application of the clustered data [4]. Therefore, application oriented matrices are to present more benefits in selection of clustering algorithms [4].

2.3 Load Forecasting

Load forecasting is an essential part of efficient power system planning and operation [59]. As the relationship between load and the factors affecting the load is non-linear and in particular with the new kinds of loads in smart grid environment, the classical estimation methods for load forecasting are not as effective as they used to be. In recent years, significant research interest has been observed in the load forecasting with new forecasting techniques to tackle emerging issues.

2.3.1 Load Forecasting Horizons and Techniques

Load forecasting can be divided into four periods based on the forecast horizons. There is no rule of thumb for classification of the forecast horizons of the four periods [60]. However, in literature the classification of the forecast is presented in terms of four forecast horizons i.e. very short-term load forecast (VSTLF), short-term load forecast (STLF), medium term load forecast (MTLF) and long-term load forecast (LTLF). VSTLF, STLF, MTLF and LTLF account for the time from seconds/minutes to several hours, from hours to weeks, from months to years and from a minimum of a year to several years respectively [60-62].

With the difference in the forecast horizon, applications of the forecast vary accordingly. Due to the forecast in near real-time, VSTLF models find their application in real-time grid management [15]. STLF models can be used for demand and generation adjustment and other power system operation applications, whereas MTLF and LTLF models are widely used for asset management, asset utilisation planning and network expansion planning [61]. Reference [59] has classified different areas of application of forecasts as given in Table 2.2.

Table 2.2 Forecasts and Applications **[59]**

APPLICATION	*VSTLF*	*STLF*	*MTLF*	*LTLF*
Energy purchasing	Yes	Yes	Yes	Yes
T&D planning	No	Yes	Yes	Yes
Operations	Yes	Yes	No	No
DSM	Yes	Yes	Yes	Yes
Financial planning	No	No	Yes	Yes

STLF and VSTLF are very important from the operational point of view. Moreover, from Table 2.2, it is evident that STLF is not limited to operations only, it also involves resource utilization i.e. economic load dispatch, energy purchasing and DSM. As the STLF covers the application areas of VSTL, often literature considers VSTLF as part of STLF rather than a separate forecast horizon.

Classic load forecasting is carried out using statistical and mathematical models whereas, new methods incorporate artificial intelligence (AI) techniques. The statistical models include regression, multiple linear regressions, stochastic time series, general exponential smoothing, and kalman filter etc. [15]. One of the simplest statistical forecasting techniques is the multiple linear regression (MLR) model [63]. In this model, the load is calculated in terms of the explanatory variables such as temperature or non-temperature dependent variables, which affect the load. The mathematical formulation of a MLR model can be given as in (2.14) [63].

$$y(t) = a_0 + a_1 x_{i1} + a_2 x_{i2+} \dots + a_{n-1} x_{i,n-1} + e_i \qquad (2.14)$$

where $y(t)$ is the load, $x_1, \dots, x_n$ are known constants, e_i is a normally distributed random variable and $a_0, a_1, \dots, a_n$ are parameters.

The AI based techniques include artificial neural networks (ANNs), fuzzy logic, expert system, grey system and many hybrids [64]. A typical static feedforward neural network model for load forecasting can be represented as below [65].

$$L_{t+1} = f(t, L_t, L_{t-1}, \dots, L_{t-n}, w_t, w_{t-1}, \dots, w_{t-r}, \widehat{w}_{t+l)+} \, \varepsilon_{t+l} \qquad (2.15)$$

where t is the time of day, l is the time lead for forecast, L_t is the load at time t, $w_t$is a vector of the weather factors observed at time t, $\widehat{w}_{t+1}$ is the weather forecast for $t+l$, and ε_{t+l} represents a random load component [65].

The statistical methods are considered to be more conventional and certain level of accuracy can be achieved using statistical methods with limited training data. Limitation of training data particularly affects the forecasting accuracy of AI based methods. However, if provided sufficient training data, AI based forecasting methods like ANN show improvement in accuracy as compared to the statistical techniques [64]. The issue of data is gradually decreasing with the increased penetration of smart meters. Authors of [66] have presented a comprehensive literature review on load forecasting. Some of the relevant literature is presented below.

A commonly used measure for load forecast error is the mean absolute percentage error (MAPE) [67]. Different studies have reported different MAPE for load forecasting study using different forecasting techniques. For example, authors of [68] have compared autoregressive integrated moving average (ARIMA), ANNs and adaptive neuro-fuzzy system techniques for VSTLF in Andrandina area in Brazil and found the MAPE as 5.07%, 1.25% and 10.21% respectively. The ANNs showed higher accuracy due to their ability to learn the complex patterns to map the functions. The result shows that ANNs outperforms ARIMA, and the adaptive neuro-fuzzy system had highest MAPE. Daneshi [69] developed a model of load forecasting using ANNs and compared it with auto regressive model. Auto regressive model proved to be less accurate as compared to ANNs. Many other researchers performed similar comparisons; however, it is difficult to compare results of any research to others since there is no standard test data set for a reasonable comparison. Reference [59] suggests that as ANN models are developed for a specific utility, their performance cannot be compared with others as it will not be meaningful. Moreover, there is no global model for load forecasting which can be applicable to all utilities with similar forecast accuracy, as the sensitivity factors of every forecast varies according to geography, weather conditions and econometric factors.

An overwhelming majority of the present literature proposes that the ANNs are generally performing better than the conventional statistical techniques and research on load forecasting using ANN is dominating the future forecasting research [4].

2.3.2 Load Forecasting Using Smart Meter Data

Although several methods and models are reported in the literature to forecast the electric power consumption, most of these models are based upon aggregated electricity consumption data without any classification of consumers. In the era of smart grids, the smart meter data provides multi-fold value to load forecasting. It enables utility to forecast load even at individual consumer level, at groups of consumers and at higher aggregation levels. The high granularity of smart meter enables the forecasting model to embed behaviour of the electricity consumption at individual consumer level.

The smart meter data provides an opportunity to realize the concept of smart grid where more localized forecast in the presence of intermittent RES is highly desirable. However, in a distribution network at lower aggregation level, there are many challenges which need new approaches for load forecasting [4].

Authors of [70] evaluated seven forecasting techniques using load data recorded every 15 minutes and concluded that neural network-based methods and support vector machines perform best for forecasting. Reference [34] used smart meter data to improve the load forecast accuracy by using k-means clustering to group consumer based on behaviour similarity. Authors of [71] used smart meter data of two hundred and twenty thousand residential consumers to develop segmentation schemes and used adaptive k-means clustering to identify representative load profiles. Authors of [72] identified smart meter data volatility due to high variability as a major problem in load forecasting. They [72] proposed a short-term memory-based deep-learning forecasting framework to address the problem. Thus, new research avenues are being explored to forecast load using smart meter data to tackle the challenges of size and variability and volatility of data.

The literature has shown that smart meter data can improve the forecast accuracy by using the clustering algorithms [34, 73]. However, the improvement in forecasting has a direct relation with the level of load aggregation [8]. At lower level of load aggregation (adding load of small number of consumers), noise, volatility and variability of data tends to be high which has an adverse impact on forecast accuracy. Importance of level of aggregation becomes paramount when the data is used for applications like DSM.

From the literature, some of the identified forecast challenges using smart meter data include the noise, volatility and high variability of the smart meter data, size of data and determining the right aggregation level of load. New data handling and mining techniques need to be developed to address these challenges such that the data can be used for load forecasting and other smart grid applications.

2.3.3 Artificial Neural Network-Based Forecasting

ANN is one of the artificial intelligence techniques that mimic the functionality of a human brain. ANNs are able to approximate the complex, hidden structures and data functions. They are considered universal approximator for any continuous function. As discussed above, the ANNs are proving to perform better for load forecasting particularly with smart meter data. They can be considered multivariate, nonlinear and non-parametric methods that can map the relationship between a given sample of input and output vectors.

Designing the ANNs based forecasting model requires selection of appropriate ANN architecture and according to [23] feed-forward multilayer perceptron (MLP) is one of such architectures that provides satisfactory results for load forecasting purpose. Typically, MLP have at least three layers, i.e. input layer, hidden layer and output layer. The number of hidden layers can be increased based on the performance of the neural network. Neural network tries to optimize the weights of the connections between the layers to minimize the error during the training of neural network. The ability of a neural network to learn the optimum values of the parameters during the training process enables it to generate higher level of forecast accuracy as compared to traditional statistical techniques [64]. However, the accuracy of the neural network depends on selection of learning algorithm and activation function.

An activation function transfers the output of one neuron of one layer as input of next layer neuron (depending on the network architecture) by limiting the output between 0 and 1 or between -1 and 1etc. (depending on the activation function). For a long time, the research around ANNs stalled because of difficulties involving non-linear activation function. This was primarily due to the fact that an activation function enables ANNs to map the non-linearity. The sigmoidal function is one of the most famous activation functions but tangent hyperbolic function has quickly taken its place [74]. With development of deep neural networks, new activation functions such as softmax, ReLU, leaky ReLU, SELU etc. are also becoming popular, however, deep neural networks are commonly applied in image recognition and other such areas

[74]. The Tanh function has a gradient stronger than sigmoid is commonly used for feed forward neural networks [75]. A Tanh function g for an input u can be given as in (2.16) [74];

$$g = tanh\left(\frac{u}{2}\right) = \frac{2}{1 + e^{-u}} - 1 \tag{2.16}$$

ANNs commonly use error backpropagation (EBP) algorithm to train the network. EBP propagates the error back to learn and improve the parameters. Propagation of error is carried out through a learning algorithm.

Selection of an appropriate learning algorithm is another key step in developing the ANNs based model for load forecasting. Commonly used learning algorithms include gradient descent, gradient descent with momentum [76], Newton's method [74], Quasi-Newton algorithm [77], Levenberg-Marquardt (LM) algorithm [78], etc.. Gradient descent is a simple and efficient algorithm which has the ability to work on the graphical processing units (GPU) which can enhance the training speed of the neural network. As discussed above, the learning in neural networks is commonly carried out using EBP algorithm. Where the gradient of the cost function or the error is calculated and propagated back to the first hidden layer [12]. The weights and biases of the neural network are adjusted in accordance to the deltas calculated from the gradients [12]. The cost function used for load forecasting in neural networks is sum of square of errors.

Selection of the appropriate input variables can significantly improve forecast accuracy. According to authors of [79], a highly significant term on its own does not necessarily generate a good forecast and it is the combination of input variables that can improve the forecast. Therefore, selection of input variables can be determined by checking sensitivity of combination of each variable. Two different approaches can be adopted in selection of variables i.e. an incremental approach or decremental approach. According to the incremental approach, a model can be started with temperature as initial variable and then rest of the variables can be added incrementally [79]. The improvements by addition of each variable can help in deciding the best combination of variables. On the other hand, a decremental approach works in opposite manner by starting with a full model. The predictive capacity of each variable is tested by dropping it from the model [79]. Thus, least significant variables can be removed from the model by adopting a decremental approach.

According to [66], the main difficulty in using the linear models for load forecasting is interpretation of the random nature of demand and its representation through mathematical equation in the model. ANNs' ability to generalize and to detect inherent non-linearities of demand results in higher forecasting accuracy as compared to linear statistical models. Availability of large amount of energy consumption data, high efficiency, small amount of time needed to set up the system and better performance of ANNs, makes them suitable for future forecasting applications. A suitable architecture of ANNs can improve the forecast accuracy provided it is supplied with sufficient training data. The smart meter data can be used to forecast the load at lower hierarchical levels of power system, which can enhance the forecast accuracy by predicting for the consumers or group of consumers who do not follow the averaged system profile.

2.4 Demand Side Management

Load forecasting provides utility with the estimation of the peak of load. If the demand increases beyond a certain level, the utility is required to turn on the peaking plants. The peaking plants are commonly of non-renewable (with exception of hydroelectric) type which are often expensive to operate and high in carbon emissions. Integration of new technologies such as electrical vehicles, heat pumps, increasing level of load and growing penetration of RES has made demand management complex due to the new uncertainties [4].

An attractive alternative to meet the increasing demand can be incentivizing the energy consumers with monetary benefits to shift their load from peak hours to off-peak hours. This needs to be managed using bi-directional communication to optimize utilization of existing generation capacity. This leads to definition of demand side management (DSM) which can be referred to as any action taken to improve the energy system at consumption side [80].

Before advent of advanced metering infrastructure (AMI) and modern grid information and communication technologies (ICTs), the DSM was more ‘utility driven’, however, present day DSM is gradually becoming ‘consumer driven’ [80]. This is mainly due to the distributed generation of RES which has made the power system load driven that was generation driven in past. The ICT in smart grids enables the operator to monitor the operational conditions in real-time and to respond appropriately to ensure safe grid operations.

DSM in the power industry mainly consists of load monitoring, analysis and response [81]. Load monitoring and analysis can provide significant benefit in the efficient resource utilization and response can help in increasing the reliability of smart grid. According to Faruqi et al. [82], five to eight percent of the installed generation capacity in Europe only handles peak, which occurs only one percent of time. Further, their [82] evaluation reveals that if this peak can be deferred to off-peak hours, savings in the capacity and transmission cost could be as much as 67 billion euros. Integration of modern technologies enables grid to monitor and analyse load and proactively respond to the arising system events with socio-economic benefits such as carbon emission reduction, renewable energy sources integration, increased grid and consumer saving etc.

2.4.1 Smart Grid Communication

Communication infrastructure plays an important role in implementation of DSM. The information and communication network in smart grids consist of different hierarchical networks. A home area network (HAN) communicates with appliances within a house whereas, a neighbour-hood area network (NAN) that communicates with HANs. Finally a wide area network (WAN) transmits metering data from all lower hierarchies to the central control centre [83]. Usually the WAN can cover area of tens of kilometres and requires high capacity transmission system.

Different levels of communication network can employ different kind of communication technologies. Usually, the HAN considers Wifi, ZigBee and Bluetooth technologies due to short range of network i.e. within home and low volume of data to be transmitted [81, 84]. NAN has a different set of candidate technologies due to higher coverage area, high volume of data and cyber security issues. The candidate technologies include WiMAX, LTE cellular network and IEEE 802.22 broadband wireless regional network [85]. The WAN having to handle high volume of data as well as maintaining high level of security, usually considers LTE wireless network, fibre optic links and power line carrier (PLC) as high transmission capacity candidate networks [83].

2.4.2 Optimization of Consumption Using Demand Side Management

Optimization of consumption using DSM is an important feature to control the peak of the load in a smart power grid. The important decision that needs to be made by the operator is that, what kind of modifications are required in the load shape to handle the peak? Load shape can be varied considering different objectives using different techniques for example peak clipping, load shifting, strategic load growth etc. [80]. Figure 2.4 shows different commonly used DSM techniques.

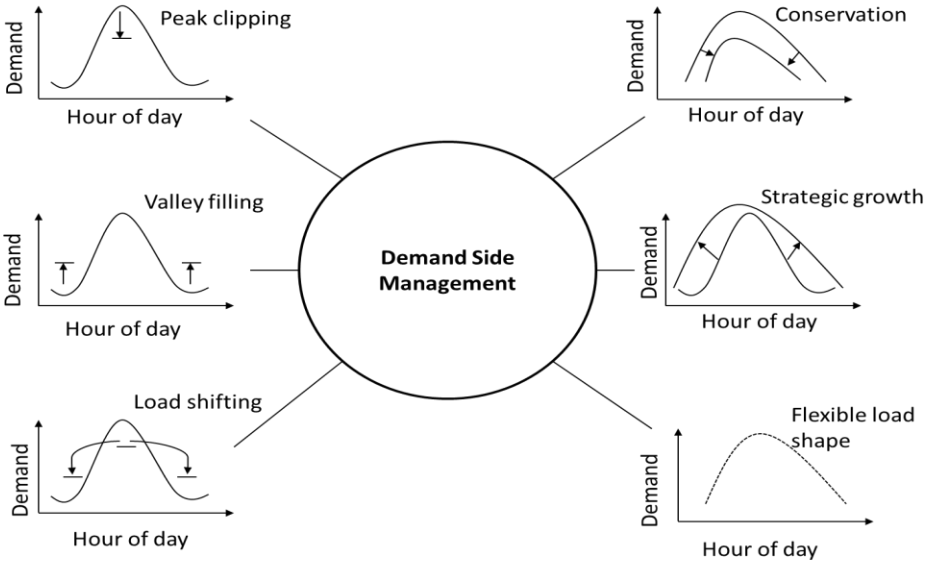


Figure 2.4 *Demand side management objectives*

As it can be seen from the Figure, peak clipping tends to curtail the load demand at specific intervals i.e. peak hours, valley filling tries to increase the load during off-peak hours and load shifting shifts the load from peak hours to off-peak hours [64]. The conservation objective reduces the overall energy demand, strategic growth increases the load and flexible load shape objective makes the load shape flexible as required by the utility [64]. The objective of peak clipping and valley filling is to reduce the peak of load and load shifting can combine benefits of both by shifting the peak load to valleys by shifting load from peak hours to off-peak hours [86].

Literature shows that the conventional primary objective of DSM is to reduce peak and operation cost [87]. If utilities can effectively incentivize the consumers, peak of load can be reduced, which can benefit the utility by achieving savings in operational costs and possible saving in infrastructure investment. It can benefit consumers as well by reducing their energy consumption cost in response to the pricing signals [64]. However, the smart grids operational paradigm has changed from a load lead paradigm to a generation lead paradigm. Previously, the generation was adjusted to support the load and at present the load can be managed to follow the generation. In such situation where the penetration of RES is increasing, the need for flexibility necessitates proactive use of DSM to ensure secure operations of power system. However, while dispatching the load using DSM, network limitations must be considered. DSM can potentially provide the network with required flexibility but improper DSM can lead to network issues including line capacity and network congestion issues.

The load shifting technique is widely used in literature with focus at appliance level using home energy management systems (HEMs) [88]. With HEMs, a consumer can respond to the signal generated by the system operator to optimize consumer saving. Consumers implement load shifting by switching off or on their deferrable loads in response to some signal from the utility. Time-of-use tariff is commonly applied by utilities to encourage load shifting from peak hours to off-peak hours. Details of dynamic pricing is given below.

2.4.2.1 **Dynamic Pricing**

Dynamic pricing refers to the idea of price of electricity varying at different times. It is considered as most effective way to manage demand by incentivizing consumers. Most commonly reported dynamic pricing schemes include time-of-use (ToU) pricing, critical peaking pricing (CPP) and real time pricing (RTP) [64]. Rates of ToU vary with the time i.e. higher rates during peak hours and lower rates during off-peak hours. This will encourage the consumers to shift load from peak hours to off-peak hours to increase their savings. User has prior knowledge about the duration of high and low prices, which enables them to schedule their energy consumption [64]. ToU tariff is pre-determined for peak and off-peak hours, therefore sometimes it is argued that ToU is not a true dynamic price. However, many researches consider it as dynamic pricing scheme [64].

CPP is a pricing signal, which is applied only on the event days and is a less pre-determined variant of ToU. It can potentially defer construction of new infrastructure that will be used only for very short duration of time. RTP is relatively new pricing scheme where the utility can send pricing signals to consumer in near real-time and consumer can adjust their loads according to the signal and their comfort. RTP provides the utility with higher flexibility in managing the demand but at the same time, it requires an active consumer participation that can lead to consumer discomfort as well [64]. However, the RTP pricing signals should be transmitted to selected consumers so that the limitations of infrastructure including issues due to RES are not violated.

2.4.2.2 **Distributed Energy Resource and Demand Side Management**

The feed-in tariff has encouraged consumers to shift from being a consumer to prosumer. This provides them with an opportunity to consider different types of energy sources including energy storage that can be installed at their premises. This in turn, reduces their dependency on the power grid but increases the uncertainty in the behaviour of load due to stochastic nature of intermittent RES. DSM relies highly on the accuracy of forecast and efficient management of load is only possible with an accurate prediction of the load.

Optimal utilization of RES is another issue, which has attracted a lot of attention from researchers [64]. DSM research should consider impact of RES integration in the system to incorporate the uncertainties in load and generation.

2.4.3 DSM Target Consumer Identification

Application of the DSM techniques discussed above has become possible only with the advent of AMI. For DSM before the advent of AMI, decisions regarding DSM (pricing etc.) were made based on monthly reading of energy consumption. The smart meter provides the opportunity to identify individual consumers for application of DSM. However, it is a difficult task to identify particular consumers, who contribute to the peak of load, amongst thousands of consumers. Authors of [89] used time-series clustering and entropy analysis for consumer behaviour identification. K-means clustering was used in [90] to identify the consumers for DSM campaign. Consumers contributing to the peak of load were identified in [91] using k-means clustering and it was observed that identification of load shapes can be achieved using basic statistical techniques.

The researches referenced above identify the target consumers based on their peaks i.e. if the peak of the cluster coincides with the peak of system, the consumers are considered as target consumers. This can result in the selection of clusters with a high number of consumers resulting in high uncertainty in terms of demand response. Moreover, reference [92] stressed upon importance of load forecasting for DSM application which is often ignored. Load forecasting at lower hierarchical levels of the network can potentially provide a forecast for the consumers or group of consumers that have different energy consumption behaviour and can play a significant role in load shifting by participating in DSM. Therefore, new approaches are required for consumer identification and these studies need to incorporate the uncertainty in terms of consumer participation as well as load forecasting.

2.5 Summary

A comprehensive literature review of different techniques used for data clustering, load forecasting and DSM is presented. Different clustering algorithms and cluster validity indices are discussed. Need for clustering techniques considering application-oriented cluster validity indices is identified as the potential research gap. Moreover, challenges for load forecasting at lower aggregation level of load in shape of noise, volatility and high variability in smart meter data profiles are highlighted. The last section of the chapter discusses different aspects of DSM and the uncertainty due to RES. It highlights the importance of load forecast for DSM and the need for identification of appropriate consumers for DSM application.

CHAPTER 3

3 Extended k-means Clustering for Load Profiling

In this chapter, an extended k-means clustering algorithm is presented. Recursive use of k-means clustering algorithm is implemented by adopting an approach similar to divisive hierarchical clustering [13]. Two new stopping rules are formulated to automate extraction of clusters. The contribution of the research described in this chapter includes the approach to exploit the smart meter data to generate clusters that have high intra-cluster pattern similarity. These clusters are used to develop typical load profiles. The typical load profiles represent the group of consumers at the distribution network level and are suitable for use in smart grid applications including load forecasting [34] and demand side management (DSM) [92] etc. A case study using the smart meter data from the Irish Commission for Energy Regulation (CER) Smart Metering Project [93] validates the applicability and robustness of the extended k-means clustering algorithm.

3.1 Complexity of Smart Meter Data

Before large scale deployment of smart meters at the consumer level, load profiles for different classes of electricity consumers in the United Kingdome (U.K.) were developed using 2500 samples of customers having smart meters with half hourly recording smart meter [94]. Consumers were divided into eight different classes and typical load profiles were generated for each class. However, the integration of renewable energy sources (RES) and demand side management (DSM) require granular load information with high accuracy to help the system operator in making the appropriate decision in real-time. At the distribution level, the load profiling methods traditionally aggregate the consumers' load profiles due to unavailability of the detailed consumer energy consumption. The aggregation of load profiles at such a higher level can result in misclassification of consumers, loss of detailed information and characteristics of the load patterns [4]. An accurate and timely half hourly consumption data from smart meters can help in efficient implantation of DSM by helping the utility in deciding a suitable DSM strategy with a higher degree of confidence [95].

One of the major benefit of the granular smart meter data for utility is the opportunity to analyse the consumer load patterns with increased visibility in the distribution network [4]. Utilities are moving towards smarter grids with smart meters to incorporate RES, ensure reliable energy supply, optimize the existing infrastructure to reduce the need for new infrastructure and expansion of the network, and to give consumers control over their energy use. However, the smart meters pose great data challenges by generating unprecedented data volume, speed and complexity [96]. Utility companies must handle these challenges and use advanced analytics to transform this data into actionable insight [96]. Mathematically, smart meter data of a consumer can be presented as in (3.1) [13].

$$y_c = [y_0(x), y_1(x), \dots, y_m(x)] \qquad (3.1)$$

where y_c is load profile of a single consumer consisting of m time series record. The variable 'y_m' represents the energy consumption magnitude at the time step x_m If *f* is the vector-valued function of y and its derivatives, then mathematically y can be formulated using a system of ordinary differential equations of order n and dimensions m as in (3.2) [13].

$$y_0 = f_0(x) = a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n$$

$$y_1 = f_1(x) = a_1 x_1' + a_2 x_1^{2'} + \cdots + a_n x_1^{n'} \tag{3.2}$$

$$\vdots$$

$$y_m = f_m(x) = a_1 {x^{m'}}_m + a_2 x_m^{2^{m'}} + \cdots + a_n x_m^{n^{m'}}$$

The vandermonde matrix form of (3.2) is given as below [13].

$$\begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ \vdots \\ f(x_m) \end{bmatrix} = \begin{bmatrix} x_0 & x_0^2 & \dots & x_0^n \\ x_1' & x_1^{2'} & \dots & x_1^{n'} \\ \cdots & \cdots & \ddots & \vdots \\ x_m^{m'} & x_m^{2^{m'}} & \dots & x_m^{n^{m'}} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_n \end{bmatrix} \tag{3.3}$$

Or

$$y_c = [F] = [X][A] \tag{3.4}$$

The mathematical formulation of a single consumers load profile as given in (3.3), shows that the energy consumption function of a smart meter data is highly non-linear and modelling such level of non-linearity is a complex task. The complexity for a system operator increases proportionally with an increase in the number of consumers. The transition from a single monthly energy reading to energy readings every 30 minutes can produce up to 48 million records per day for every million consumers and the U.K. is aiming to provide more than 50 million consumers with smart meters by 2020 [97]. According to authors [96] it is estimated that 1 million smart meters will produce 1.3 TB data and storing and processing such volume of data adds to the existing data complexities.

Handling big data brings new computational challenges and extraction of knowledge from the information present in the shape of the big data poses even more challenges. However, it is not possible to analyse each consumer’s load profile at the individual level due to the computational resources and time required for such a large amount of data. Therefore, typical profiles of consumers are required that can represent standard consumer energy consumption patterns for use in power system studies. Extraction of such typical profiles from voluminous data is a multifaceted challenge that requires computational techniques for data feature extraction and data abstraction [2].

Such challenges can be dealt with the use of advanced data mining techniques such as decision tree induction, support vector machines, neural networks, data clustering etc. [22]. Data mining seeks to extract important data features from large, noisy, real-world databases, which can potentially have high inconsistent and missing entries. As the smart meter data is unlabelled and no prior classification information is available, unsupervised data mining techniques can be used to extract the required profiles. Data clustering is a commonly used technique in the domain of load demand profiling for electricity consumers [98] and its application in load demand profiling is discussed in the proceeding section.

3.2 Data clustering

Data clustering is an unsupervised learning technique that has the ability to determine the intrinsic groups in unlabelled smart meter data [2]. Added advantage of the data clustering is its ability to detect data anomalies. It does not require any prior knowledge and can group consumers based on their load patterns [40]. Different clustering algorithms are used for load demand profiling in power systems but k-means clustering algorithm [34, 99-101] and its variants have proven to be most effective for load demand profiling in many comparative studies [98, 102-104].

Hierarchical clustering has an advantage over k-means clustering i.e. it does not require number of clusters to produce a clustering solution [42]. This does not necessarily outweigh the disadvantages, which are present in the form of high computation time, complexity of the algorithm, complexity of selection of the linkage criteria and disregard for global shape or size of cluster [42]. A detailed discussion on the k-means clustering is presented below.

3.2.1 k-means Clustering

k-means clustering is one of the most popular clustering algorithms which finds its applications in many domains including power systems and particularly categorization of electric load [45]. Researchers have used k-means clustering algorithms for electricity load profiling and studies have shown the k-means clustering algorithm to be better than many other clustering algorithms [98, 102, 103]. For example, reference [105] compared k-means, hierarchical, k-means and fuzzy c-means clustering techniques for electric load profile classification and established that amongst these techniques, k-means clustering was superior to others in terms of processing speed and robustness. The simple operating principle, fast processing speed, robust to temporal resolution effects and efficiency in shape preserving [104, 105] makes k-means favourable for clustering the electricity consumer load demand profiling.

3.2.2 k-means Clustering Working Principle

k-means clustering is a simple and robust partitional clustering algorithm. This algorithm assigns each pattern to nearest centroid i.e. cluster centre. A conventional k-mean clustering algorithm follows principle as shown in Figure 3.1 [10].

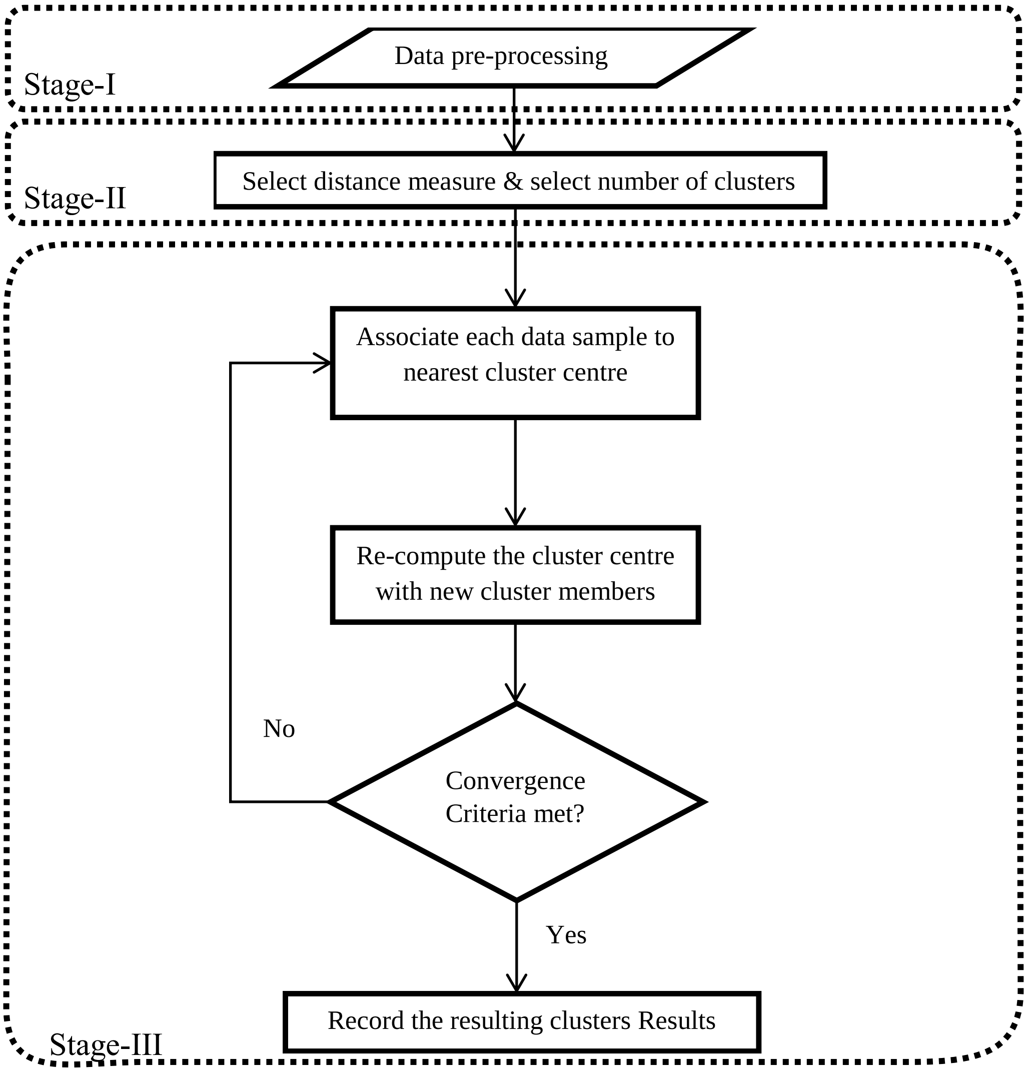

Stage-I
Data pre-processing
Stage-II
Select distance measure & select number of clusters
Associate each data sample to nearest cluster centre
Re-compute the cluster centre with new cluster members
No
Convergence Criteria met?
Yes
Record the resulting clusters Results
Stage-III

Figure 3.1 *k-means clustering process flowchart [10]*

As shown in figure 3.1, k-means clustering is divided into 3 stages. As the first stage of k-means clustering, the input data is pre-processed which involves sorting the data into a format where the application of k-means clustering is possible. Real-world smart meter data requires data sorting to sort and re-arrange into a form where the application of a clustering algorithm is possible. The data is re-arranged into a $n \times p$ matrix to make the data format compatible with clustering application. The n is number of rows which represents number of smart meter consumers in the dataset and p represents the number of columns which in this case are data features.

Apart from re-arranging the data, removing missing and erroneous data, and detection of outliers is performed as part of k-means clustering to ensure data quality. Another pre-processing step involves scaling the data between 0 and 1 to remove effect of magnitude. However, in some applications like DSM, representation of magnitude of electricity demand in a cluster is as important as the representation of patterns. Applications like DSM require reflection of consumer pattern as well as magnitude to select DSM target consumers and potential strategies.

Different distance measures used for k-means clustering include Euclidean distance, Squared Euclidean distance, Minkowski distance and Mahalnobis distance as shown in Table 2.1 (Chapter 2, Section 2.2) [23]. However, due to its ability to detect strays, squared Euclidian distance is one of the most popular and commonly used distance measures [52]. It uses the sums of squares of distances to compute the distance between a cluster member and cluster centroid. For two different data point x_i and x_j, squared Euclidian distance [52] can be mathematically represented by (3.5) [13].

$$d_{sqeuc}(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} - x_{jk})^2 \qquad (3.5)$$

The ability of squared Euclidean distance to notice outliers makes it more favourable for k-means clustering. The second stage of k-means clustering includes selection of the distance measure and selection of number of clusters. A detailed discussion selection of appropriate number of clusters is given in chapter 2 (Sections 2.2.4 and 2.2.5)**.** Other parameters that require to be specified include a number of iterations and repetition of the algorithm to find the global optima.

After selection of number of clusters, the third stage starts with random allocation of number of centroids (cluster centres) equal to the number of clusters. The data points nearest to each centroid are allocated to the centroid. After allocation of the data points, the position of centroids is re-computed to minimize the distance between the cluster members and the centroid. As a result of change in centroid position, the data points change their affiliation of centroid. Thus, at each iteration, membership of a cluster can change. In this stage, k-means tries to minimize the error function that can be given as in (3.6);

$$E = \sum_{i=1}^{k} \sum_{x \in c_i} d(x, \mu(C_i)) \qquad (3.6)$$

where $\mu(C_i)$is the centroid of cluster C_i; $d(x, \mu(C_i))$ shows the distance between data point x and the centroid of cluster i. Calculation of the distance depends on the selected distance measure.

k-means clustering will re-iterate the process untill it finds the minimum value of the error, E. Once the k-means clustering algorithm finds optimum clusters with minimum value of E, the cluster membership will not change. This will indicate that the algorithm has converged to a solution. Finally, the k-means clustering algorithm records the results of clustering.

Although the principle of k-means clustering is simple, however, the algorithm faces some challenges including selection of number of clusters and convergence to local minima. These are discussed below in detail.

3.2.3 Number of Clusters

Selection of number of clusters i.e. k is an important issue that is often contentious. Different methods of selection of k are used but the most commonly used ones include the elbow method and X-means clustering [50]. The elbow method determines the number of clusters by explaining the percentage of variance as a function of k. X-means clustering [50] keeps splitting the clusters unless the Akaike information criterion (AIC) or Bayesian information criterion (BIC) are reached. The computation involved in X-means makes it complex and expensive method, which sometimes makes it unfavourable for use in clustering big data sets.

Selection of number of clusters is a factor, which is data and application driven. It is governed by the definition of an appropriate cluster that can only be defined by either the application of the clusters or by an expert in the domain. For example, for demand side management (DSM) application, the required level of detail of load variation is significantly higher than that of power system reliability. Therefore, domain and application-based criteria can possibly present a better solution for the selection of k for use in the particular domain.

3.2.4 Convergence to local minima

k-means clustering is sensitive to the initialization of the centroid [43]. The initial position of the centroid can lead to sub-optimal solution. Poor results can be avoided by performing clustering sufficient times with new random seeds each time for initialization of centroids at different starting positions [106].

3.3 Extended k-means Clustering

The above discussed issues require new approaches to extend the use of k-means in the adoption of big data like electricity consumption data clustering for power system applications. Aim of electricity consumption data clustering is to extract clusters from the data that not only achieve high intra-cluster pattern similarity but also isolate the outliers. However, intra-cluster pattern similarity is highly dependent on the number of clusters.

Applications of k-means clustering usually do not contain any explanation or justification for selecting pre-specified values for k [13]. As discussed in the literature review chapter 2 (section 2.2.5) different cluster validity indices can be used to validate the clustering solution. The relative criteria proposes different cluster validity indices including Caliński and Harabasz index [57], Davies-Bouldin index [107] and Silhouette index [41]. However, these methods tend to evaluate the value of k from the pre-specified range of numbers. Rather than allowing the data to dictate the number of clusters, a judgement is made based on pre-specified assumptions of a number of clusters. Therefore, a new approach is required to adopt the number of clusters for load demand profiling based on the pattern similarity in the data to be clustered.

The three cluster validity indices i.e. Caliński and Harabasz, Davies-Bouldin and Silhouette index were used in the initial case study to evaluate the best values of k using the smart meter data from [93]. Empirical evaluation reveal that although these criteria define the compactness of the clusters for general data clustering purposes, they are still not able to meaningfully discriminate the dissimilarity in the patterns of the volatile smart meter data [13]. This is primarily due to the level of intra-cluster pattern similarity required in specific applications [13].

To resolve the issues faced by the standard k-means clustering algorithm, an extension of the existing k-means clustering is proposed in this research. The approach is inspired by divisive hierarchical clustering, which initially considers the entire dataset as a single cluster and keeps splitting each cluster into two clusters until each data point represents a cluster [20]. This approach resembles the classification process in the human brain [21]. The divisive hierarchical clustering is not commonly used due to their computational complexity and in addition, divisive hierarchical clustering cannot undo the splitting of clusters once it is done [41]. On the other hand, recursive application of k-means on the smart meter data can produce divisive hierarchy that is based on the construction of a binary tree [13].

The existing k-means clustering requires to be extended for recursive application and thus the new clustering algorithm is called extended k-means clustering algorithm. However, the splitting of clusters should be subjected to certain rules, which can define concept of a meaningful cluster for the specific application. This will eliminate the requirement for pre-specified cluster numbers that is a major issue with k-means clustering.

Different stopping rules are used in hierarchical clustering and are discussed in chapter 2 (Section 2.2.5). However, these rules are not suitable for k-means clustering due to the intensive computation involved in calculations. Therefore, new stopping rules are required to extract clusters which can be used in load forecasting and DSM.

Having a high number of clusters can potentially produce more compact clusters but it can render the clustering process futile a with large number of clusters. Thus, a balance between the number of clusters and pattern similarity must be maintained to minimize the compromise on the accuracy while reducing the data [13]. Therefore, there should be a threshold number of consumers below which the cluster should be considered having sufficient intra-cluster pattern similarity [13]. However, determining a minimum number of consumers in a cluster requires some empirical evaluation. As discussed in chapter 2 (Section 2.2.4), visualization of data can help in understanding the data and potentially lead to discovery of certain features of data. Moreover, data visualization can be considered as most direct method to estimate number of clusters. However, a large number of samples and the dimensionality of the data poses challenges. A small experimental data can be used to visualize the data and to decide the efficacy of clustering [54] and from multiple experiments, inference on the minimum number of consumers in a cluster can be made.

From initial clustering trails using k-means for clustering the smart meter data, it is deduced that the intra-cluster pattern similarity of clusters having less than one percent of the total population becomes sufficient for pattern extraction [13]. Therefore, if during the recursive application of k-means clustering, the number of consumers in a cluster become lower than one percent of the total population, the cluster will not be sub-divided and considered as final cluster. The condition of having a minimum or lower number of consumers in a cluster defines the first stopping rule for the recursive application of k-means.

Although, the first stopping criterion stops the clustering process from dividing into too many clusters, there is a possibility of having cluster having high number of consumers and high intra-cluster pattern similarity as well. Therefore, an additional stopping rule is required that can reflect the intra-cluster pattern similarity. A second stopping rule is defined based on the notion of intra-cluster pattern similarity. The commonly used measure of similarity of two load curves for the purpose of load forecasting in power systems is mean absolute percentage error (MAPE) [108] which can be given as in (3.7)

$$E_k = \frac{100}{m} \times \left[abs \sum_{i=1}^{m} \left(\frac{C_{k_i} - \mu(C_k)}{C_{k_i}} \right) \right] \qquad (3.7)$$

where E_k is the error of cluster *k* with respect to mean of cluster centre $\mu(C_k)$ and C_{ki} represents the consumer 'i' of cluster *k* and *m* represents total number of consumers [13].

The second stopping rule is based on the saturation of MAPE in a cluster. At each stage of the ramification of the cluster, the MAPE is computed for both parent and child clusters. During the process of re-clustering, the MAPE declines towards a minimum, beyond which it saturates such that [13];

$$E_p \geq E_c \qquad (3.8)$$

where E_p is the error of the parent cluster and E_c is the error of the child cluster and (3.8) gives the stopping rule 2. Therefore, if the re-clustering process does not satisfy stopping rule 1, the re-clustering process is continued to determine the minimum MAPE, i.e. saturation level of MAPE. An important aspect of the second stopping rule is splitting of clusters and then deciding that the previous cluster should not have been split. The ability of the algorithm to retrieve the previous cluster makes it superior to the divisive hierarchical clustering. The flow chart for the extended the k-means clustering by incorporating both stopping rules is given in Figure 3.2. [13].

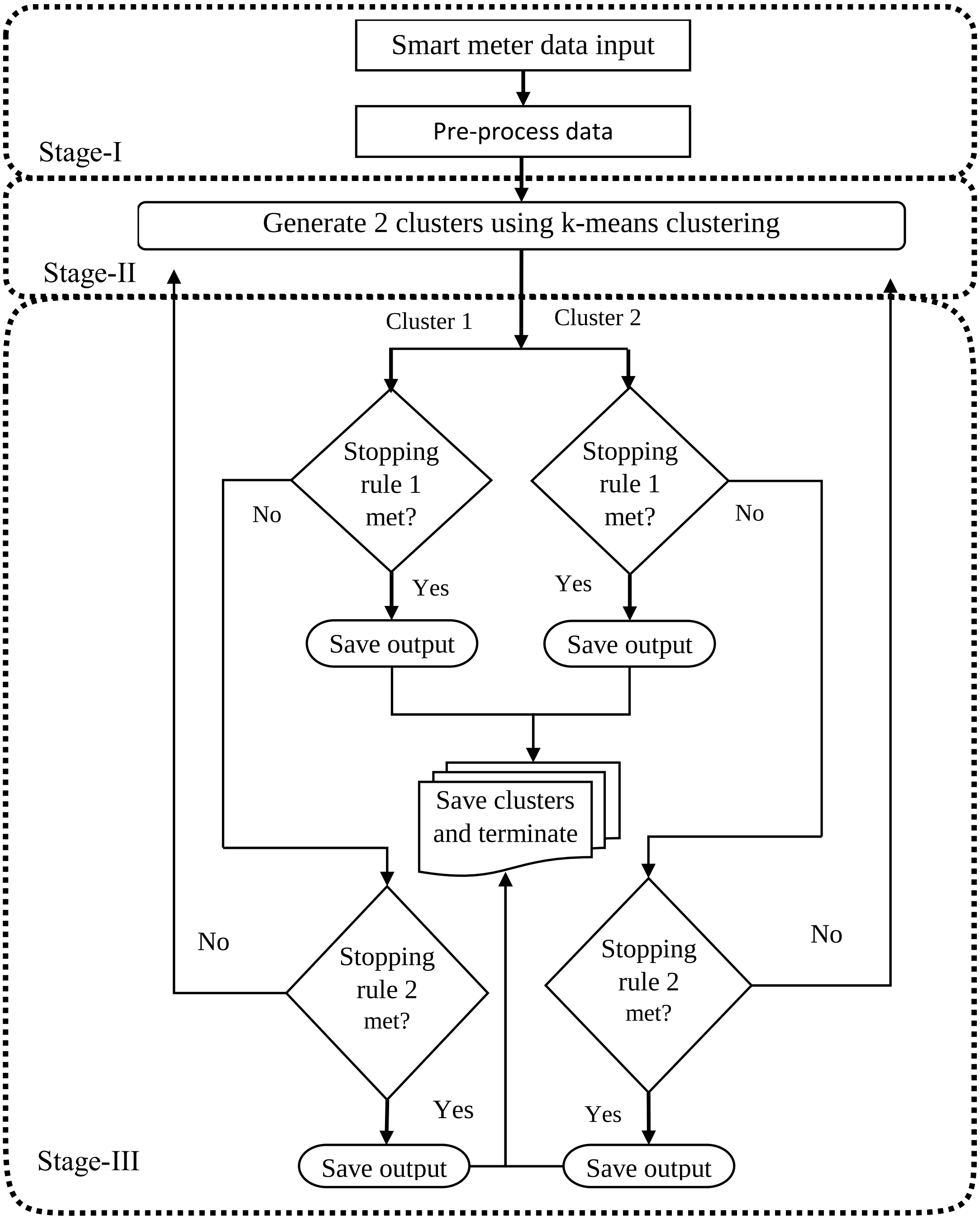


Figure 3.2 *Extended k-means clustering algorithm (adopted from [13])*

From Figure 3.2, the entire process is divided into three stages. At first stage, the data pre-processing is carried out. Data pre-processing is discussed in detail in section 3.2.2. At second stage, the k-means clustering algorithm is applied with value of '*k*' i.e. number of clusters as 2. The third stage evaluates the clusters against the stopping rule. At first, it is checked whether the clusters generated have number of consumers less than 1% of the total population, if true, the cluster is considered as terminal cluster. Otherwise, the cluster is checked against the second stopping rule i.e. value of MAPE in parent and child clusters is compared. Naturally, the MAPE declines with splitting of clusters as the clusters become more compact. If MAPE of the child cluster becomes higher than the parent cluster, i.e. the MAPE saturates, the cluster is considered as terminal cluster otherwise the child cluster is sent back to second stage. The child cluster sent back to the second stage will be considered as parent clusters and recursive application of k-means clustering will be continued on all the child cluster which do not meet one of the stopping criterions. Eventually all clusters will meet one of the stopping criterions. All of the terminal clusters are saved as final cluster for extraction of load profiles. Thus, by combining the speed and efficiency of k-means and divisive hierarchical approach clusters having high intra-cluster pattern similarity are extracted. The algorithm allows data to dictate the number of clusters which produces more meaningful clusters for load forecasting.

The extended k-means clustering process divides the entire population into two clusters such that some consumers, which can be represented as $F_{c1}, F_{c5}, F_{c12}, F_{c13}, F_{c18}, \ldots$ are in one cluster, while the other consumers represented as $F_{c2}, F_{c3}, F_{c4}, \ldots$ are assigned a second cluster. The re-clustering process again splits each cluster into two such that the two clusters resulting from division of cluster one become $F_{c1}, F_{c12}, F_{c15}, \ldots$ and $F_{c5}, F_{c13}, F_{c18}, \ldots..$ Similarly, each cluster is subdivided until all clusters meet one of the stopping criteria. A typical cluster is formulated as in (3.9) [13].

$$\begin{bmatrix} F_{C1} \\ F_{C12} \\ F_{C15} \\ \vdots \end{bmatrix} = [X_1 \quad X_{12} \quad X_{15} \quad \ldots] \begin{bmatrix} a_1 \\ a_{12} \\ a_{15} \\ \vdots \end{bmatrix} \qquad (3.9)$$

3.4 Average Demand Profile

The clusters generated from the above-discussed extended k-means clustering are used to generate average demand profiles. The average demand profiles represent each cluster using a single profile and thus reduce the data burden from thousands of profiles to a few. An averaged profile from a cluster can be obtained using (3.10);

$$P_{ck} = \frac{1}{n}\sum_{i=1}^{n} F_i \qquad (3.10)$$

where P_{ck} represents load profile for cluster k, n is the total number of consumers in the cluster k and F_i represents the profile of consumer i in cluster k.

The profile developed using (3.10) are referred to as typical load profile for the cluster. This profile can be used for many system studies including load modelling [109], load forecasting [75], system contingency and reliability studies [110] etc. These studies tend to incorporate the distribution system level data to incorporate the impacts of load patterns in the system studies by clustering large datasets for data reduction. The representative profiles of the clusters reduce the computational burden by representing the entire members of the cluster by a single profile.

A case study is simulated using the real-world smart meter data to validate the clustering approach developed in this research i.e. extended k-means clustering. Details of the case study are given in the proceeding section.

3.5 Extended k-means Clustering Application and Validation

To validate the applicability of the extended k-means clustering algorithm, a case study is conducted by incorporating a dataset from Irish Commission for Energy Regulation (CER) Smart Metering Project [93] (available from Irish Social Science Data Archive (ISSDA)). The Smart Metering Electricity Customer Behaviour Trials (CBTs) took place during 2009 and 2010 with over 5,000 Irish homes and businesses participating [93]. The data set used in this study consists of 6 comma-separated values (CSV) files with 180 million rows of data for more than 5,000 consumers. The data consists of 25,728 time series half-hourly records of energy consumptions for each individual consumer including domestic and small businesses consumers for 1.5 years starting from the first of January 2009. The details of the dataset are given in Appendix A.

According to reference [111], '*Big Data concern large-volume, complex, growing data sets with multiple, autonomous sources*' and according to this definitions, the smart meter data qualifies as Big Data [4]. However, due to the limited size of dataset i.e. only 25,728 time series records, the dataset does not require adoption of special platforms required by Big Data. MATLAB platform is used to convert the data from the comma-separated values file format to MATLAB format. The data is sorted into $n \times p$ matrix form to ensure compatibility with the requirement for application of k-means clustering. Each row of the dataset represents a consumer whereas the columns represent the features or time step.

The next step involves the application of extended k-means clustering. To validate the applicability and robustness of the clustering algorithm, three different scenarios of clustering are considered. Three scenarios include a weekly scenario, which shows higher level of variations in the load to model weekly load profiles and can be used for applications requiring short term load profiles, for example, demand side management. A monthly scenario that captures the seasonal trends with a broader horizon is considered as the second scenario and finally third clustering scenario considers clustering for an entire year which can be used in long term system studies like load forecasting. The dataset considered in this study contains only one year of data so that all of the scenarios can be compared.

In the first scenario, the pre-processed data is clustered on a weekly basis. The load data for one year is divided into 52 weeks. Thus, 52 clustering solutions were generated for the first scenario. The profile for each week with a temporal resolution of 30 minutes has a dimensionality of D = (60×168)/30 i.e. 336 [13].

In the second scenario, the time window for the clustering process is enhanced from one week to one month to capture the seasonal effects for medium term studies. The dataset is split into 12 segments depending on the duration of each month. Clustering is performed on a monthly basis using one year’s smart meter data. Each monthly load profile has different dimensionality depending upon the days in the month for example, for January (31 days) the dimensionality will be D = (60×744)/30 i.e. 1488 and for February 2009 (28 days) the dimensionality will be D = (60×672)/30 i.e. 1344 [13].

Thirds scenario simulated clustering of entire population for one year. The temporal resolution remains unchanged and the dimensionality, for third scenario is very high as compared to other two scenarios i.e. D = (60×24×365)/30 i.e. 17520 [13] .

In all three scenarios of clustering, there were several clusters with only one or two consumers. Such clusters can result from high magnitude of the load or due to error in data collection [24]. It is possible for some smart meter consumers to have unique energy consumption pattern, which does not resemble any other consumer. The extended k-means clustering explores the smart meter data deep enough to identify consumers with unique energy consumption patterns that are statistically termed as outliers [13]. Apart from the uniqueness of the patterns, the outliers can potentially result from bad data [24]. Moreover, from these clusters, there were certain clusters that contained zero values for long duration. All such consumers which are considered outliers or have missing or zero value, are excluded from the dataset to ensure high data quality.

As the variability of load demand is higher in the first scenario, it produces highest number of clusters. This higher variability is captured in extended k-means clustering due to short span of the time and consequently the patterns tend to have higher MAPE, which results in higher number of clusters. The clusters produced using the weekly clustering data tend to be more compact and have high intra-cluster pattern similarity as show in Figures 3.3 to Figure 3.8.

The Figures 3.3 to Figure 3.8 show that the clusters resulting from the extended k-means cluster are compact and consumers having similar patterns are clustered into same cluster. The dominant pattern can be clearly identified by visual inspection of the plot of the cluster and profiles of the clusters are also presented in the Figures using bold black lines. The plotted Figures i.e. Figures 3.3 to Figure 3.8 are randomly selected from different weeks and clusters in those weeks. It can be observed from Figures 3.3 and Figure 3.6 that the pattern of both clusters seems to be similar. Their profiles show almost similar pattern but with difference in the magnitude of energy consumption. The cluster in Figure 3.3 is selected from week 17 whereas cluster in Figure 3.6 is selected from week 34 and similarity of the pattern shows that both clusters might have same consumers with variation in level of their energy consumption. This shows that the clustering solution produced using extended k-means clustering algorithm is robust and can identify similar patterns over different period in time.

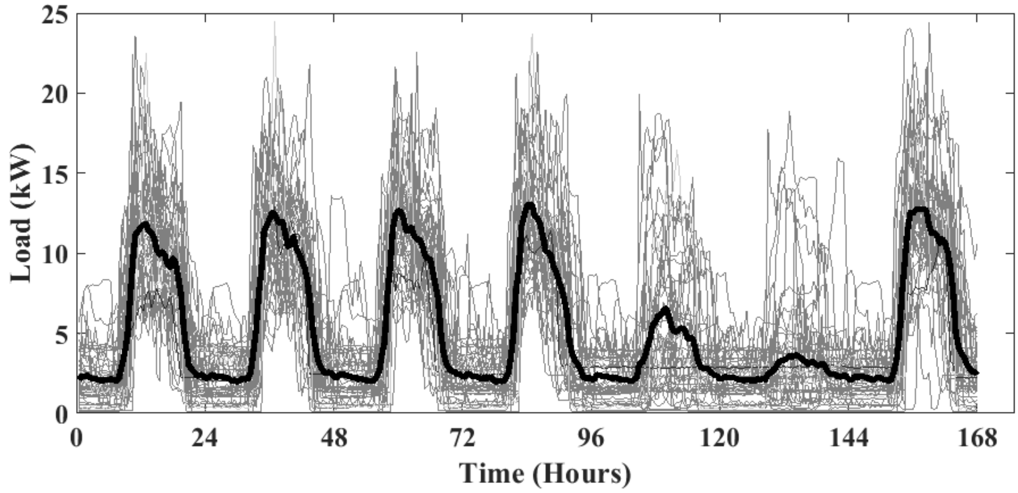


Figure 3.3 *Plot of Cluster 11, week 17*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

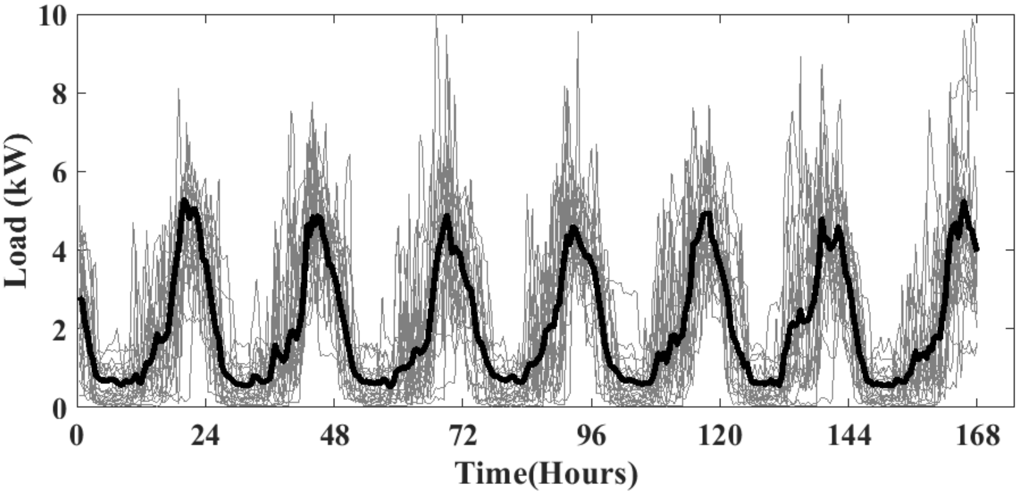


Figure 3.4 *Plot of Cluster 24, week 25*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

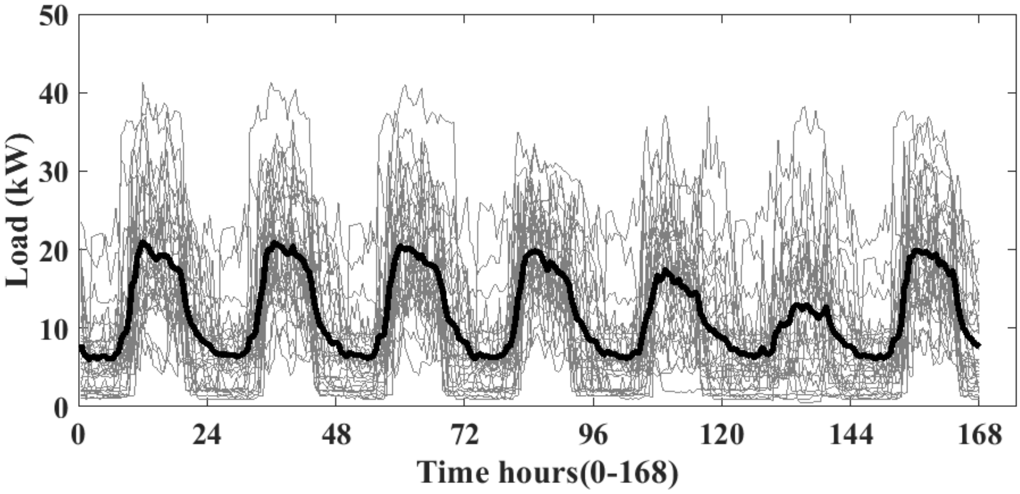


Figure 3.5 *Plot of Cluster 5, week 30*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

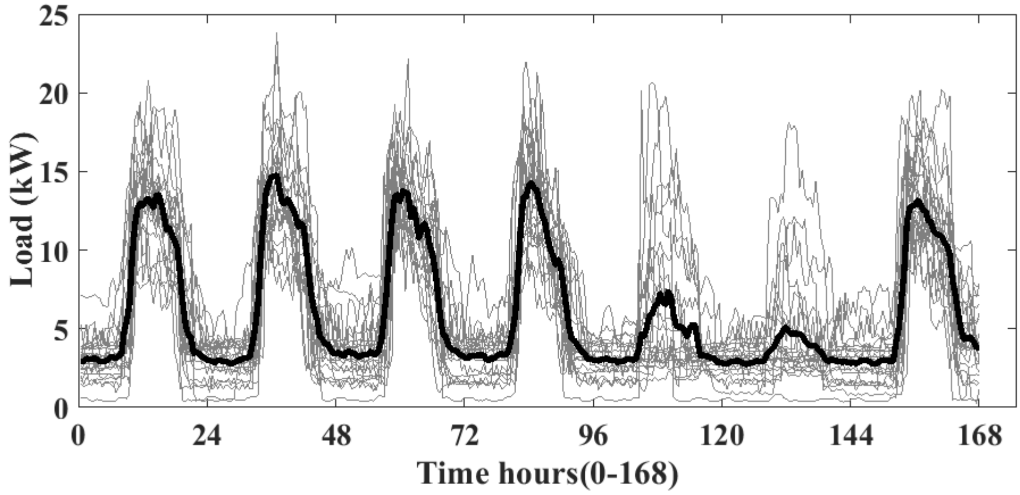


Figure 3.6 *Plot of Cluster 22, week 34*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

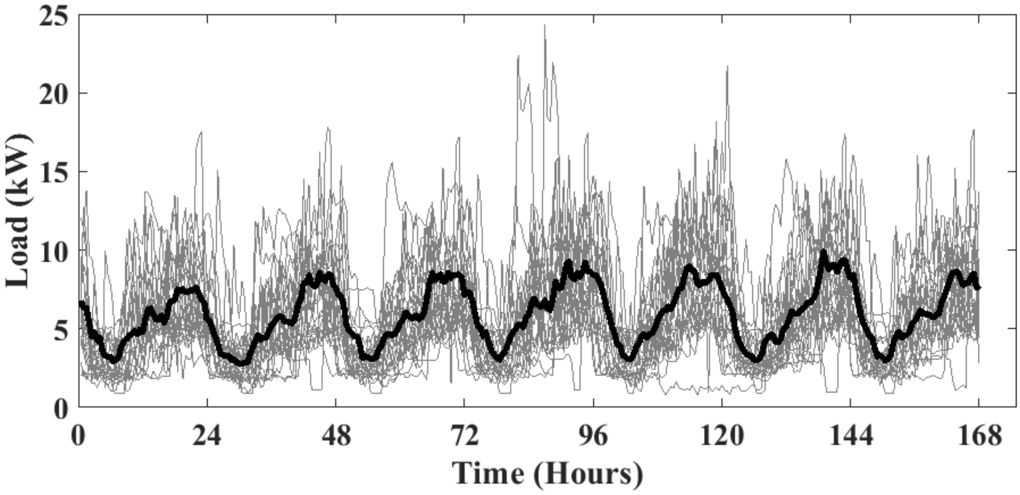


Figure 3.7 *Plot of Cluster 12, week 45*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

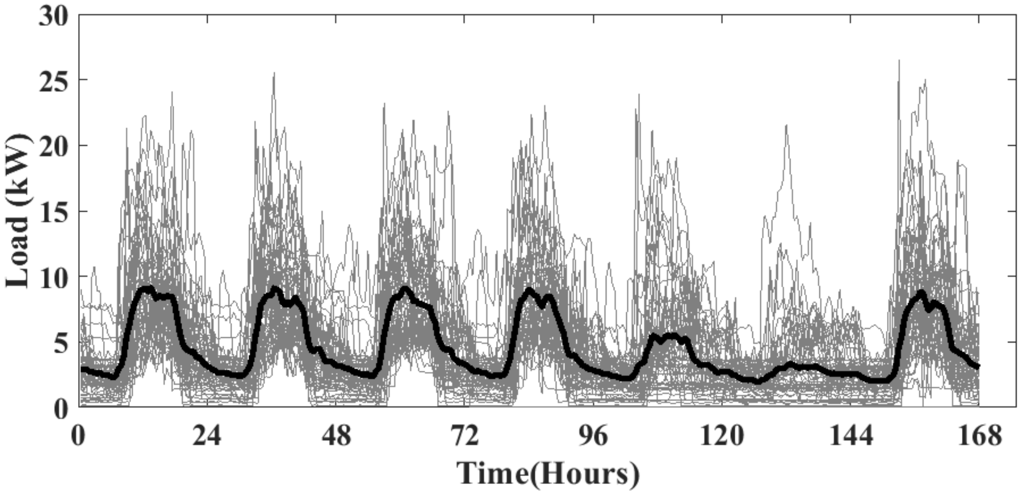


Figure 3.8 *Plot of Cluster 11, week 52*

Grey lines show load profiles of the individual consumers in the cluster and thick black line shows the typical load profile for the entire cluster

The second scenario produced lower number of clusters as compared to the first scenario. With the longer time span, the clusters tend to be less compact and the MAPE saturation level is achieved with higher number of consumers as compared to the first scenario.

A similar trend has been observed in third scenario where the number of clusters produced from extended k-means clustering are lowest [13].

The simulations for all three scenarios are repeated several times to verify the consistency of the results and stability of the solutions provided by the algorithm. The simulation showed consistent clusters for all scenarios, which shows that the extended k-means clustering algorithm produces clusters that are dictated by the data. The intra-cluster similarity of the patterns is evident from visual inspection of weekly clustering results. Thus, extended k-means clustering approach provides robust clustering solution with clusters having consumers with high intra-cluster pattern similarity.

3.6 Summary

An innovative clustering approach is presented in this chapter. The approach extends the existing k-means clustering algorithm by generating binary trees of clusters using recursive application of k-means clustering. Two stopping rules are formulated to stop the splitting of the clusters such that compact clusters with high intra-cluster patterns similarity can be extracted. The proposed algorithm is validated using real-world smart meter data by simulating three clustering scenarios. The results for all three scenarios validate the applicability, stability and robustness of the proposed clustering approach by producing compact clusters having high intra-cluster patterns similarity.

CHAPTER 4

4 Novel Load and Generation Profiling Approach

In this chapter, novel approach for modelling the load profiles is proposed. Raw load profiles extracted from smart meter data using extended k-means clustering are used to model them into the proposed alternative profiles. A mathematical framework is proposed to model these alternative profiles which are concatenation of linear profiles and can be used for different applications including Monte-Carlo simulation applications, load forecasting and demand side management. The proposed approach of alternative load profiles is extended to model intermittent renewable energy sources (RES) profiles into alternative generation profiles to reduce the variability of the intermittent RES profiles.

The major contribution of the research described in this chapter includes the development of novel load profiling approach and innovating the approach for RES generation profiling. The proposed approach can reduce complexity, noise, variability and volatility of smart meter data and due to linear nature of proposed profiles, increase convexity for use in energy optimization. Case studies are simulated to validate use of proposed approach in power system applications.

4.1 Demand Profiling in Big Data Era

Conventional load profiling in power grids has mainly focussed on the development of techniques for consumer characterization [112]. However, with large scale deployment of smart meters, the computational burden has become manifold. Increased data resolution inevitably brings the increased volumes of data at high frequency with added data complexity which impacts the efficiency of data processing [113]. In the era of smart grids, the planning horizon has become narrow with sudden variations in RES generation and deployment of new technologies like electric vehicles, heat pumps require more data to ensure high accuracy but faster decision with efficient computation and lesser processing time [13]. The smart grids need to perform localized studies to implement distributed control but with millions of smart meters, the computational and data complexity pose great challenges for such studies.

Although the accuracy of system study relies on the load model being reflective of the actual network loads, direct use of big data of smart meters for detailed large-scale simulation is computationally expensive [13]. Data mining techniques, particularly data clustering reduces the burden of big data of smart meter by creating clusters of similar load profiles. Apart from the size of data, noise, volatility and high variability of smart meter data particularly adds to the existing complexity of the data functions and consequently, the efficiency and accuracy data driven application in power systems is significantly affected. Moreover, smart meter data is becoming heterogeneous with increasing integration of intermittent RES, making it more volatile and variable.

The highly non-linear complex energy profiles prove to be computationally expensive and difficult to handle directly by linear system study methods particularly for power system applications that are purely data driven [114]. In smart power grids, modelling energy consumptions at every node is challenging, particularly for applications requiring linear processing of the non-linear data functions. The high-resolution smart meter data provides granular information about consumers' load and can yield better results if used for studies that incorporate load data. Thus, the key challenge in modelling smart meter data is the determination of an appropriate representation of the load to decrease uncertainty margins with reduced complexity and computational burden [13] .

Considering the aforementioned, this research proposes a novel and uniform approach to model the load profiles extracted using extended k-means clustering. The approach simplifies the complexity of data while maintaining the precision of representation close to the original data as compared to conventional linear approximation techniques. The approach proposes an alternative representation of the load profiles resulting from the extended k-means clustering by modelling them as concatenation of linear profiles. The new refined load profiles are hereafter called alternative profiles and the development process of alternative profiles for smart meter-based demand profiles as well as RES generation profiles is described in proceeding sections.

4.2 Development of Alternative Load Profile

Challenge of large volume of data of thousands of smart meters can be managed by data clustering for extraction of averaged load profiles. However, issues of noise, volatility, variability and complexity of data persist despite clustering of smart meter data. This limits linear processing of this data in different linear methods of power system studies [13]. In order to address these issues, researchers have used linear approximation of the load curves such that non-linear problems can be solved using linear techniques such as linear programming [115].

The energy data generated from the smart meter is deemed proportional to load data [116]. The hourly smart meter energy data is considered equivalent to load data, as the magnitude of total energy and power consumption during the hour is the same [116]. Many researchers have tried to linearize the load profiles. Load curve linearization in a multiple energy system was achieved using the the Douglas-Peucker algorithm for use in expansion planning by reference [115]. Often the linearization if carried out by selecting typical days instead of the full year [117, 118] or by using a stepwise approximation of load duration curve [119]. However, authors of [120] argue that although stepwise approximation of the curves should be carried out as close to the original curves as possible and they recommend to select typical days due to computation burden thus avoiding the linearization of yearlong curves [13].

Review of literature shows that no study has been conducted in power systems using smart meter data to address these issues. Linearization of load profile at a higher aggregation level using substation load profile is much different than smart meter-based load profile. Conventional linearization approaches use either stepwise linearization or linear approximation of load curves, which are averaged profiles at high aggregation level. The example of stepwise and linear approximation techniques are shown in Figure 4.1 [121] and Figure 4.2 [115] respectively.

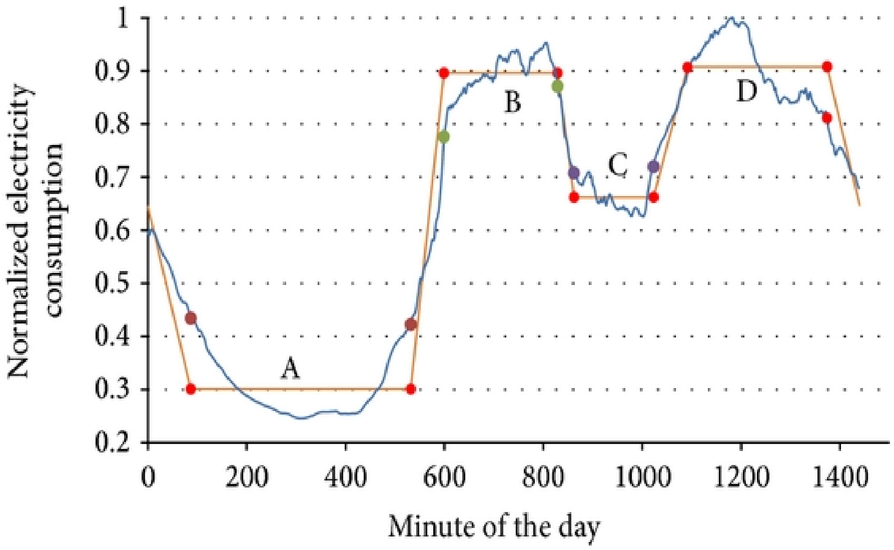


Figure 4.1 Stepwise linearization of energy consumption profile[121]

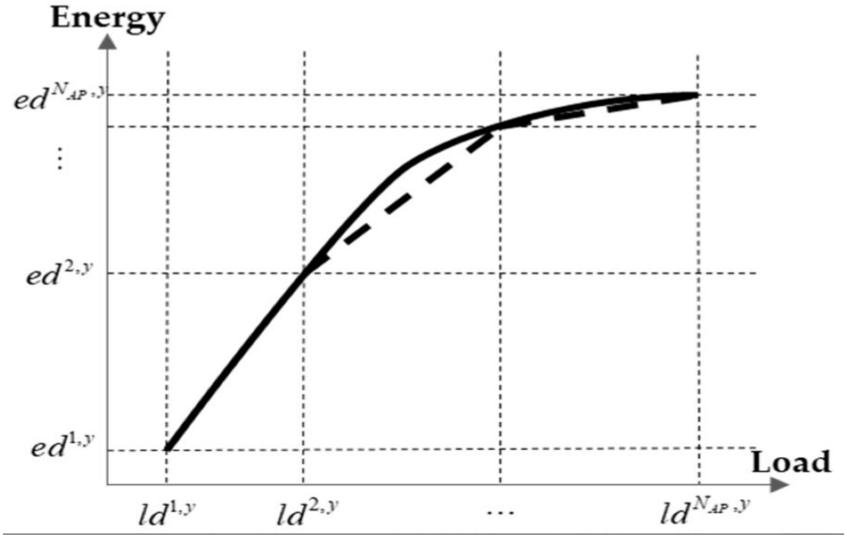


Figure 4.2 Piecewise linear load-energy curve [115]

The 24 hours demand profile in Figure 4.1 (substation level) and the energy load-energy profile in Figure 4.2 (multi-energy system) are comparatively much smoother than profiles at lower aggregation level that are generated using smart meter data. Figure 4.3 shows load profile of a cluster which aggregates the smart meter data in at distribution level.

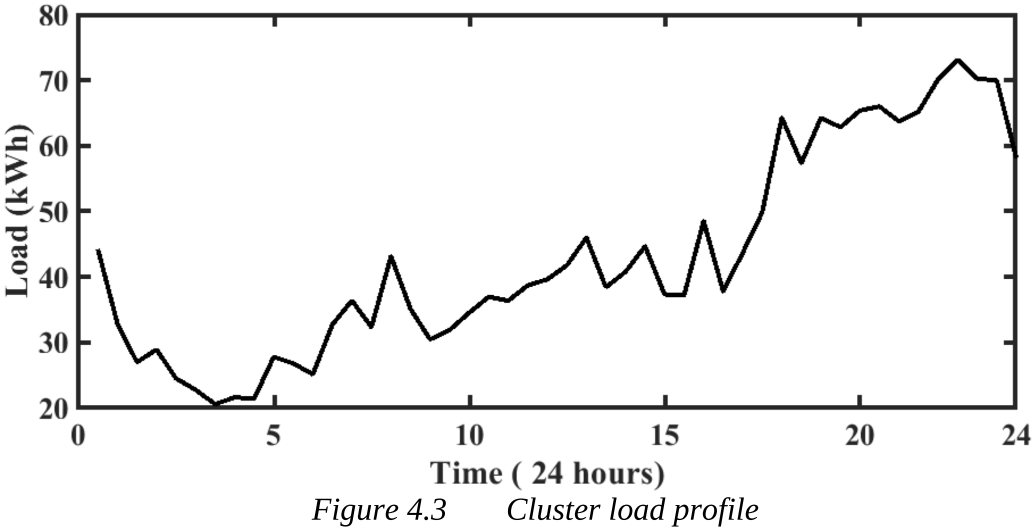


Figure 4.3 Cluster load profile

From Figure 4.3, it can be seen that at the lower aggregation level the volatility of smart meter data is pronounced. Linearization of such patterns requires different approaches than the existing approaches because apart from the inaccuracy in representation, these existing approaches can either increase the computational burden by adding segments in the piecewise linearization or lose track of the trajectories of the load curves.

To address the problems faced in linearization of the smart meter-based load profiles, a novel approach is proposed to generate alternative load profiles. Figure 4.4 shows the flow chart detailing each step of the development of alternative load profiles. The process can be broadly divided into five stages. At the first stage in development of the alternative load profiles, extended k-means clustering is applied to extract the raw load profiles. The details of application of extended k-means clustering are provided in Chapter 3. The rest of the stages in development of alternative load profile are discussed below in detail.

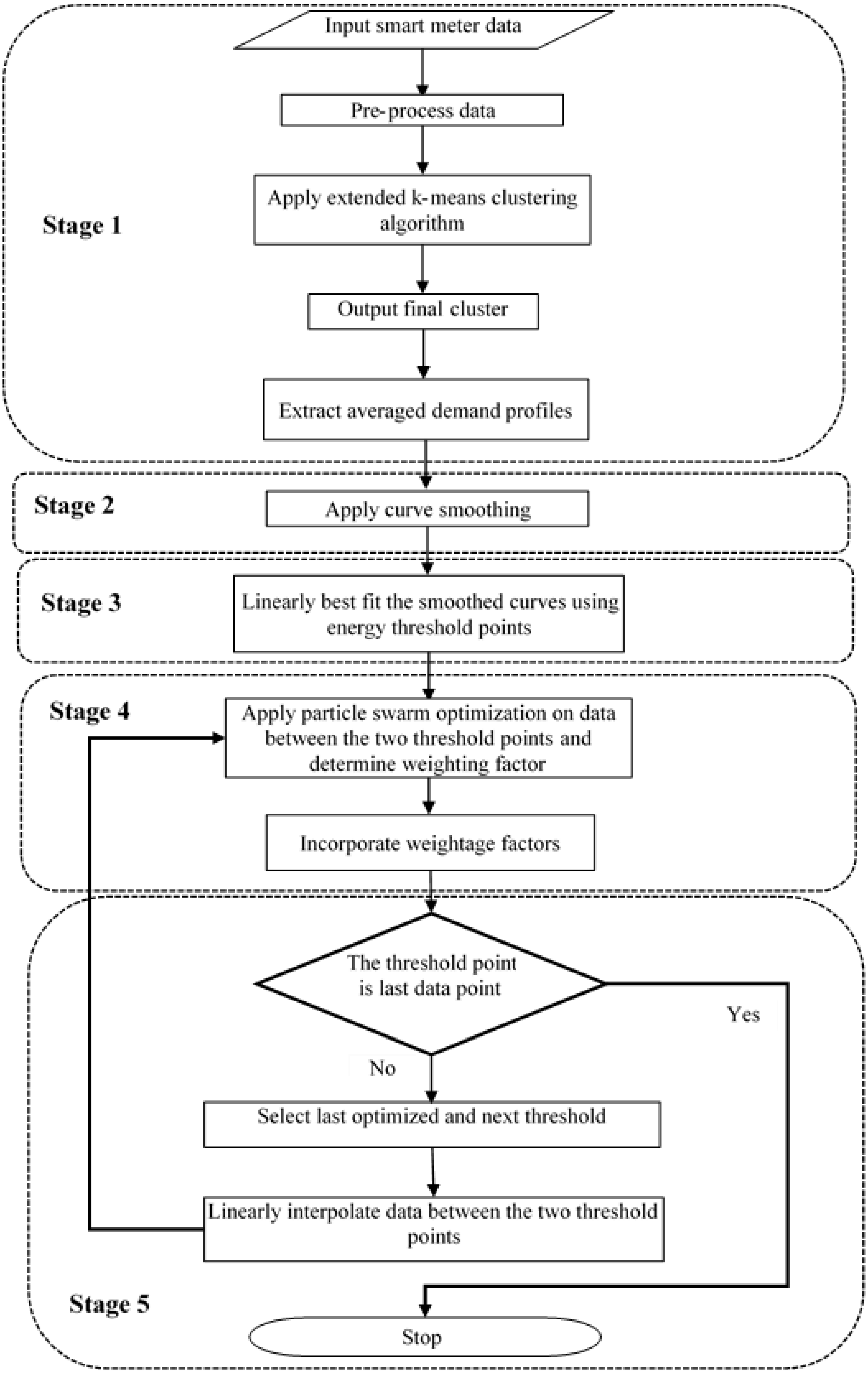

Input smart meter data
Pre-process data
Apply extended k-means clustering algorithm
Stage 1
Output final cluster
Extract averaged demand profiles
Stage 2
Apply curve smoothing
Stage 3
Linearly best fit the smoothed curves using energy threshold points
Stage 4
Apply particle swarm optimization on data between the two threshold points and determine weighting factor
Incorporate weightage factors
The threshold point is last data point
Yes
No
Select last optimized and next threshold
Linearly interpolate data between the two threshold points
Stage 5
Stop

Figure 4.4 Alternative energy profiling process [13]

4.2.1 Curve Smoothing and Linearization (Stage 2 &3)

The raw load profiles generated by averaging the consumption of the consumers in clusters have significant fluctuations in them. These fluctuations can often obscure the developing load trends and clarity of patterns within the data. The noise in the data can be reduced by curve smoothing techniques like polynomial curve fitting [122] and moving average smoothing [123] etc. The polynomial fitting with lower degree of polynomial struggles to capture the higher variation whereas, Runge's phenomenon [122] restricts the use of a higher degree of the polynomial fitting. Noises can be attenuated by averaging [123], and noise, in this case, can be identified as unsought variations in the load profile. Because of its simplicity and accuracy [123], moving average smoothing is a suitable candidate for averaging the data [123] and is used in the second stage of the alternative profile development. The second stage takes the raw load profiles and applies moving average smoothing to reduce the noise in the data. Moving average smoothing is a convolution process and can be represented mathematically by (4.1) [123].

$$\bar{y}_j = \frac{\sum_{i=-m}^{m} C_i y_{j+1}}{N} \qquad (4.1)$$

where y is the variable, C_i is the convoluting integer, j is the running index of data and N is the number of time periods [123].

After smoothing, the load profiles are linearized at the third stage of the alternative load profile development. Linearization of non-linear data functions is achieved using the Taylor series linearization process [124]. The energy threshold points are considered as the operating points for linearization. The energy threshold points are determined such that for a threshold point y_i [13].

$$y_{i-1} < y_i > y_{i+1} \tag{4.2}$$

or

$$y_{i-1} > y_i < y_{i+1} \tag{4.3}$$

where y_i represents the magnitude at time step i, y_{i-1} is magnitude at timestep $i-1$ and y_{i+1} is magnitude at timestep $i+1$ [13].

To minimize the number of linear patterns, energy threshold points within three consecutive time steps (three consecutive values of load) in load curve are ignored for linearization. The energy threshold points are extracted from the raw load profile and are used for modelling the alternative load profiles. These points are taken as the operating points for the Taylor series expansion up to the first degree only. The high order terms are neglected as their effects is considered negligible for linearization. The missing data at each time step between the threshold points is linearly interpolated to create the alternative profile [13].

The linearization process results in the generation of the concatenation of continuous linear curves with discontinuous derivatives. The linearized curve can be represented by a concatenation of the tangent lines such that at operating point (threshold point) a_i, linear curve forms which can be given as in (4.4) [13].

$$y = \sum_{i=0}^{n} (f(a_i) + f'(a_i)(x - a_i)) \tag{4.4}$$

where $f'(a_i)$ gives the derivative of a_i. From (4.4), it can be seen that linearized profiles do not require higher order derivatives, as they are linear in nature. The linearized profiles simplify the complexity of load variations leaving only eminent variations in load curve [13].

4.2.2 Enhancing the Accuracy Using Particle Swarm Optimization (Stage 4 & 5)

An important aspect of demand profiling is the accuracy of representation. The accuracy is quantified by comparing energy captured by raw and alternative load profiles. The energy captured by alternative profiles in the form of (4.4) varies significantly from the raw profiles. This variance needs to be minimized such that linearity of the alternative profile is preserved and the energy of the alternative profiles equates the energy of raw profiles. In order to minimize the energy variation between two profiles, an additional factor for each linear pattern needs to be incorporated. The problem of minimization of energy difference can be formulated as an optimization problem.

Particle Swarm Optimization (PSO) an algorithm that operates iteratively to find the local best. In this case, PSO generates random numbers for each iteration which is used to determine the solution of the objective function. In each iteration, the best value of the random number i.e. resulting in the minimum value of the objective function is considered to be the local best. The local best is stored, and the iterative algorithm updates the global best by comparing with the stored local best. After predefined maximum iterations, the global best, which is the best out of all local bests, is considered as the optimum solution. Some of the advantages of PSO as a numerical solving tool over other optimization algorithms are ease of implementation, fewer parameters to adjust, enhanced memorization ability to store local and global best solutions, low computer memory requirement and fast running speed due to only primitive mathematical operators [125], [13].

The objective function of the problem is to minimize the difference of energy between the raw and alternative load profiles by determining weighting factors for alternative profiles. To get the best fit of energy, a weighting factor is determined using PSO for each individual pattern/segment rather than considering single weighting factor for longer periods of a day, a week or entire horizon of the load curve.

In order to minimize the energy variation at sub-hourly intervals, a random number is generated for each individual pattern in the alternative profile, which provides a weighting factor to minimize the objective function. The optimum weighting factor results in the minimum value for the objective function. In the fourth stage of alternative profile development in Figure 4.4, the iterative procedure is repeated for all segments of the linear load profile. The optimization problem to minimize the difference of energy between raw and alternative profiles can be formulated as a minimization problem as given in (4.5) [13].

$$Minimize\sqrt{\left(\sum y_0 - w_i \sum y_l\right)^2} \tag{4.5}$$

Subjected to:

$$0.5 \leq w_i \geq 2.0 \tag{4.6}$$

where y_o represents the raw data and y_l represents the linearized data. w_i represents the weighting factor where 'i' is a particle of the swarm. After analysing the difference in energy of all segments it is concluded that the difference in energy ranges from 50% to 200%. Therefore, a constraint is introduced which limits the weighting factor between 0.5 and 2.0 to limit the search space. Constraining the value of weighting factor reduces the processing time of PSO and the target can be achieved using a lower number of particles with fewer movements to attain the global optima [13].

At the fifth stage, the optimization process proceeds iteratively with weightage factor determination initially for only the first two threshold points (assume, A and B) i.e. first straight line (AB) only. These weighting factors are incorporated in the data points of the tangent line AB. At the next step, to avoid any data loss, the overlapping of the last threshold point (B) is considered to determine the weighting factor of the next linear segment (assume, BC).

The overlapping process does not use the raw threshold point B. After incorporating the weighting factor for the first segment (AB), the values of point B can potentially be changed and the new value of B i.e. B* is used to optimize the next linear segment (B*C). However, incorporation of weighting factor for segment B*C will again result in a change in the value of point B*. If point B* is not considered in optimization i.e. optimization is carried out for BC, there will be a sudden variation after point B. On the other hand, by considering the point B* for optimization of B*C, final optimized point of AB (B*) changes. Therefore, to neutralize the effect of overlap and avoid sharp variation, the value of B is taken as average of the value of two new points for B as determined by optimizing segments AB and B*C.

The process is carried out for all linear segments of the profile until the last threshold point is reached. The optimized profiles at the operating point a_i can be obtained by incorporating the weights w_i can be mathematically represented by (4.7) [13].

$$y = \sum_{i=0}^{n} w_i(f(a_i) + f'(a_i)(x - a_i)) \qquad (4.7)$$

Alternative load profiles can be modelled using the simplistic approach and mathematically represented by a set of linear profiles as given in (4.7). Transformation of non-linear curves into the linear ones reduces the intricacy of the data and converts the complex non-linear functions as (3.3) (given in chapter 3 section 3.1) into simplified linear functions as (4.7). Thus, reduction in data complexity with reduced data complexity and a high degree of accuracy is achieved in alternative demand profiles.

4.3 Alternative Generation Profiling

Generation profiles for conventional bulk generation systems are smooth due to very little variations in their pre-specified operations schedules. On the other hand, the RES like solar PV or wind generation tends to be intermittent. Their intermittency can cause large variations over small period of time making the generation profile highly variable and intricate, therefore complexity of these generation profiles tends to be higher. The drive for reduction in carbon emissions essentially requires penetration of more RES, which can change the power generation profiles by enhancing their variability and intricacy. In future grids, power system studies need to incorporate load at lower aggregation level and the intricacy of load and generation profiles as well as their interaction needs to be reduced [4]. Size and complexity of the load profiles are reduced using data clustering and alternative load profiling, the complexity of RES generation profiles needs to incorporate a similar approach for the reduction in complexity of the profiles.

Hence, to address the data complexity of the generation profiles of intermittent RES, alternative generation profiling approach is developed. The half-hourly solar PV generation profiles are modelled as alternative generation profile using an approach like the one used for modelling alternative load profiles. However, the PV profiles are intermittent and have long hours where the generation remains zero. The load profiles do not tend to have zero values thus the algorithm designed to model alternative load profiles, cannot be directly used for the development of alternative generation profiles. After the zero generation, in an effort to find the next threshold point, the algorithm divides the energy over the period starting from first zero to the first maxima after long zero generation hours. Therefore, in the case of the generation profiles, alternative profile development the algorithm requires to ignore the continuous zero generation duration and consider it as only two data points, i.e. first and last zero indices. The alternative profile development procedure is performed using the approach adopted for alternative load profiles and the incremental optimization of each linear pattern is performed using PSO. The ignored zeros are added to their corresponding indexes after incorporation of weighting factors. Thus, the alternative generation profiles are developed by linearly best fitting the raw generation profiles and optimizing them [14].

4.4 Application and Validation

A case study is presented to validate alternative profiles for energy classification for Monte-Carlo simulation applications. Details of the case study are presented below.

4.4.1 Alternative Load Profile

Raw load profiles resulting from the extended k-means clustering are used to assess the viability of the approach. The profiles generated for clustering scenarios of all three cases discussed in chapter 3 are used in this chapter. The quality of linearization is evaluated based on the accuracy of energy capture by the alternative profiles when compared with the raw profiles. Figure 4.5 shows sample raw and alternative load profiles for a week.

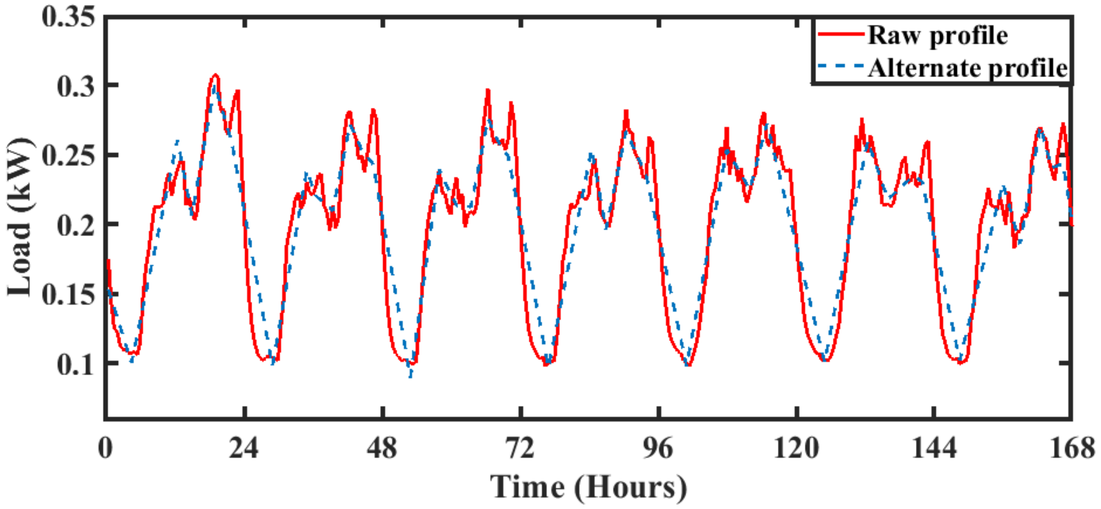


Figure 4.5 Raw vs alternative load profile for one week

From the Figure 4.5, it can be observed that alternative profile (dotted blue line) linearly fits the raw profile. The profiles for monthly and annual scenarios show similar results.

The accuracy of representation of the alternative load profiles is evaluated for all three scenarios independently. Accuracy of energy capture by alternative load profiles for weekly scenarios is given in Figure 4.6. The maximum level of error stays below 2% with an average slightly above 1%. The low error shows that the linearization process has approximated original profiles with high accuracy. The errors for each individual cluster are evaluated using error of each segment/linear pattern and average of error of the cluster profile represents the cluster error. The average error of all clusters is used to calculate the error for the entire period which is presented in Figure 4.6 [13].

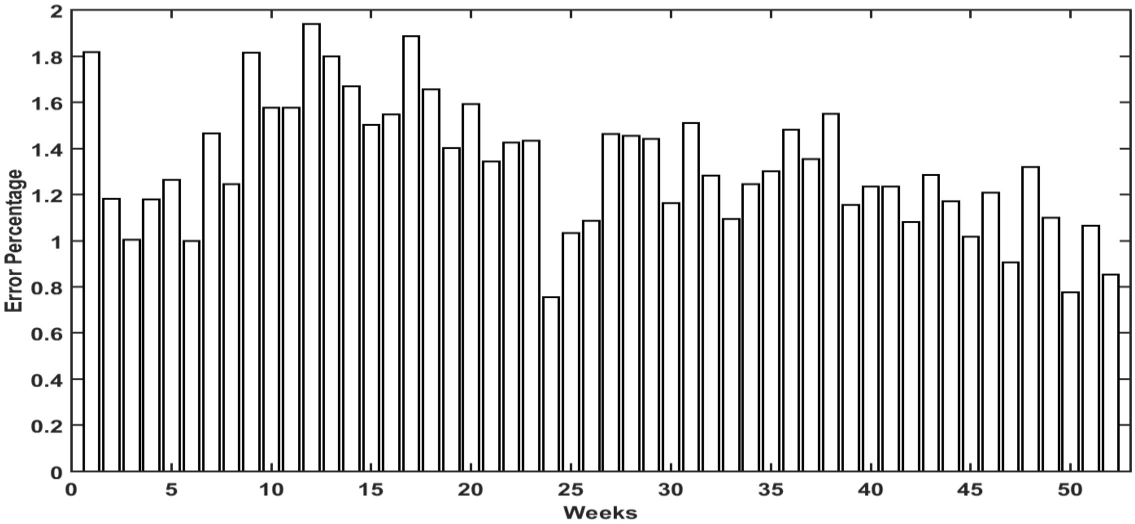


Figure 4.6 Weekly percentage error in energy capture [13]

In scenario 2, the monthly load profiles are considered for the development of alternative profiles. The errors in energy capture for scenario 2 are given in Table 4.1 [13]. The error in energy capture of alternative monthly load profiles is also low and validates the robustness of the approach [13].

Table 4.1 Weekly percentage errors in energy capture [13]

Month 1	**Month 2**	**Month 3**	**Month 4**	**Month 5**	**Month 6**
1.253	1.547	1.766	1.601	2.052	1.019
Month 7	**Month 8**	**Month 9**	**Month 10**	**Month 11**	**Month 12**
1.638	1.369	1.460	1.576	1.311	1.151

In case 3, the error in energy capture for the load profile of a year is calculated to be 1.66 % [13]. It is important to observe that the errors in the energy capture should ideally be nullified by the weighting factors determined by PSO. However, due to consideration of overlap of threshold point to mitigate the sudden variations and to avoid data loss, the value of each threshold point except for the first and last threshold points changes from optimum to an average of the two consecutive optimums. This phenomenon changes the energy balance and thus the error increases in the representation of the profiles.

4.4.2 Energy Classification and Validation for Monte-Carlo Simulation Applications

Monte Carlo simulation is the process that uses random number for calculation of the structure of a stochastic process [126] . A stochastic process can be defined as a sequence of states whose evolution is determined by random event sampling [126]. Monte Carlo method is commonly used in many stochastic processes such as power system reliability assessment, where the method estimates the indices by simulating the actual process and random behaviour of the system states [127]. The method simulates the process by treating it as a series of experiments, which are often referred to as Monte Carlo simulations.

The alternative load profiles are used to create energy classifiers. The energy classifiers are created for each alternative load profile to validate the applicability of the approach in Monte-Carlo simulation applications. For an alternative load profile, the energy classifier is designed by computing the magnitude and frequency of each threshold point. This information is used to define the energy classifiers that hold magnitude, frequency and classification numbers per individual profile [11].

Each classification number represents a magnitude of threshold point such that classification no. 1 has the highest probability of occurrence. Energy classifiers are designed for profiles of all clusters. Finally, all classifiers are combined to create a single classifier, which represents all profiles by realizing the information from all alternative load profiles. The information contained by this single energy classifier is a good representative of the entire population of smart meter data for the duration of entire year. The final energy classifier can be used in stochastic modelling applications preserving a significant accuracy, benefiting with a very little data as compared to the original data and making the process manifold faster that saves computational time [13].

Random sampling is applied to the computed information from the final energy classifier by incorporating probability of occurrence of each class. Each class represents magnitude of an energy threshold point. The results of different number of trials, as shown in Figure 4.7, reveal that the sampled and computed probabilities are almost similar. Different number of trials for random sampling are used and Figure 4.7 shows results of 10,000 trials. With the increase in number of trials, the sampled probability becomes closer to the computed probability. Thus, the results of random sampling verify the applicability of energy classifier for stochastic applications [13].

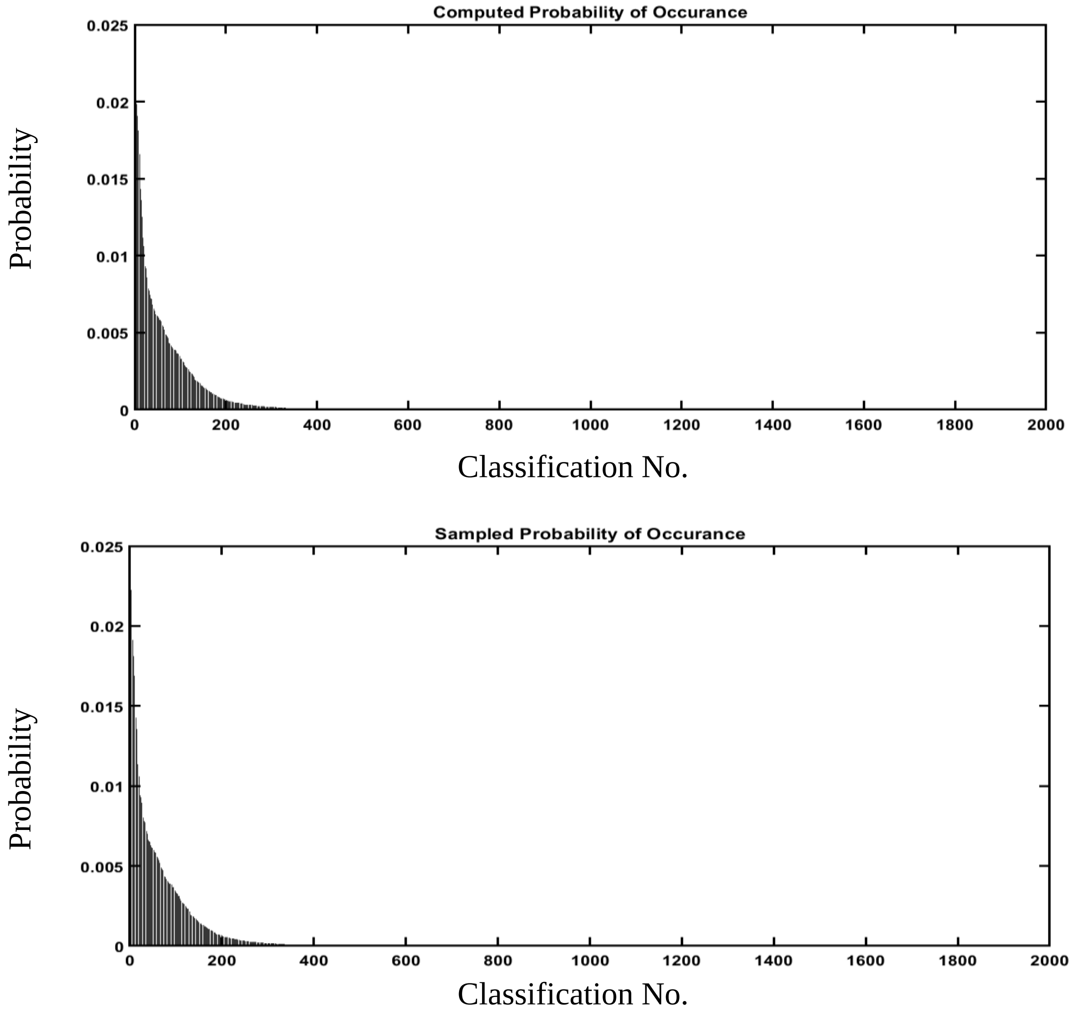


Figure 4.7 Computed and Sampled Probabilities for profiles of one year [13]

4.4.3 Alternative Generation Profiles

Half-hourly weather data including air temperature and total solar radiation was incorporated estimate PV power generation profiles [128]. Converting the raw PV generation profiles to alternative PV generation profiles resulted in high accuracy. For 1-year PV profile, the difference between the energy of raw and alternative PV generation profiles turned out to be 1.37%. Figure 4.8 shows plot of raw and alternative PV generation profiles for comparison.

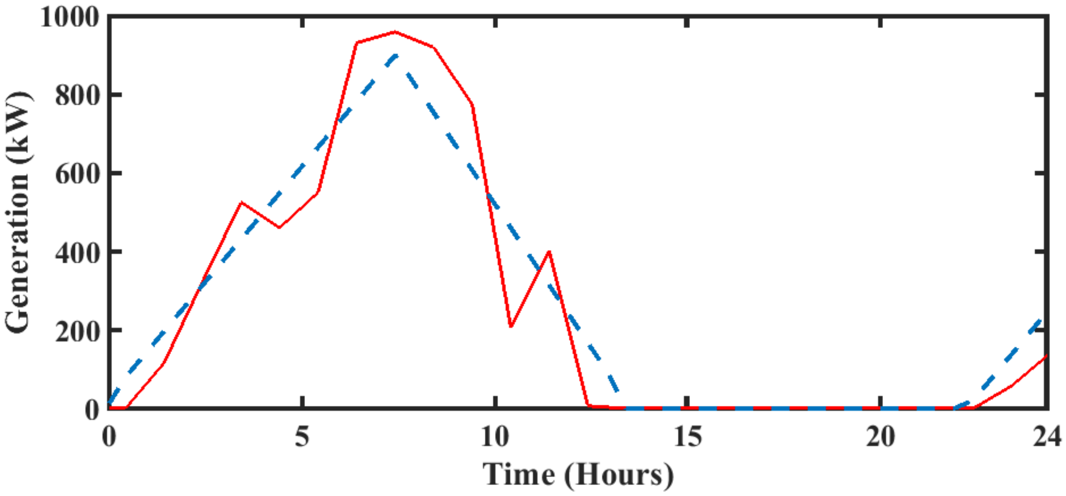


Figure 4.8 Raw and alternative PV generation profiles

(Red line shows the raw PV generation profile and the dotted blue line shows alternative PV generation profile)

From Figure 4.8, it can be clearly seen how the non-linearity of the generation profiles is reduced by linearizing and thus leaving only eminent variations in the generation profile. The variations in the PV profile are comparatively lower than the load profiles of consumers which results in better accuracy as compared to alternative load profiles. However, depending on the weather, the intermittency can cause significant variations in the PV generation profiles.

4.5 Summary

A novel approach for the development of alternative load and PV generation profiles is presented. The approach is successfully used to develop alternative energy profiles for both load and PV generation using a mathematical framework that reduces the complexity of the profiles and converts them into a concatenation of linear profiles. The accuracy of representation in alternative profiles is significantly high. Energy classifiers are introduced for the alternative profiles that contain detailed information of the load profile and use case of these classifiers for Monte-Carlo simulation applications is proved by incorporating random sampling.

CHAPTER 5

5 Alternative Load Forecasting Approach

This chapter presents an alternate approach for load forecasting at lower aggregation level in distribution network using a novel load profiling approach. The reduced complexity of the smart meter data profiles at lower aggregation level is expected to improve the forecast accuracy and reduce the computation burden. Feed forward neural networks are used to forecast load at the individual cluster level. Different load forecasting scenarios considered in this chapter include load forecasting without any renewable, with the inclusion of 10%, 20% and 30% PV penetration at the individual cluster level. Efficacy of approach is validated by comparison of forecast accuracy of the proposed approach and conventional approach for load forecasting.

5.1 Load Forecasting Using Smart Meter Data

Load forecasting has been a vital process in electricity utilities and particularly with the deregulation of the energy industry, the utilities tend to be conservative about the upgrade and expansion of existing infrastructure [129]. This is leading to stressed utilization of different power system components and consequently the importance of accurate load forecast has become paramount for secure and reliable operation and planning of power system.

Typically, the load forecast model is developed using the weather data and load history at higher aggregation level such as substation. The higher-level aggregation of load often fails to embed the different factors such as the impact of switching different appliances and consumers' energy consumption habits. Smart meter data provides two-fold benefits by enabling the utility to forecast at individual consumer level and by using the granular load information to improve forecast at high aggregate level [4].

The system load profile at higher aggregation level of load is smooth as the noise in the profile is cancelled in the aggregation process. On the other hand, at lower aggregation level, i.e. individual consumers or group of consumers, the load profiles are extremely volatile. The volatility of load profiles is driven by many factors including operational characteristics of the loads, energy consumption behaviour of the consumers, time of the day, holiday and many others that contribute to the high error in load forecast [4].

The relationship between the forecast accuracy and level of aggregation was investigate by [8]. The authors show that the forecast accuracy drops rapidly with an increase in the number of consumer when the number of the consumers is low than 1,000 and does not decrease significantly when the number of consumers is above 1,000. Thus, the forecast error depends on the spatial aggregation level of the prediction. A load forecasting study using neural network by [130] shows that mean absolute percentage error (MAPE) for 10 to 100 consumers can be between 33.6% to 11.1%. Authors of [131] also show that forecast error for smart meter can be up to 38%. This shows that at lower aggregation level, load forecasting requires different approach from the conventional approaches. Despite high forecast error at low aggregation level, the need for load forecast at these levels cannot be ignored with increasing penetration of intermittent renewable energy sources (RES). Integration of RES, particularly with intermittent nature of PV and wind generation, the stochasticity of load is enhanced which intensifies the existing uncertainties in the distribution system. This necessitates need for improved localized load forecast to reduce the uncertainty in the system directing towards to load forecast at lower aggregation level using smart meter data.

To forecast load using smart meter data, energy consumption habits of the electricity consumers using the smart meter data can be embedded using data clustering algorithms. However, the noise, volatility and variability of smart meter data are driven by the aggregation level. Therefore, reduction in noise, volatility and variability of the smart meter data can benefit the load forecast at lower aggregation level. The innovative clustering algorithm proposed in chapter 3 (Section 3.3) and novel load profiling approach proposed in chapter 4 (Section 4.2) can benefit load forecasting at lower aggregation level by clustering consumers with similar energy consumption habits and by refining the representative profiles to reduce the non-linearity, volatility and variability. The proposed and conventional load forecasting approach using smart meter data are given below in Figure 5.1.

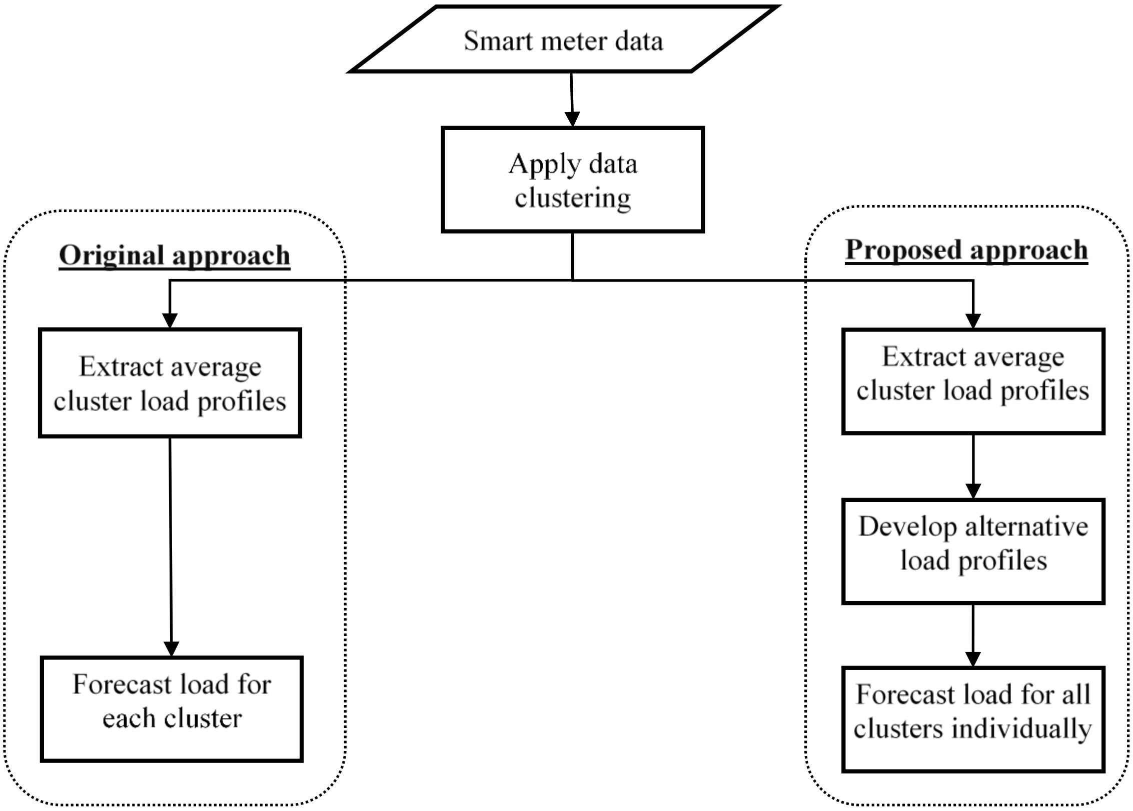

Smart meter data
Apply data clustering
Original approach
Extract average cluster load profiles
Forecast load for each cluster
Proposed approach
Extract average cluster load profiles
Develop alternative load profiles
Forecast load for all clusters individually

Figure 5.1 Conventional and Proposed approach for Load Forecasting

The proposed approach uses the extended k-mean clustering and alternative profiles which are refined form of the raw profile and are linear in nature with reduced variability and volatility. Approximating such load profiles which are linear in nature is less complicated and can be achieved with reduced computational complexity and higher precision than highly non-linear profiles (as raw smart meter profiles). To compare the forecasting capability of the proposed approach, the smart meter data is clustered using extended k-means clustering. The load is forecasted using artificial neural networks (ANN). A discussion on ANN is given in the proceeding section.

5.2 Load Forecasting Using Artificial Neural Networks

Load forecasting can be performed using different methods including many statistical and artificial intelligence based methods [64]. Artificial neural network (ANN) is artificial intelligence based method which is commonly used for load forecasting [38]. The ANNs try to learn the parameters for the optimal approximation of any continuous data function. The learning of the parameters is driven by the data [132]. The added advantage of ANNs being data driven technique makes it a suitable technique for load forecasting using smart meter data.

A simplified model of feed forward multilayer perceptron neural network is presented in Figure 5.2. The input variables are represented by X_i and the arrows connecting the input layer to hidden layer incorporate the weights of input to the layer j i.e. w_{ij}. Y represents the output which in this case will be the forecasted value of load. The weights are determined by training the neural networks to determine the global minimum of error function. The commonly used error function is sum of square of errors.

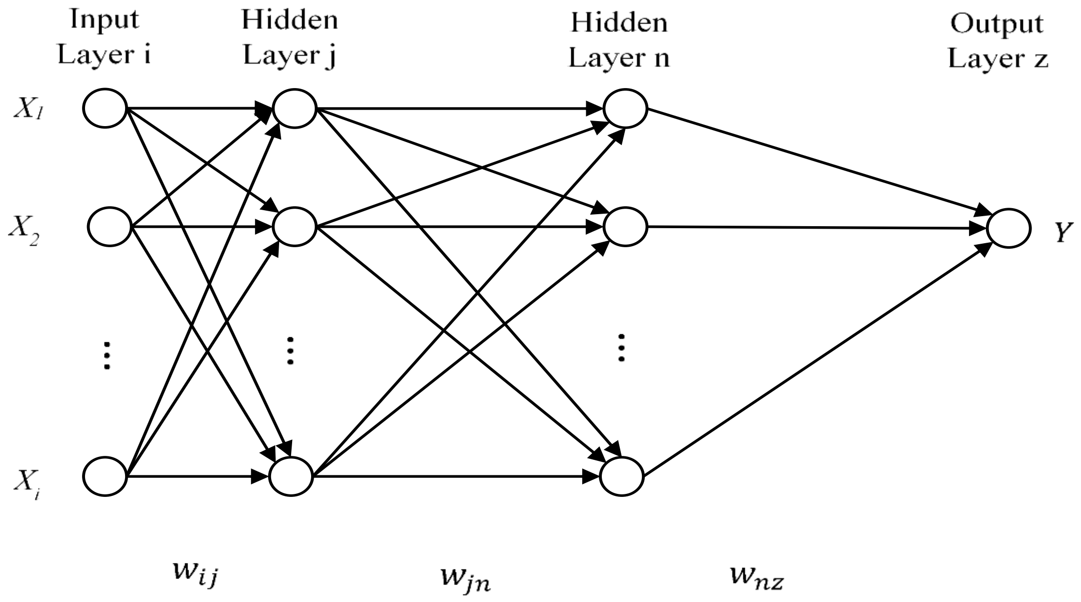


Figure 5.2 Typical Model of Multi-Layer Perceptron

The feed forward multilayer perceptron (MLP) is a commonly used ANN technique that produces satisfactory results for load forecasting applications [75]. A detailed discussion on the architecture of the neural networks is presented in chapter 2 (Section 2.3.3) and after reviewing the literature, feedforward MLP with Tanh function as activation function is selected which can be given as in 5.1.

$$g = tanh\left(\frac{u}{2}\right) = \frac{2}{1 + e^{-u}} - 1 \qquad (5.1)$$

Gradient Descent algorithm is selected due to its efficacy and ability to work on the graphical processing units (GPU) which can enhance the training speed of the neural network. The learning is carried out by back propagation (BP) algorithm where the gradient of the cost function is calculated at the output layer and is propagated back to the first hidden layer. The BP algorithm has two segments for each step of training as shown in Figure 5.3. The forward calculation segment calculates the output using the present weights and error propagation segment propagates the error from output layer towards input layer, thus weights are adjusted according to the error at each layer.

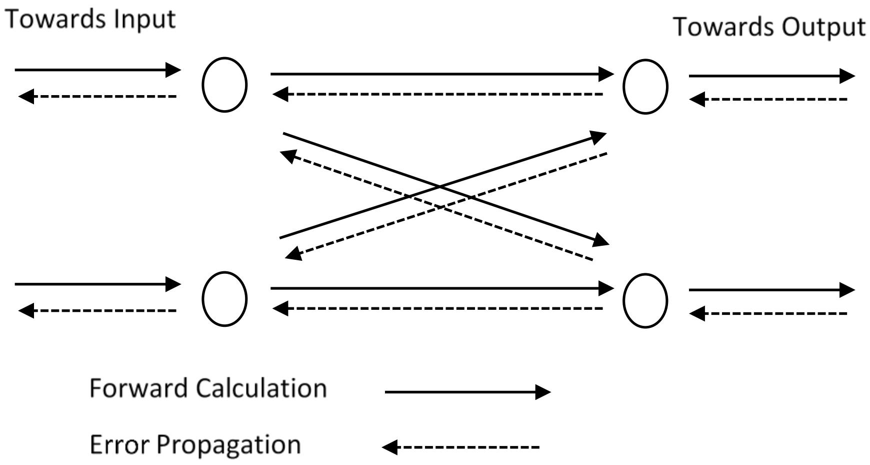


Figure 5.3 Forward Calculation and Error Propagation in Back Propagation

The cost function used for load forecasting in neural networks is sum of square of errors. There is no rule of thumb to determine the number of neurons in the hidden layer and number of hidden layers. The ability of neural network to forecast load can increase with the increase in number of hidden layers, however, a trade off should be considered for the improvement in the forecast accuracy and the computational complexity of the network. After simulating forecasting scenarios for different clusters, number of layers for ANN is selected as 4. The improvements in forecast accuracy beyond 4 layers do not provide significant benefits. The number of neurons in the hidden layers are selected through a forward method discussed in chapter 2 (Section 2.3.3). Moreover, as the initialization of the weight matrix is random, each model for the number of neurons and layers is simulated 10 times. Multiple initialization of the neural network can result in a better solution and the best network was selected based on the best out-of-sample MAPE. The optimal network is selected using the raw load profiles and the same networks are used to test the conventional and proposed approaches. The final input variables selected for final forecasting model are listed below:

- Half hourly smart meter readings (Historical load values)
    - Previous day same hour load
    - Average load of previous 24 hours
- Half hourly temperature
- Month of the year
- Day of the week
- Hour of the day

In addition to the above variables, in cases where the integration of PV is considered, an additional variable i.e. the expected value of the PV generation is also considered to include the impact of PV generation.

5.3 Load Forecasting Numerical Application

The temperature data used for this study is obtained from the Irish Meteorological Service [133]. The hourly temperature data is linearly interpolated into half hourly data. The ANN model is used to forecast short-term load for every half hour for one week. A commonly used measure for load forecast error i.e. mean absolute percentage error (MAPE) is used to evaluate the forecast accuracy [67]. The results presented in this research show MAPE for the entire week for conventional approach i.e. using raw load profiles and for the proposed approach i.e. alternative load profiles. The forecast accuracy for raw load profiles is measured by comparing the forecasted load against the future raw load profile and for alternative load profile is evaluated against the future alternative load profile.

The load is forecasted for each cluster individually by simulating four different cases. Each case considers two scenarios, where the first scenario considers training the MLP using raw profile and second scenario considers the training of MLP using alternative profile. Apart from the raw and alternative profiles of the smart meter data, the raw and alternative PV generation profiles developed in Chapter 4 (Section 4.3) are also incorporated for different PV penetration levels in this research. Incorporation process of PV profiles in the load profiles is detailed below.

5.3.1 Incorporating PV Profiles

The PV generation profiles are used to simulate the scenarios of distributed generation with different levels of PV penetration at the individual cluster level. The magnitude of PV generation for incorporation in load profile of a cluster is computed based on the average load of the cluster. The initially simulated PV generation profile is scaled between 0 and 1, which is then multiplied with the average load of each cluster during 24 hours to determine the capacity of PV generation. The resulting values of the PV profiles are multiplied with the level of PV penetration such that for power after integration of PV generation i.e. P_{res} becomes;

$$P_{res} = (P_{(i,j)} - PV_{(i,j)}) \qquad (5.2)$$

where $P_{i,j}$ is the original raw/alternative power for a cluster i at time step j, $PV_{i,j}$is the power input from the PV generation for cluster $i$at time j . $PV_{i,j}$ for a cluster i can be calculated as in (5.3).

$$PV = \left(\frac{\sum_{j=1}^{24} P_{(i,j)}}{24}\right) \times \propto \times\, PV_s \qquad (5.3)$$

PV_s is the initial simulated PV generation (scaled from 0-1) and $\propto$ determines the level of PV penetration i.e. for 10% PV penetration, $\propto$ will be 0.1. For case 3 and 4 the value of $\propto$ will be 0.2 and 0.3 i.e. 20% and 30% PV penetration levels. Moreover, the raw load profiles will use raw PV generation profile and the alternative load profiles will use alternative PV generation profile.

In each case, the forecast is generated using raw profiles and then using alternative profiles. Results of both forecasts are evaluated for comparison of forecast accuracy of proposed and conventional approaches. Different cases simulated for comparison and with scenarios of using raw and alternative profile forecasting for different cases are shown in Table 5.1 and discussed in proceeding sections.

Table 5.1 Cases for Load Forecasting with two scenarios

Forecasting Scenarios		**Case I**	**Case II**	**Case III**	**Case IV**
i	**Raw profile**	No PV	10% PV	20% PV	30% PV
ii	**Alternative profile**	No PV	10% PV	20% PV	30% PV

5.4 Case I: No PV

In the first case, raw and their alternative load profiles are considered without the integration of PV generation. The forecast for raw load profiles of all clusters is carried out individually which is followed by the forecast for alternative load profiles. Forecast accuracy in terms of MAPE for the raw and alternative profiles for different clusters is shown in Figure 5.4.

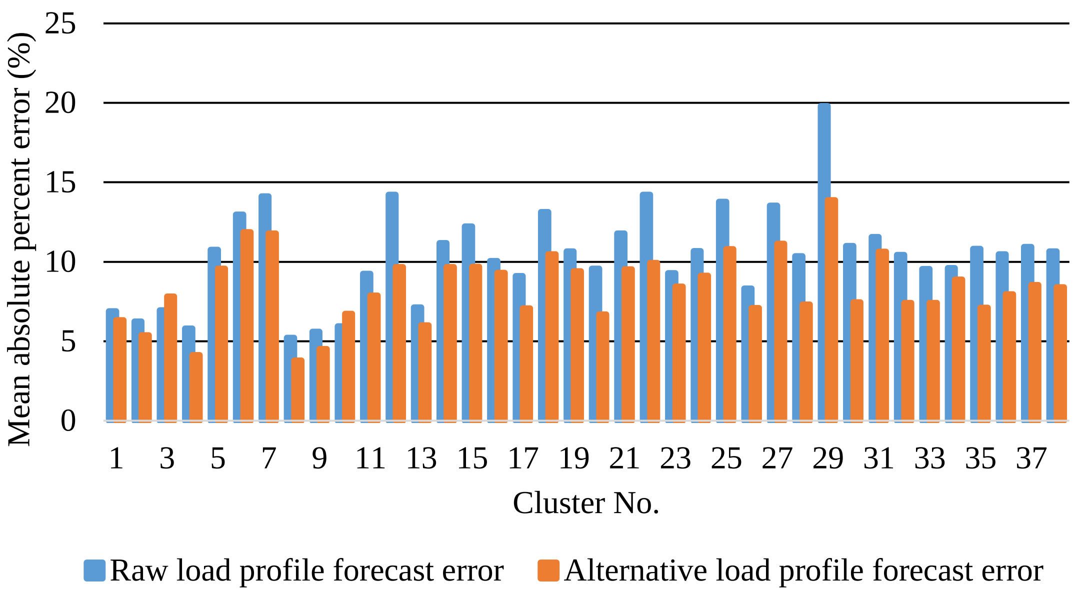


Figure 5.4 Raw and alternative profiles forecast error

Figure 5.4 shows that forecast error for clusters using raw load profiles tends to be higher than that of the alternative load profiles. Only two instances show that the cluster forecast using raw load profile is better than that of the alternative profile. The two clusters, which have comparatively higher MAPE for alternative profiles have smoother profiles due to a high number of consumers. The profiles for these clusters are smoother, as the noise in the data is cancelled in the aggregation of load demand. Moreover, the difference between the MAPE of raw and alternative profiles for these two cases tends to be lower than 0.8%, which is significantly smaller as compared to the difference between the raw and alternative forecast error for other clusters.

In the case of raw profiles, the average of MAPE for the 38 clusters is 10.36%, whereas for alternative profiles it is 8.39%. Therefore, on average the alternative profiles forecast with 1.97% more accuracy than raw profiles. Improvement in forecast accuracy using alternative profiles as compared to raw profiles is shown in Figure 5.5. From the Figure 5.5 it can be seen that cluster 3 performs worst for alternative profiles with reduction of 0.87% in forecast accuracy for alternative load profile, where the raw profile-based forecast has a MAPE of 6.95% and alternative load profile-based forecast has a MAPE of 7.82%. For cluster 29, the alternative profile shows the best improvement of 5.92% accuracy with MAPE for the raw profile being 19.80% and for alternative profile 13.88%.

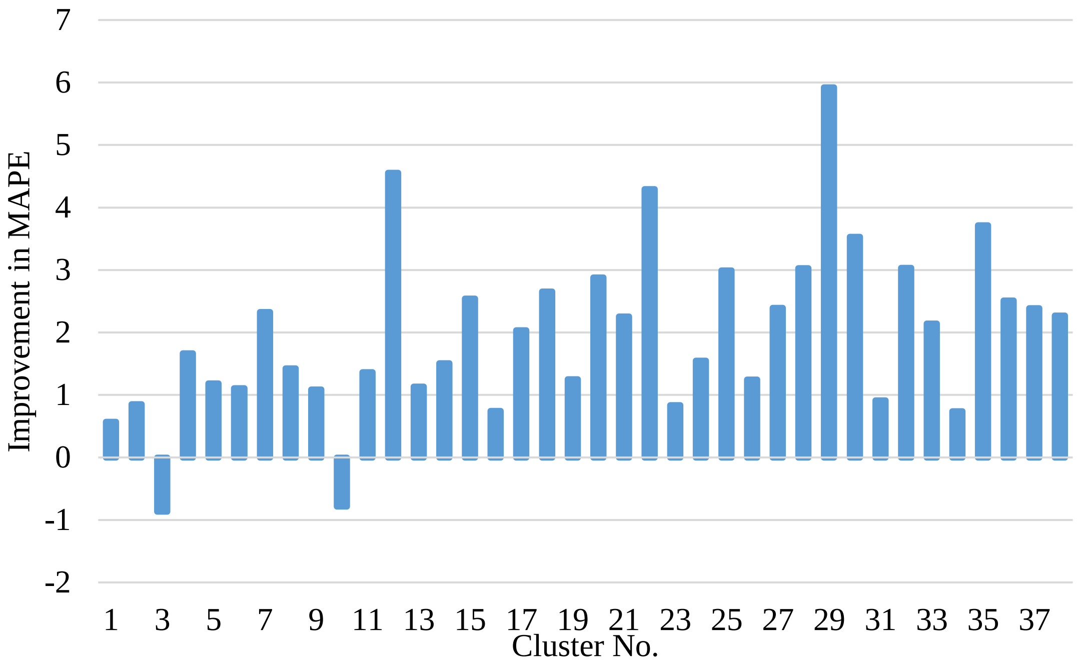


Figure 5.5 Improvement in forecast accuracy using alternative load profile

The results suggest that the proposed approach shows significant improvement in the load forecast at lower aggregation level. The reduced complexity of the alternative load profiles alleviates the noise in data, which results in the improved mapping of the relationship between load and different variables considered for load forecast. Apart from the improvement in forecast accuracy, the proposed approach has shown added advantage of the reduced computational time for neural network training.

5.5 Case II: 10% PV

In Case II, 10% PV penetration is considered by incorporating the PV generation profiles in load profiles. An important change made in the forecasting process is the addition of a new variable. The new variable is added to incorporate the PV generation variable in the MLP forecasting model. The variable contains the expected PV generation values. The values of PV generation considered in this research are the simulated values that are incorporated in the raw and alternative load profiles. By adding this variable, if the forecasting accuracy is affected, this will be due to the error in mapping the non-linearity of PV or the noise. Moreover, the intermittency of the PV generation can potentially increase the non-linearity of the profiles and can have an adverse effect on the forecast accuracy of conventional approach i.e. using raw profiles as well as proposed approach i.e. using alternative profiles.

Figure 5.6 shows the forecast accuracy for both raw and alternative load profiles. In Case II, once again the forecast accuracy for alternative load profiles outperformed the forecast accuracy using raw load profiles. However, an overall increase is observed in forecast error for both raw and alternative profile forecast. The minor increase in the forecast error for both profiles does not significantly influence the overall difference in forecast error. The average improvement in forecast accuracy for all cluster using the alternative forecasting approach marginally increases from 1.97% in Case I to 1.98% in Case II. The two clusters identified as having better performance for raw load profiles as compared to alternative load profiles showed marginally better forecast accuracy. However, the decrease in forecast accuracy for raw load forecast of these clusters, shows the impact of intermittency and non-linearity of the PV generation.

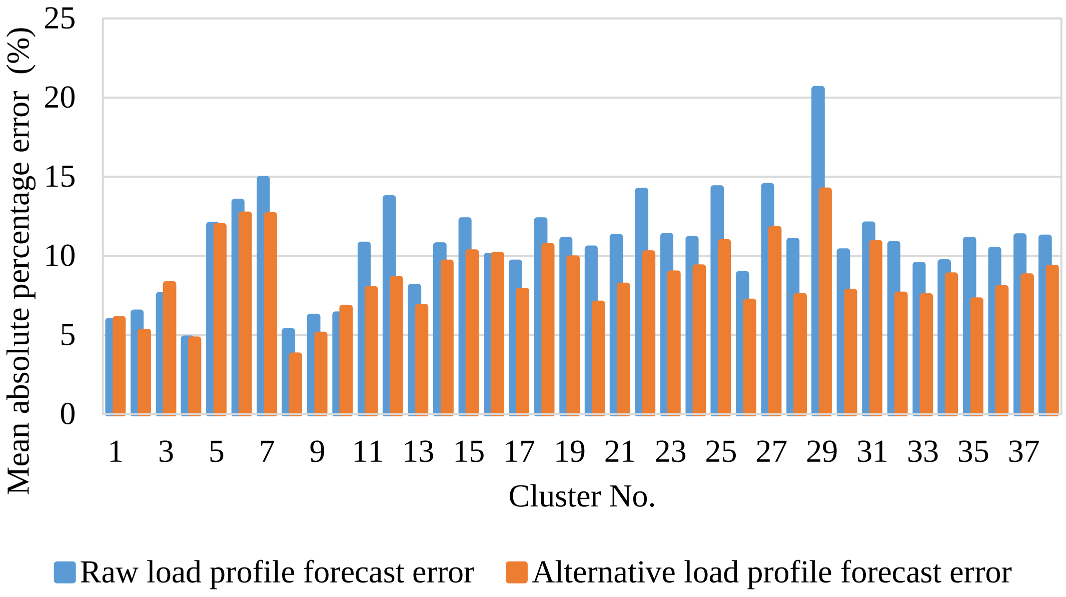


Figure 5.6 Forecast error for Raw and alternative profiles with 10% PV

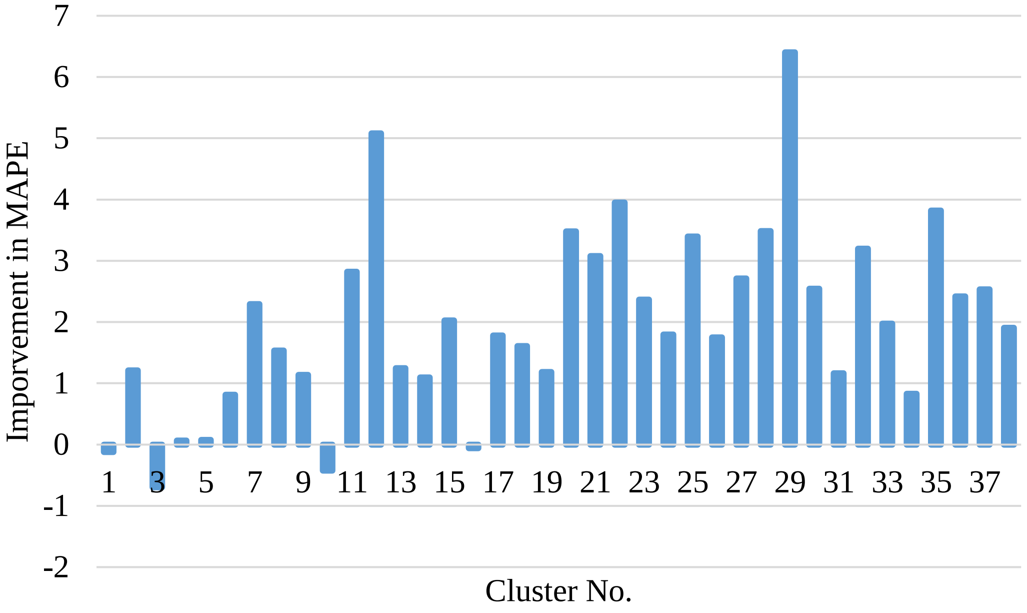


Figure 5.7 Improvement in forecast accuracy using alternative load profile (10% PV)

The improvement in forecast accuracy for each cluster by using alternative profile in case II is shown in Figure 5.7. It can be clearly seen that the best improvement in forecast accuracy has increased from 5.92% to 6.4%. Thus, the alternative profiles prove to be robust against the intermittent nature of the PV generation with improved load forecast accuracy when compared with the load forecast using raw profiles.

5.6 Case III: 20% PV

The PV penetration level in Case III is increased from 10% to 20%. The increase in penetration level of PV generation increases the non-linearity in the profile making the variability due to PV intermittency in the load profile more pronounced. This is evident from the results of the forecasting shown in Figure 5.8. The average of forecast errors for raw profiles increases to 11.07% as compared to average of forecast errors using alternative profiles i.e. 8.99%. Thus, the averaged improvement in forecast accuracy using alternative profiles increases beyond 2%. This can be seen in Figure 5.9. An overall increase in forecast error has been observed in this case for raw as well as alternative profiles. However, the variation in forecast error is not proportional to Case I or Case II. The variations are non-linear which is a consequence of the difficulty in mapping the increased non-linearity. As non-linearity of raw profiles is higher as compared to alternative profiles, the performance of raw profile in forecasting is not as good as alternative profiles. The increase in forecast error for alternative profiles is characterized by the incorporation of alternative PV generation profiles in alternative load profiles. Incorporation of alternative PV generation profiles decreases the linearity of the alternative profiles and consequently the forecast error increases for the alternative profiles. Once again, the robustness of the forecasting results for alternative profiles become clearer when improvements in the forecast accuracy are observed.

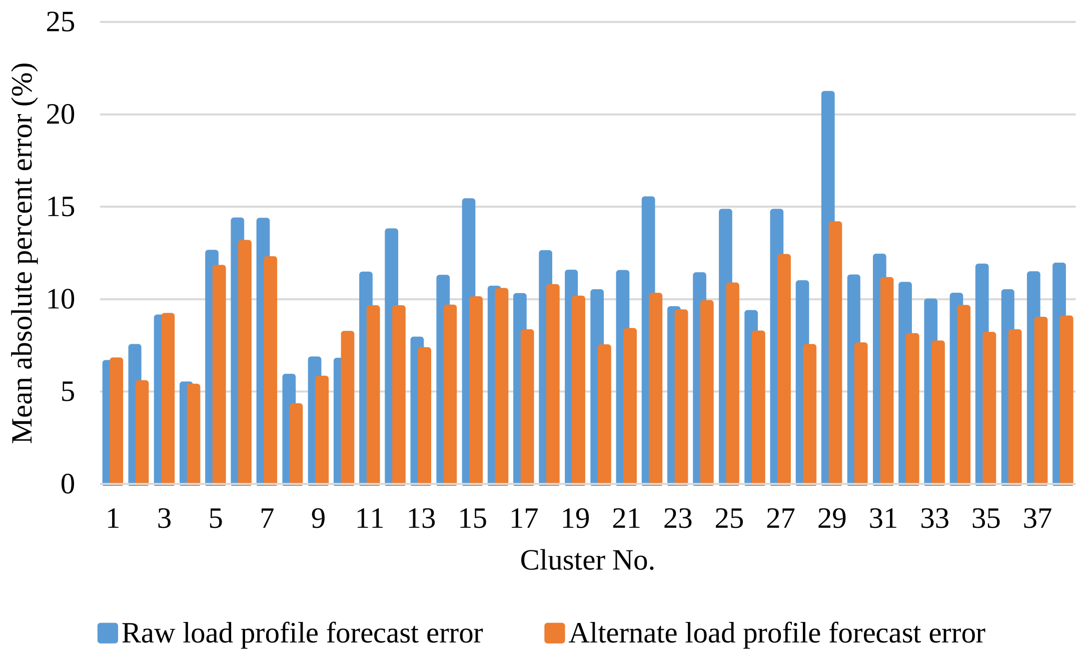


Figure 5.8 Raw and alternative profiles forecast error with 20% PV

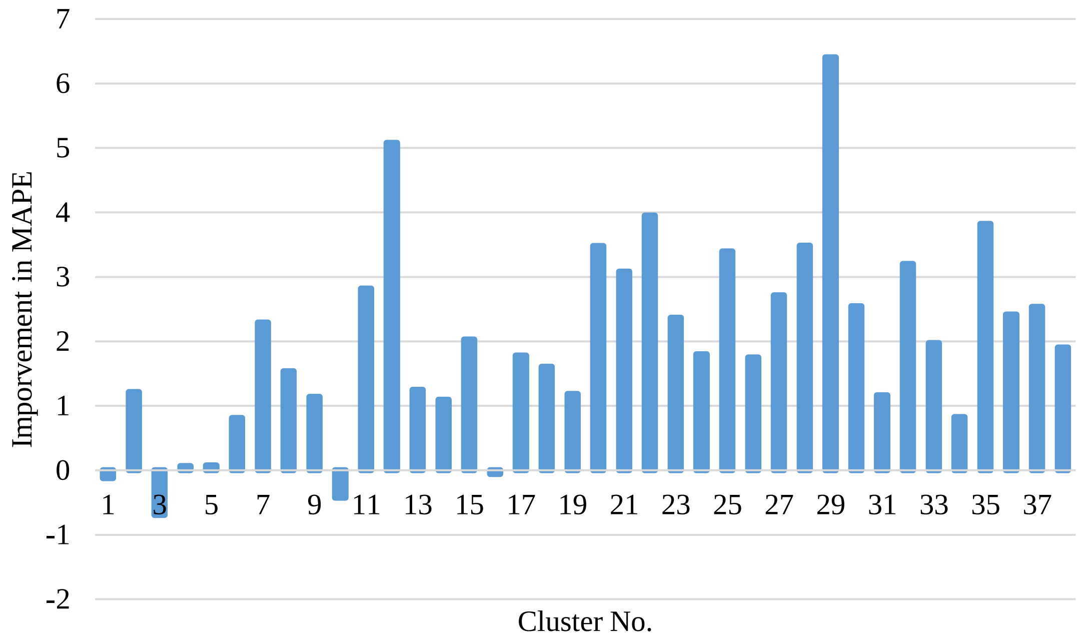


Figure 5.9 Improvement in forecast accuracy using alternative load profile (20% PV)

5.7 Case IV: 30% PV

In Case IV, the penetration of PV generation is enhanced to 30% of the average cluster load. The individual cluster forecast for raw and alternative load profiles show mostly increase in forecast error as shown in Figure 5.10. The proportional increase in error for raw load profiles is higher as compared to alternative profiles, which is evident from the improvement in forecast accuracy as shown in Figure 5.11. The average of improvement overall clusters tends to be highest of all scenarios i.e. 2.26%.

Apart from one cluster i.e. cluster 16, the alternative profiles performed better than raw profiles in terms of forecasting accuracy. In case of cluster 16, the forecasting process using raw profiles showed only 0.4% of improvement as compared to the alternative profiles. Overall, once again alternative profile forecast has proved to be robust by being less sensitive to intermittent PV generation as compared to the raw profile forecast.

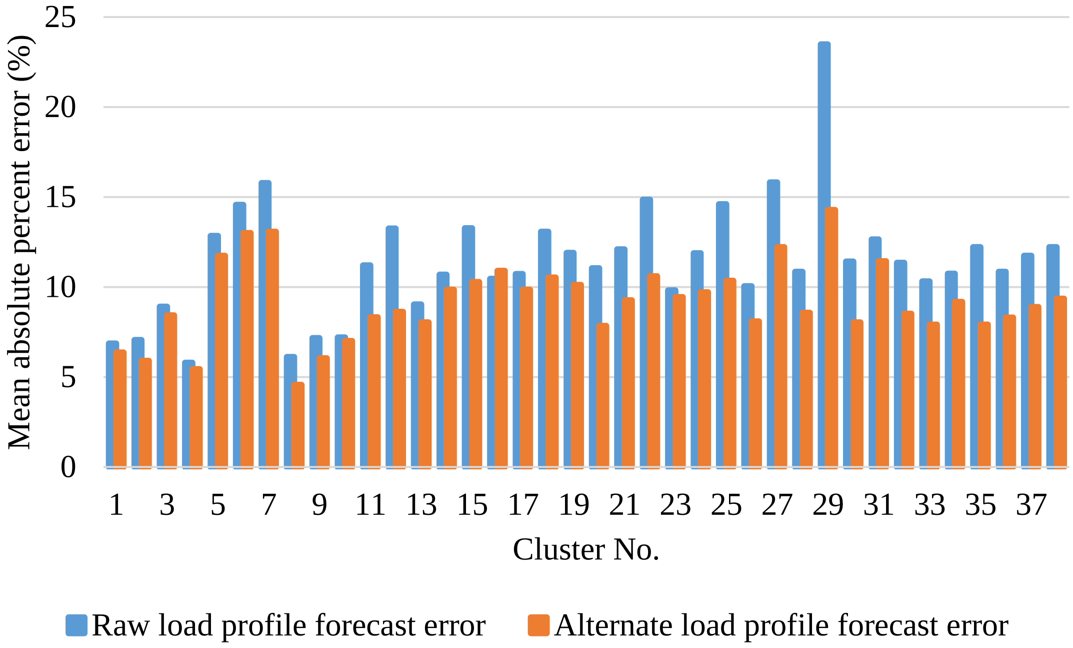


Figure 5.10 Raw and alternative profiles forecast error with 30% PV

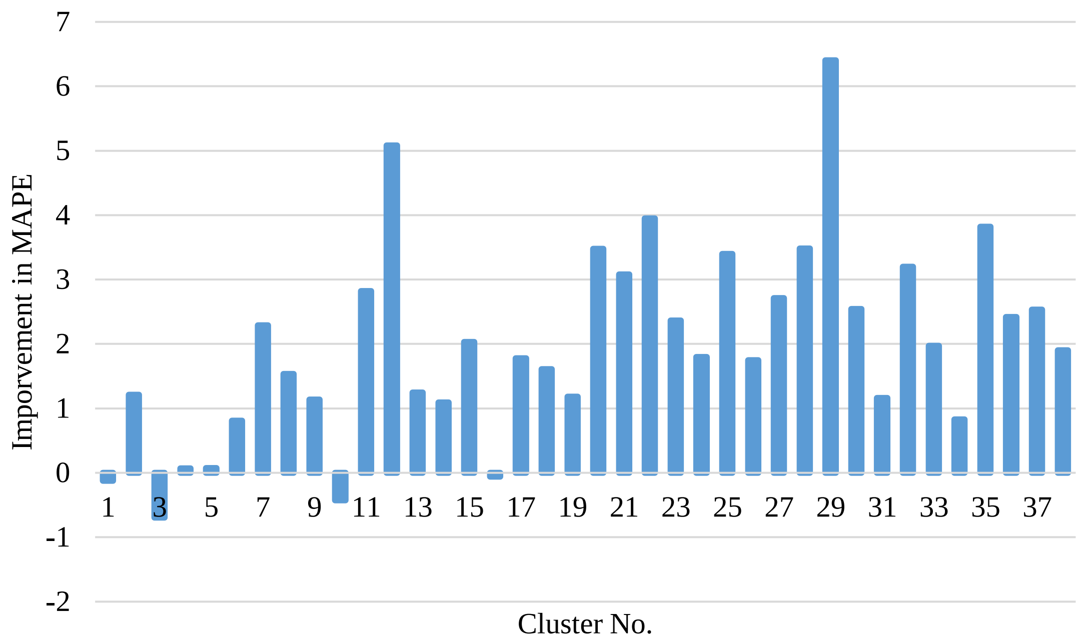


Figure 5.11 Improvement in forecast accuracy using alternative load profile (30% PV)

A comparison of all four cases shows that the forecast error using alternative load profiles is lower as compared to the one using raw load profiles. The results show that forecast accuracy is affected adversely by incorporation of PV generation profiles in load profiles. This is due to the intermittent nature of PV profiles, which introduces non-linearity in load profiles. Although, the alternative generation profiles are concatenation of linear profiles, upon incorporating in alternative load profiles, the intermittency of the PV generation adds some level of non-linearity to the alternative load profile. Overall, the forecast accuracy of alternative profiles is not greatly affected by the incorporation of PV generation profiles. They remain robust in terms of increase in forecast error. For alternative load profiles, the maximum forecast error for any cluster in all cases ranges from 13.88% to 14.28%. On the other hand, for raw profiles the maximum forecast error for clusters ranges from 19.80% to 23.49%. This show that the forecast accuracy for scenarios including PV generation can be increased by using the alternative approach.

From the case studies, it is observed that in all scenarios, the alternative load profiles provide higher forecast accuracy as compared to the raw load profiles. The high forecasting accuracy of the alternative profiles is characterized by the linear nature of the profiles. As clusters used in the case studies contain different number of consumers ranging from as little as 10 to more than 600, the forecast accuracy for each cluster varies as well. At lower aggregation level, the volatility of smart meter does not allow the ANN to map optimum relationship between the input and output. This is clear from the results in case of raw load profiles. On the other hand, at the same level of aggregation, the linear nature of the alternative load enables better mapping of the input variables to the output by reducing the variability of the load profiles. This enables the alternative load profiles to produce better forecast accuracy as compared to raw load profiles.

The disparity between the forecast accuracy of the two approaches is significant in cases with low number of consumers. In cases where the number of consumers is high, the noise in the data for raw load profiles is cancelled in load aggregation process. Consequently, the raw load profiles are smoother, and, in some cases, they outperform the alternative load profile forecast. However, the non-linearity of the PV generation can greatly affect the forecast accuracy of the raw load profiles. The results have shown that as compared to raw load profiles, the alternative load profiles are less sensitive to different levels of PV penetration and the decline in forecast accuracy for alternative load profiles is not as eminent as for raw load profiles. Moreover, the training time for the ANN for both approaches suggests that on average for alternative load profiles, ANN are up to 20% faster than that of raw load profiles. The linearity of the alternative profiles enables quicker learning for neural networks. Computational efficiency for quicker process turns out to be an added merit of the alternative forecasting approach. Therefore, the proposed alternative approach for load forecasting using alternative load profiles perform better in forecasting using smart meter data at lower aggregation level.

5.8 Summary

This chapter presents an alternative approach based on alternative profiles for load forecasting. The alternative profiles are concatenation of linear profiles with reduced complexity, variability and volatility. The forecast performance of the proposed approach is compared with the conventional approach i.e. using non-linear raw profiles using ANN. The results show that the proposed approach performs better in terms of forecast accuracy due to reduced variability and complexity of the profiles. The alternative forecasting approach shows more robust and stable behaviour even in the presence of PV profiles whereas, the raw profile-based forecast is significantly affected with incorporation of PV generation. Apart from the forecast accuracy, the training time required for alternative profiles is lower as compared to training raw profiles. Overall, the results suggest that the proposed approach is better than conventional approach in terms of forecasting accuracy, robust to variability of PV profiles and computationally efficient for load forecasting.

CHAPTER 6

6 Demand Side Management Using Smart Meter Data

This chapter presents a novel approach for demand side management (DSM) in a power distribution system. The approach uses smart meter data of consumers in power distribution systems to create clusters of consumers by using extended k-mean clustering algorithm. Load profiles of these clusters are modelled into alternative profiles that are used to forecast load. The forecasted load profiles are used for the DSM at cluster levels. A new cluster selection index is developed for the selection of clusters to handle the challenge of scalability of approach with a large number of consumers and to incorporate the impact of uncertainty in the load forecast of the cluster. Novelty of the proposed DSM approach originates from the combination of the extended k-means clustering, alternative load profiling, alternative load forecasting and new cluster selection index. The case studies for the proposed approach incorporate the impact of varying demand flexibility for each cluster in combination with varying penetration levels of PV. Impact of the cluster level DSM on the system profile is also studies with a different number of clusters. The demand is managed via load shifting in response to time-of-use tariffs through linear optimization. Comparison of the proposed approach with the conventional approach is carried out in terms of cost saving for consumers and load peak reduction.

6.1 DSM in Smart Grids

The status quo of centralized generation system is being replaced with a paradigm within which flexibility is paramount for secure and reliable operation of the power network. Electricity markets are established to use generation resources more effectively. With the decentralized generation and distributed network operators (DNOs), the interconnected system can potentially help to maintain the fluctuations and the uncertainties in the demand. Flexibility in demand can help in the operations of the electricity market [134]. The underlying principle behind implementation of DSM, within the context of smart grids, is to improve the system efficiency, security, reliability and sustainability [87, 135]. Efficient DSM can potentially optimize the utilization of existing infrastructure and support deferral of construction of new infrastructures for generation, transmission and power distribution systems [87].

Apart from the deferral in construction of new infrastructure, efficient management of demand can also benefit the smart grids in operational efficiency and financial gains. For example, five to eight percent of the installed generation capacity in Europe only handles load peak, which occurs only one percent of time [82]. Often the generation plants handling peak load, for example diesel generators etc., are expensive to operate and are not environment friendly. Reduction in the peak load by changing the shape of load curve can potentially save huge financial investments and optimize the system efficiency. Moreover, application of efficient DSM can provide socio-economic benefits such as reduction in carbon dioxide emission, integration of renewable energy and cost savings for DNOs and consumers.

DSM in the current smart grids comprises of load monitoring, analysis and response. Advanced bidirectional communication infrastructure can provide an opportunity to DNOs to improve operational efficiency from the real time load monitoring and control, whereas the consumers can optimize their energy use to achieve monetary benefits. However, implementation of DSM requires prior knowledge about the expected load profile of the consumer and sophisticated coordination between the consumer and utility. For a network having large number of consumers, scalability of direct DSM application is a major challenge. Moreover, the size, high dimensionality, volatility and heterogeneity of the of smart meter data [46] pose great computational challenges for DSM application at consumer level.

The size of the profiles can be managed using appropriate computational techniques to cluster similar load profiles which are extensively discussed in chapter 2 (Section 2.3) and an innovative clustering algorithm is presented in chapter 3 (Section 3.3). Clustering the load profiles reduces the size of data but intricacy of high dimensionality, volatility and variability and non-linear nature of energy consumption patterns pose challenges for linear processing of the data.

In order to address the non-linearity for DSM optimization using linear techniques, a piece-wise linear cost function is introduced by [136] and many other studies focus on linearizing the constraints for DSM optimization. None of the prior studies tried to address the non-linearity of profiles for linear optimization application. Augmenting the linearity of profiles can potentially provide an improved DSM solution using linear technique.

Considering the challenges discussed above, this research presents a novel holistic approach for DSM in smart grids using smart meter data. The proposed approach uses smart meter data to generate typical load profiles by clustering the smart meter data (Chapter 3, Section 3.3). The cluster profiles are systematically linearized and modelled as alternative profiles (Chapter 4, Section 4.3). These alternative profiles are used to forecast cluster load (Chapter 5, Section 5.3). The load forecast is used to plan a day ahead DSM plan. A new cluster selection index is proposed to reduce the scalability of the problem. Details of the approach are given in 6.2.

6.2 Demand Side Management Algorithm

Optimization of electricity consumption using DSM is an important feature to control the peak of the load in a smart power grid. The shape of load curve can be changed by exploiting the latent demand flexibility using DSM. The important decision that needs to be made by the operator is, what kind of modifications are required in the load shape to handle the peak? Load shape can be varied considering different objectives using different techniques for example peak clipping, load shifting, strategic load growth etc. [80].

Due to its effectiveness, load shifting is one of the most commonly used and widely applied load management technique in power distribution systems [87, 88]. This technique shifts the load from peak hours to off-peak hours by applying the time independence of the load. Consumers implement load shifting by switching their deferrable loads on or off in response to some signal from the utility which most commonly is a price variation signal.

Application of DSM using load shifting at consumer level has become possible only with the advent of advanced metering infrastructure (AMI). Smart meter provides the opportunity to identify potential individual consumers for application of DSM, by providing granular information about their energy consumption. Although smart meters provide granular information about the load of consumers, number of consumers makes the scalability of such direct DSM application challenging task. To reduce load peak, a possible solution is to identify specific consumers, who contribute to the peak of load, amongst large number of consumers and manage demand of selected consumers only.

As discussed in literature review (Chapter 2, Section 2.4), some researchers have adopted different methods to identify potential consumers for DSM application based on peak load of the consumers. However, prior knowledge about the cluster load for DSM is possible with load forecast which can have significantly high errors at lower load aggregation level and can create great uncertainty in the load profile of cluster. Therefore, impact of the error in forecasting needs to be included while selecting clusters for DSM application. A new algorithm is presented in Section 6.2.1 for cluster selection while considering the impact of forecast uncertainty in this selection.

6.2.1 Cluster Selection

The primary objective of the DSM is reduction of system peak load and operational cost [80]. Different dynamic pricing schemes (discussed in Chapter 2, section 2.4) can be used to implement load shifting such as real time pricing, critical peak pricing, time-of-use pricing. This study incorporates time-of-use tariff as it does not require consumer to be very pro-active and even consumers without home energy management system can benefit from this tariff [64].

As discussed above, one of the most important aspects of DSM application in power distribution systems using the smart meter data is the selection of the appropriate consumers. The appositeness of the consumers is decided by the cluster load profiles. For instance, all clusters that have high load during peak hours can be considered as suitable clusters. The clusters of consumers can be selected using different indices based on energy consumption, for example based on overall maximum energy consumption, high-energy consumption during peak hours etc. A cluster having overall maximum energy consumption or the one with highest energy consumption can be used for DSM to achieve greater changes in load. However, such criteria can also result in selection of clusters with a large number of consumers and consequently either ineffective response to demand signals or consumer discomfort can arise. Moreover, it can result in greater uncertainty in response as well. Therefore, the aim of selection of the clusters for DSM should be to cause limited disruption to the consumers and avoid customer fatigue while ensuring the requisite reduction in peak with increased certainty in demand response can be achieved.

The objective of the cluster selection algorithm should allow DSM to achieve maximum cost saving with minimum consumer disruption and reduction in peak with increased certainty. To achieve this, an algorithm (given in Figure 6.1) is proposed for cluster selection and DSM application. The algorithm selects clusters based on an index, which considers energy density in each cluster during the peak load hours and impact of the forecast uncertainty on energy density. Energy density will enable selection of consumers having low numbers but high load during peak hours. As the first part of the cluster selection algorithm, the peaks hours are identified from the system load curve. In this study, the periods with load beyond 85% of the maximum system load are considered peak hours. The peak is the result of combined load by all clusters as given in (6.1);

$$P_T = \sum_{k=1}^{n} PC_k \qquad (6.1)$$

where P_T represents the total energy consumed by the entire system during peak hours and PC_k represents the energy consumed by cluster k during peak hours. Average energy consumed by each member of cluster 'k' during peak hours is quantified as in (6.2);

$$EC_k = PC_k / N_k \qquad (6.2)$$

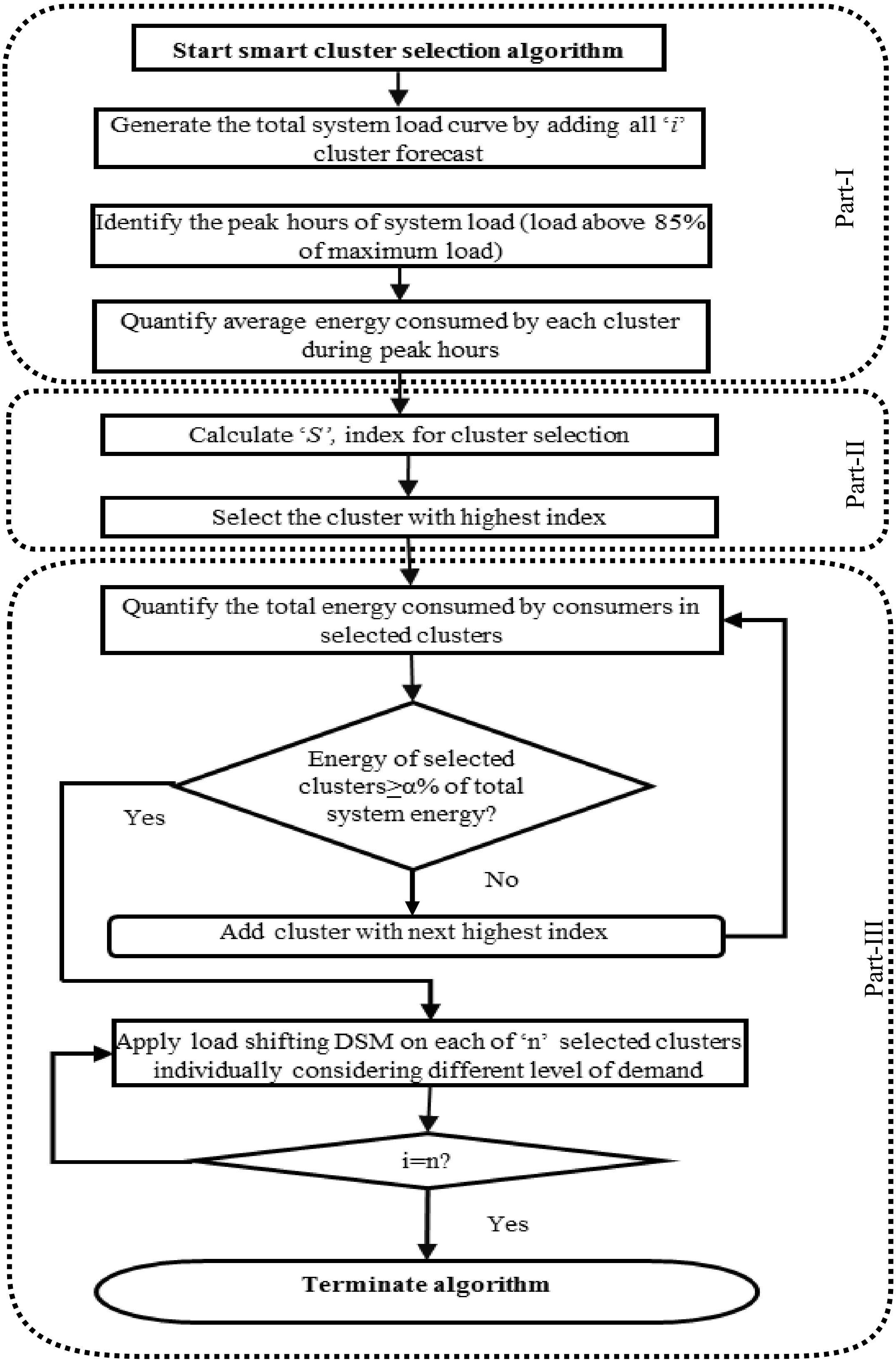

Start smart cluster selection algorithm
Generate the total system load curve by adding all 'i' cluster forecast
Identify the peak hours of system load (load above 85% of maximum load)
Quantify average energy consumed by each cluster during peak hours
Part-I
Calculate 'S', index for cluster selection
Select the cluster with highest index
Part-II
Quantify the total energy consumed by consumers in selected clusters
Energy of selected clusters≥α% of total system energy?
Yes
No
Add cluster with next highest index
Apply load shifting DSM on each of 'n' selected clusters individually considering different level of demand
i=n?
Yes
Terminate algorithm
Part-III

Figure 6.1 Smart cluster selection algorithm for DSM

EC_k represents the average energy of cluster k during peak hours and N_k is number of consumers in cluster 'k'. The second part of the algorithm computes the selection index S for a cluster k is using (6.3).

$$S_k = EC_k + EC_k \times (MAPE/100) \qquad (6.3)$$

The cluster selection index S incorporates the impact of the MAPE by adding the energy difference due to the MAPE. MAPE is the mean absolute percentage error in forecast taken from the available historic forecast error of each cluster (from Chapter 5, Section 5.3). The index S helps in selection of clusters with high energy density, low number of consumers and high uncertainty in forecasting. Selection of clusters with high MAPE can help in reducing the uncertainity of load forecasting by managing the load of selected consumers. Clusters are selected using the selection index S and are prioritized for selection based on highest value of S.

In the 3rd and final part of the algorithm, initially a single cluster is selected, and the selected cluster is evaluated for the percentage of the system energy. If the energy of the first cluster is less than α percentage of the system energy, cluster with the next high value of the index is added to the selection. The decision of value of α is made in consideration with two factors. Firstly, the percentage of peak that needs to be reduced. Secondly, the flexibility in demand at consumer end. Impact of the consumer participation in DSM can be quantified using different levels of demand flexibility, thus it affects selection of value of α. Upon addition of each cluster in list of selected clusters, load magnitude of selected cluster is checked to be a minimum of α percentage of the system load. This ensures that when multiplied with the demand flexibility variable a minimum of $\alpha \times 0.1$ percent impact on the system load can be achieved (where 0.1 represents 10% demand flexibility). So, if demand flexibility is lower, the magnitude of load can be increased i.e. more clusters can be considered for DSM to reduce peak and vice versa. Different demand flexibility levels ranging from 10% to 90% in combination with different levels of load are studied in this research. Once the selection is finalized, the selected clusters are applied with DSM.

6.2.2 DSM by Load Shifting

The objective of DSM application is set to minimize the energy consumption during peak hours in response to energy price so that maximum saving for the consumer can be attained. DSM optimization problem is formulated as a cost minimization problem. The optimization problem can be mathematically formulated as in (6.4);

$$Min. C = \sum_{i=1}^{k}\sum_{j=1}^{24} P_{(i,j)} \times Cd_j \qquad (6.4)$$

where C represents the total cost of energy consumed during 24 hours, $P_{(i,j)}$ represent the load demand of cluster i at time j. The price of electricity at time 'j' is given by Cd_j.

Subject to:

Equality constraint:

$$\sum_{i=1}^{k}\sum_{j=1}^{24} Pnew_{(i,j)} = \sum_{i=1}^{k}\sum_{j=1}^{24} Pold_{(i,j)} \qquad (6.5)$$

where $Pnew\ and\ Pold$ refer to total energy after and before DSM respectively. The equality constraint introduces here is an energy constraint. It ensures that the inequality constraints can apply the demand flexibility using the power values, however, the energy balance before and after DSM is maintained.

Inequality constraints:

Demand flexibility lower bound;

$$Pnew_{(i,j)} \geq Pold_{(i,j)} \times df_l \qquad (6.6)$$

df_l is the demand flexibility ranging from 0.90 to 0.05 in decreasing steps of 0.05. This sets the lower bound according to the required demand flexibility.

For off-peak hours;

$$Pnew_i \geq Pold_i \qquad (6.7)$$

Demand flexibility upper bound level;

$$Pnew_{(i,j)} \leq Pold_{(i,j)} \times df_u \qquad (6.8)$$

where df_u is the demand flexibility ranging from 1.10 to 1.90 in steps of 0.05. This sets the upper bound according to the assumed demand flexibility. Any changes made in the upper bound will correspond to a similar level of change in lower bound.

For peak hours,

$$Pnew_i < Pold_i \qquad (6.9)$$

This constraint ensures that the peak load does not exceed the original peak load. An additional constraint that is considered for Case II of the case studies presented in this research limits the upper bound to 95% of the original load i.e.

$$Pnew_i \leq 0.95 \times MaxPold \qquad (6.10)$$

where $MaxPold$ is the maximum value of the $Pold$.

DSM scenarios are also considered where different levels of PV penetration are considered. The optimization objective for the PV integrated scenarios can be formulated as;

$$Min.C = \sum_{i=1}^{k}\sum_{j=1}^{24} Pres_{i,j} \times Cd_j \qquad (6.11)$$

where $Pres_{i,j}$ is the energy after renewable energy integration and given in (5.3) (Chapter 5, Section 5.2.1).

6.3 Application of Novel Demand Side Management Approach

As a first step of the DSM application, the clusters are selected for DSM application. The forecasted profiles from the clusters simulated in chapter 5 (Section 5.3) are used to calculate the cluster selection index. All forecast scenarios discussed in chapter 5 (Section 5.3) i.e. load with and without different levels of PV penetration are considered. These scenarios are used to simulate two different constraint-based cases for DSM optimization with varying levels of demand flexibility and load. The details of cases are given in Section 6.3.2.

6.3.1 Cluster Selection Results

For the DSM application, the forecasted load of one day is considered for each cluster's raw and alternative profiles. The cluster profiles for a particular day are aggregated to generate a consolidated system profile that can be used to observe the impact of DSM participation by consumer of the selected clusters. The clusters are selected using the cluster selection index S based on system load profile of the day of DSM application.

One of the most important aspect of cluster selection index is the selection of clusters having high load density during peak hours that results in selection of low number of consumers for DSM participation. The low number of consumers and high cluster energy density will result in higher flexibility in load demand with a lower number of consumers to disturb. This is reflected in Figure 6.2 where the impact of proposed cluster selection index is represented in terms of percentage of peak load contribution against the percentage of consumers. Selected clusters given on x-axis are sorted in descending order of magnitude of S and cumulative percentage of population and contribution in peak load is given the on y-axis. Figure 6.2 shows an exponential rise in load consumption during peak hours with the first three clusters selected by S. The first three clusters having less than 3% of total consumers constitute more than 18% of the total energy consumption during peak hours. This shows that the proposed cluster selection index can effectively select clusters having a low number of consumers which are big contributors to the peak of the load. It can benefit DSM to achieve greater flexibility with minimum consumer discomfort thus higher certainty in response.

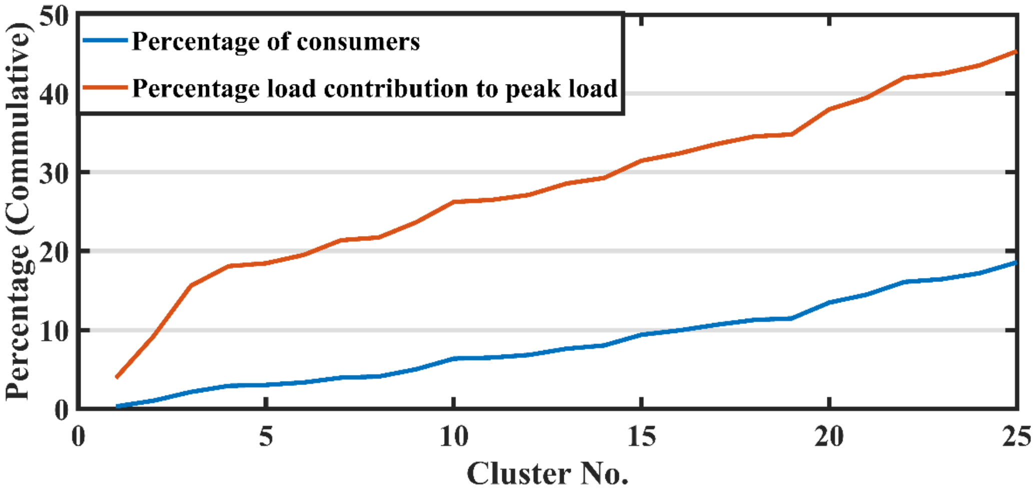


Figure 6.2 Cumulative percentage of consumer numbers against load contribution

Moreover, Table 6.1 shows the impact of including MAPE in cluster selection. With the inclusion of MAPE in cluster selection, clusters with higher MAPE or uncertainty are prioritized over clusters with lesser uncertainty.

Table 6.1 Impact of MAPE on cluster selection

Cluster No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
***S* with MAPE %**	134.36	72.52	58.48	**43.33**	**41.33**	35.88	35.33	33.81
***S* without MAPE %**	126.36	69.64	54.24	**35.92**	**36.27**	32.10	34.04	30.83

Without considering the MAPE in cluster selection, cluster 5 is selected as it has higher value of '*S*' as compared to cluster 4. On the other hand, if MAPE is included in cluster selection, cluster 4 is prioritized over cluster 5. The benefit of inclusion of MAPE can be observed from Figure 6.3 where cluster 4 presents a better shaped cluster for DSM participation as compared to cluster 5. Although the cluster 4 has comparatively higher number of consumers with almost the same level of energy density, the coinciding peak of the cluster 4 with system peak makes it more suitable for DSM. Similar cases are also observed for cluster 6 and 7 as well where better load shape is prioritized only due to incorporation of the forecast uncertainty. Thus, the inclusion of MAPE helps in selecting better clusters.

The results of cluster selection from Figure 6.2 and 6.3 suggest that the proposed cluster selection index works effectively and including the forecasting uncertainty can benefit the selection of clusters for DSM.

The original system load profile along with the system profile without the 10 selected clusters is presented below in Figure 6.4. From the figure, it can be clearly seen that the selected clusters closely follow the shape of system profile which results in a uniform reduction in the system profile, except the from 09:00 hours to 1400 hours. This indicates that the usual trend in the system load is partly set by these selected clusters and the proposed cluster selection index can significantly reduce the peak of the load.

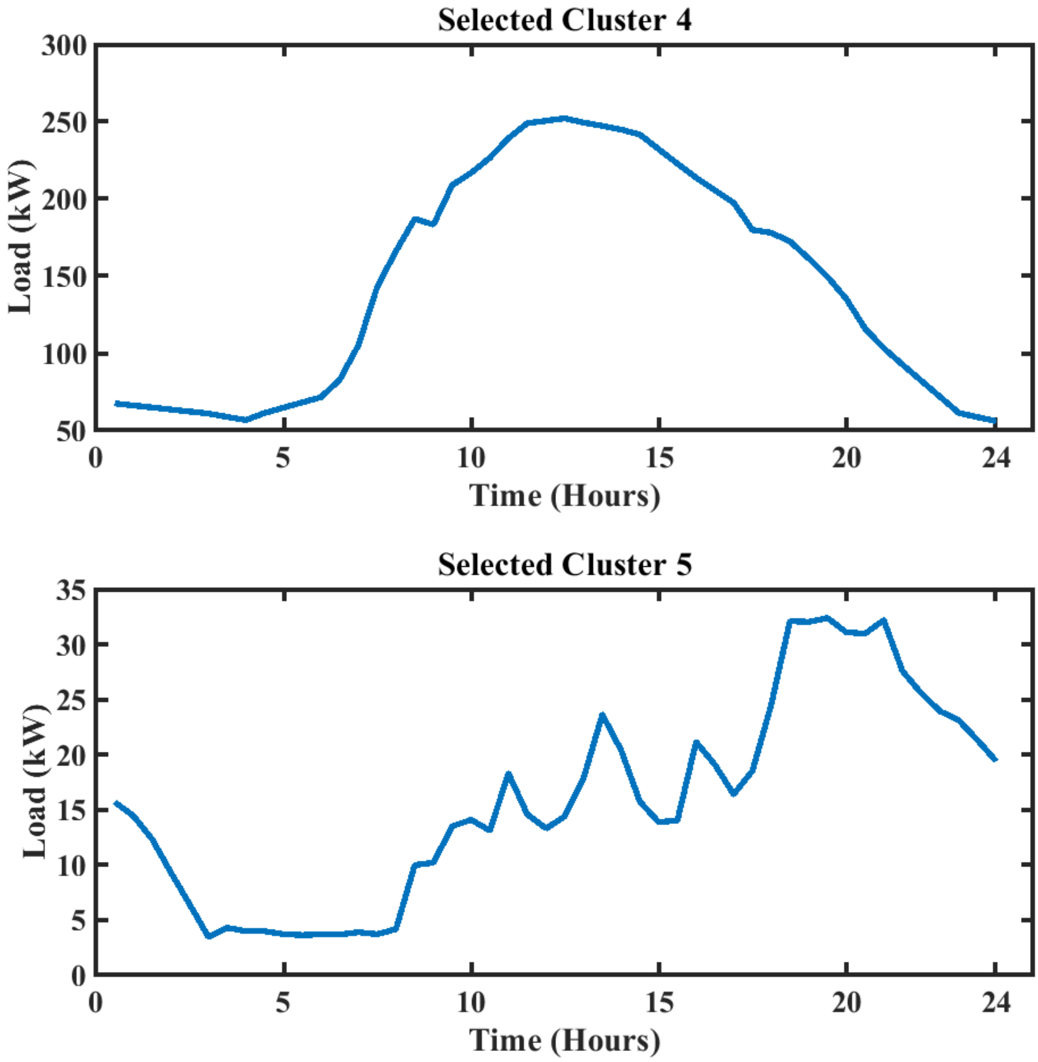


Figure 6.3 Clusters selected with and without forecast uncertainty

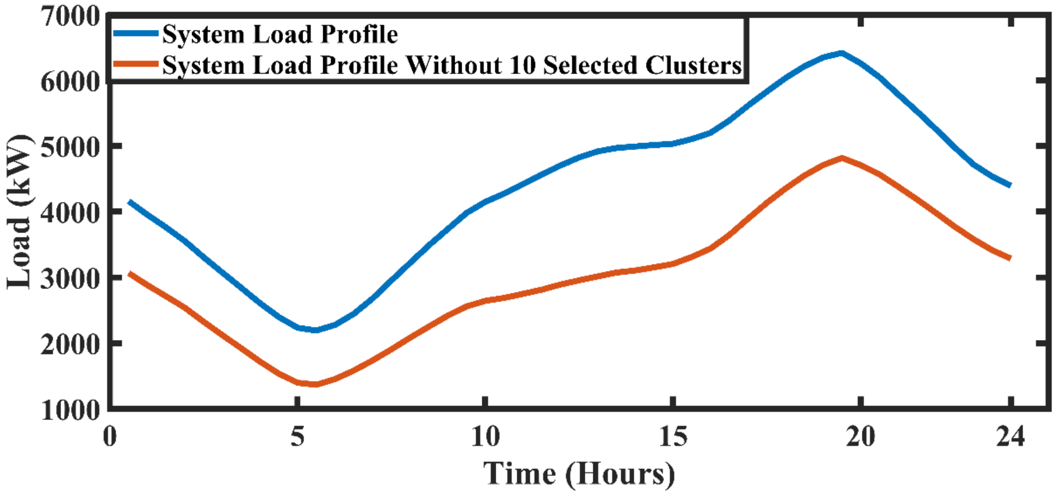


Figure 6.4 Plot of real system load profile and without 10 selected clusters

6.3.2 DSM Application

Upon selection of the clusters, load shifting is applied to shift the load from peak hours to off-peak hours. This reduces the peak of the load and increases savings for the consumers as well as utility. The Time-of-Use tariff, which is taken from the office of gas and electricity markets (Ofgem) in the United Kingdom, is used to reduce the peak of load by load shifting [137] . Raw forecasted profiles require the processing of the constraints to make them convex, whereas linearity of alternative forecasted profiles allows linear processing of the raw profiles. To evaluate the benefits of proposed DSM approach linear programming, which is an efficient linear optimization technique is used to optimize the DSM solution. Two different approaches are compared in this research to analyse better DSM approach for smart meter data. The two approaches are presented in Figure 6.5.

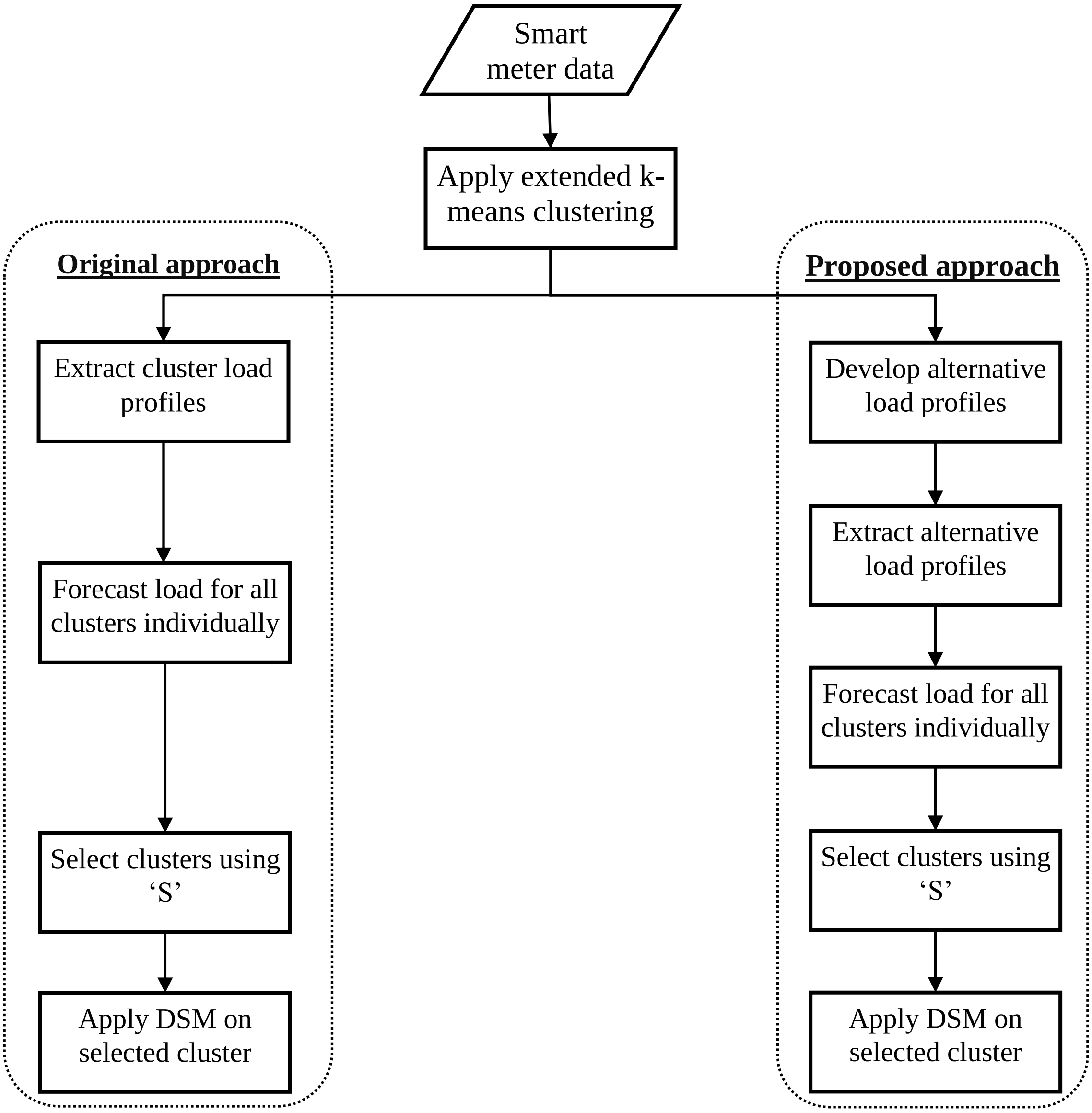

Smart meter data
Apply extended k-means clustering
Original approach
Extract cluster load profiles
Forecast load for all clusters individually
Select clusters using 'S'
Apply DSM on selected cluster
Proposed approach
Develop alternative load profiles
Extract alternative load profiles
Forecast load for all clusters individually
Select clusters using 'S'
Apply DSM on selected cluster

Figure 6.5 Original and alternate approach for DSM application

As the index '*s*' has proved to be effective, for a fair comparison of both approaches, both approaches used '*s*' for cluster selection and results suggested no difference in cluster selection. This also shows that the alternative profiles are refined profiles with high accuracy of the representation of raw profiles. For each approach, two DSM cases are simulated with each case considering different scenarios of DSM for raw and alternative profiles.

Table 6.2 shows the different scenarios considered for DSM simulation with both cases (cases are based on different optimization constraints). At first the optimization constraints are selected i.e. Case I or Case II. After selection of the Case, one selected cluster out of the number of clusters to be optimized is retrieved and the required level of PV generation is incorporated in the cluster profile. The profile after incorporation of PV profile is optimized for cost minimization with demand flexibility ranging from 10% to 90% in steps of 5%. In the next step, the PV penetration is increased to 10% and same process is repeated. After 30% PV generation, demand fexibility is increased to 10% and all combinations are repeated with a step increase of 5% in demand flexibility till demand flexibility reaches to 90%. In total 21,760 simulations are carried out for each case (for both approaches) with a combination of seventeen demand flexibility scenarios, four PV integration and application of DSM on up to 20 individual clusters.

Table 6.2 Cases for DSM

Case I (Demand flexibility 10%-90 %) **Raw load profile & Alternative load profile DSM**					
DSM Constraints	**PV**	**Load (Clusters)**			
Upper bound (6.8), lower bound (6.6) and equality constraint (6.5)	0%	5	10	15	20
	10%				
	20%				
	30%				
Case II (Demand flexibility 10%-90 %) **Raw load profile & Alternative load profile DSM**					
DSM Constraints	**PV**	**Load (Clusters)**			
Upper bound (5% peak shaving) (6.10), lower (6.6) and equality constraint (6.5)	0%	5	10	15	20
	10%				
	20%				
	30%				

The optimization constraints given in Case I consider the upper bound to be the original peak of the cluster and in Case II, the peak is clipped to 95% of the total peak load by shaving

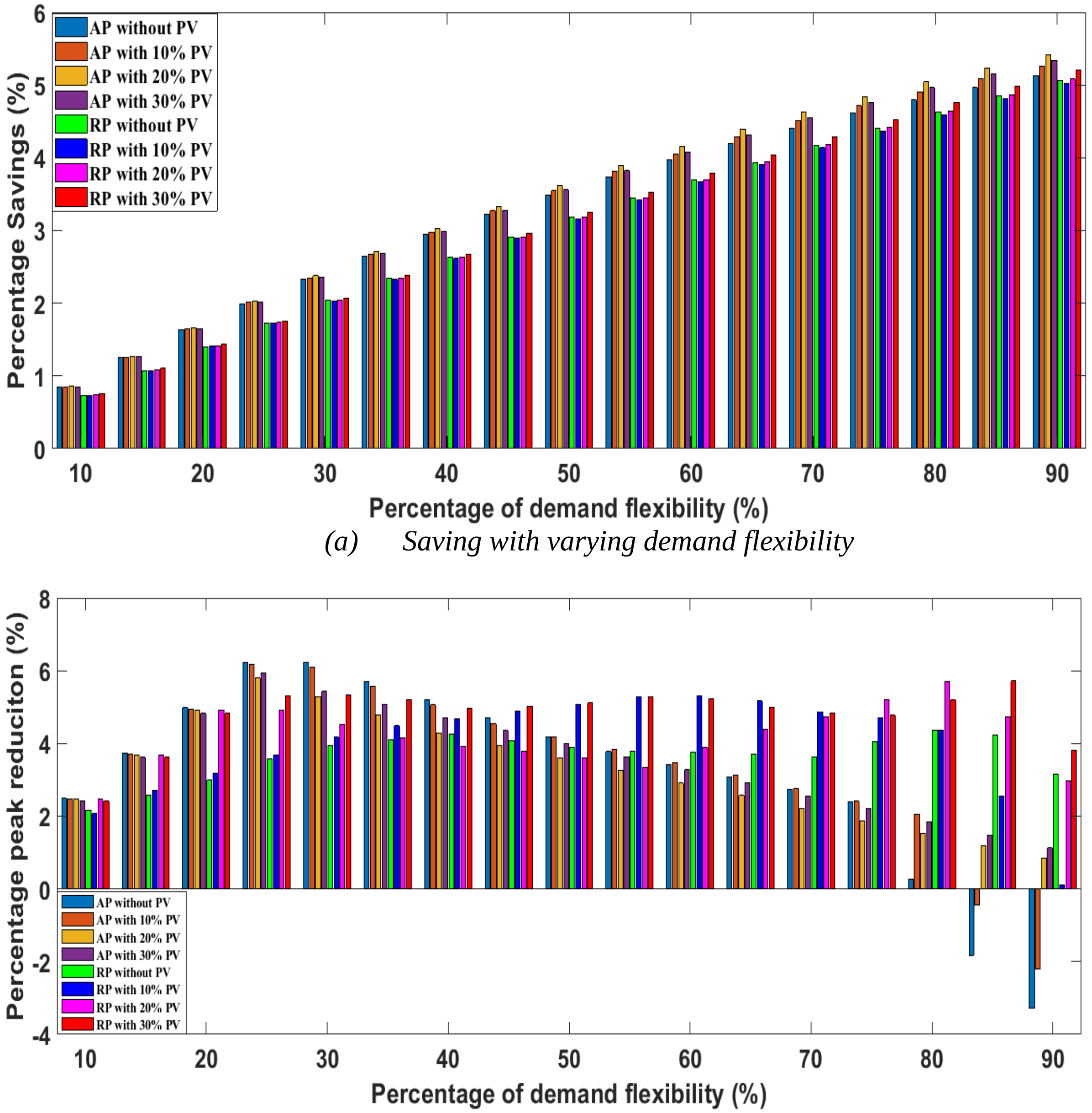
5% of the peak load. The DSM application considers the selected number of clusters individually for application. Each cluster is applied with DSM individually, however the results are consolidated in groups of clusters to see the combined impact on the savings and peak reduction at the system level. After application of DSM on all selected clusters, the final profiles of all clusters, including the profiles of clusters which are not applied with DSM, are aggregated to generate a post DSM system profile. The new system profile is compared with the old system profile to compute savings for consumers and a reduction in load peak. DSM simulation for each scenario is repeated for different levels of demand flexibility. Varying demand flexibility quantifies sensitivity of savings and peak reduction to the responsiveness of the consumers in the selected clusters. Moreover, the impact of PV penetration is also considered by using the forecasted profiles having different levels of PV penetration (Chapter 5, Section 5.3) Comparison of savings and reduction in peak for different levels of demand flexibility is carried out to ascertain the performance of both DSM approaches using linear optimization. Results of both cases are discussed below.

i. CASE I

In the first case, the forecasted load is applied with load shifting using cost minimization objectives given in (6.4) and (6.11). Different scenarios considered for DSM in CaseI include load forecast without PV and with 10%, 20% and 30% PV integrated load forecast. The load shifting is applied to both raw and alternative forecasted profiles to ascertain the impacts of the profiling approaches on the optimization of DSM. Figure 6.6 shows the results of 10 clusters participating in the DSM for cost saving maximization. The savings and peak reduction shown in the Figure 6.6 reflect the system level savings and peak reduction instead of cluster level.

It can be clearly seen from Figure 6.6 (a) that in all scenarios, the cost saving using the forecasted alternative profiles is higher as compared to the raw profile forecasts for all scenarios of PV integration. It is pertinent to mention that savings of up to 2.31%, 4.06% and 5.98% are already achieved by the integration of 10%, 20% and 30% PV respectively for alternative profiles. These savings are not considered in this comparison. The savings for PV integrated profiles in Figure 6.6 are exclusively due to the DSM application so that a fair comparison of peak reduction and savings can be carried out for all scenarios.

The cost savings come with the load shifting from peak to off-peak hours. As the clusters are individually optimized and they are oblivious to other cluster's load, potentially a second peak or rebound effect can occur. In case I, it can be seen from Figure 6.6 (b) that initially peak reduction is achieved, however, with increase in demand flexibility the peak reduction percentage reduces and beyond 25% demand flexibility, the peak reduction starts to decline for alternative profiles. All scenarios of alternative profiles achieve the maximum reduction in peak at 25% demand flexibility, whereas for raw profiles, some profiles even require demand flexibility of up to 80% to achieve the maximum reduction in peak. The maximum peak reduction achieved by raw profiles at higher levels of demand flexibility i.e. 80%, is lower than that of alternative profiles at 25% demand flexibility. Thus, alternative profiles present better optimization solution with higher cost saving at all levels of demand flexibility and higher peak reduction at a lower level of demand flexibility.



(a) Saving with varying demand flexibility

(b) Peak reduction with varying demand flexibility

Figure 6.6 Savings and peak reduction without peak clipping

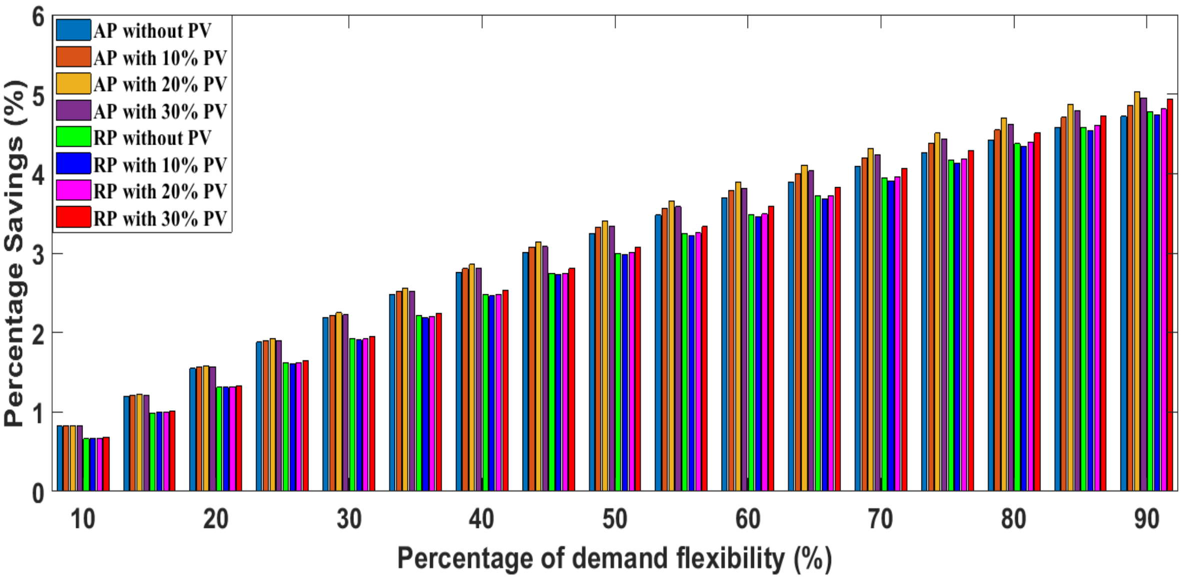
(AP: alternative profile, RP: raw profile)

The key driver for higher saving and peak reduction using alternative profiles is linearity of the profiles. Due to the linearity of the profiles, the convexity of constraints increases which sets the linear boundaries and provides freedom to find global optima. From the above, it can be deduced that to maximize the benefits for system operator i.e. reduction in peak, the demand flexibility of 25% is sufficient using alternative profiles. Whereas, due to non-linearity of raw profiles, the maximum reduction in peak is achieved at higher levels of demand flexibility and peak reduction is less than that presented by alternative profiles. The benefits for consumers are presented in terms of cost savings and higher savings are achieved using alternative profiles as compared to raw profiles.

ii. CASE II

An additional constraint of 5% peak shaving is considered for Case II. The additional constraint is implemented by limiting the maximum magnitude of the load to 95% of the maximum of the peak load of the cluster. Different scenarios described in Case I i.e. varying levels of demand flexibility (10%-90%), varying levels of PV penetration (0%-30%) and a different number of clusters (5-20), are simulated with this additional constraint.

Figure 6.7shows the results of cost saving in all scenarios with 10 clusters only. The savings show similar behaviour as Case I but lesser than that of Case I. The reduction in savings is due to the additional constraint that limits the savings. Peak reduction achieved in this case shows better performance by alternative profiles. Maximum peak reduction by alternative profiles is attained at 30% demand flexibility whereas the peak reduction for raw profiles remains relatively low. The maximum peak reduction for raw profiles is attained between 65-70% demand flexibility. An interesting factor observed in this Case is the pronounced rebound effect for both raw and alternative profiles. With the reduced margin for peak to grow during the peak hours, a new peak emerges in off-peak hours. Restricting the peak of individual clusters to 95% of the original peak does not stop the rebound effect. The rebound occurs due to coincidental peaks of the clusters applied with DSM. This shows that selection of the appropriate level of demand flexibility is essential to avoid rebound of peak. This appropriate level of demand flexibility can also be interpreted as coordinated DSM approach where different consumers cooperate with grid but by communicating with each other to minimize the peak.



Saving with varying demand flexibility

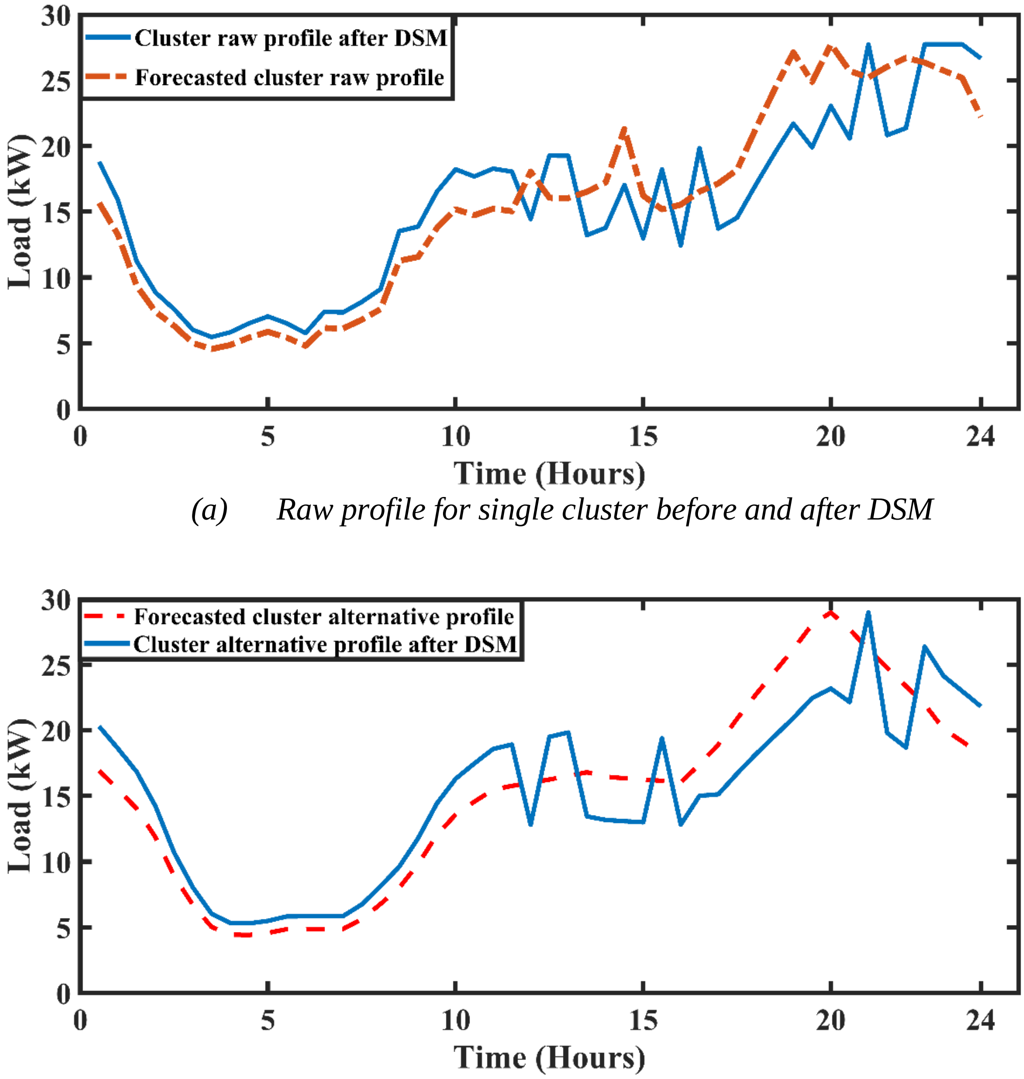
Peak reduction with varying demand flexibility

Figure 6.7 Savings and peak reduction with peak clipping (AP: alternative profile, RP: raw profile)

From Case II, it becomes clear again that linearity of the profiles significantly influences the optimization results. The complexity of raw profiles does not provide appealing peak reduction solution when compared with the solution provided by raw profile. This is again due to non-linearity of the profiles that is further exacerbated by the additional constraint of peak shaving. The additional constraint limits the boundaries of the solution thus reducing the feasible region, which is highly non-linear in case of raw profiles. This non-linearity and reduced freedom lead the linear programming solution for raw profiles to suboptimal solutions. This is evident from the non-linear behaviour of peak reduction with demand flexibility in both cases for raw profiles. On the other hand peak reduction in alternative profiles, shows a comparatively linear and uniform increasing and decreasing trend. Extended studies with higher levels of load using 15 and 20 clusters also showed that the alternative profiles provide better DSM solutions with high cost savings and high peak reduction at lower demand flexibility level. Impacts of the optimization of the cluster and system load profiles are discussed in the proceeding section.

6.3.3 Impacts of DSM on Load Profiles

The results for both cases and the extended studies indicate that being refined and linear, alternative profiles provide better cost saving as well as peak reduction. The financial benefits from the linear nature of alternative profiles come in the shape of higher forecast accuracy and convex data functions. Therefore, an alternate approach for DSM i.e. using the alternative profiles, provides benefits for both electricity consumer and network operator by higher savings and peak reduction. However, additional benefit of DSM at cluster level and system level for network operator can be observed from Figure 6.8 and Figure 6.9 respectively. From Figure 6.8 it can be seen that the non-linearity of the raw profile of a single cluster results in DSM solution with high variability. This variability has an impact on the final system profile, which turns out to have high variations in the system profiles requiring frequent ramping of the generation. The high variability of the post DSM cluster profile mainly due to the inherent variability of the profiles and intermittency of the PV generation. The alternative profiles require lower number of variations in system profile to minimize the objective function. Thus, a comparatively uniform solution with less variability is achieved using alternative profiles.

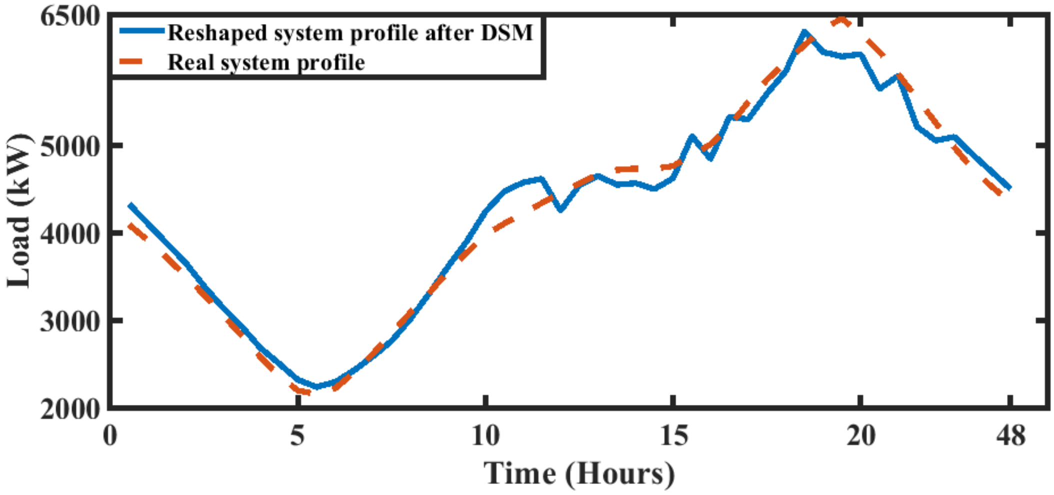


(a) Raw profile for single cluster before and after DSM

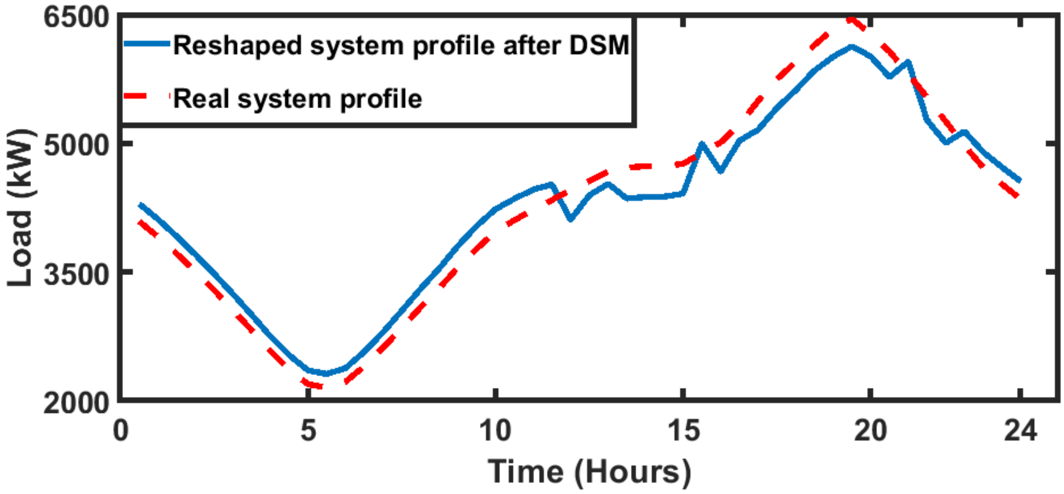
(b) Alternative profile for single cluster before and after DSM

Figure 6.8 Raw vs alternative profile after DSM (Case II, 15% demand flexibility, 10% PV penetration)

Figure 6.9 shows the pre and post DSM system profile for both original and alternative approaches. Visualizing system profiles provides the impact of non-linearity of individual clusters on the entire system. The system profile before DSM tends to be smooth as the load aggregation process cancels the noise in the load data. However, DSM application introduces variations in the load, which will reduce the smoothness of the system profile.



(a) Raw profile before and after DSM



(b) Alternative profile before and after DSM

Figure 6.9 Reshaped system profiles using raw and alternative profiles (Case II, 10 clusters, 15% demand flexibility, 10% PV)

As can be seen from Figure 6.9, the raw system profile lacks smoothness and uniformity in re-shaped profiles. The solution for DSM should tend to reduce the peak load but, in some cases, where load growth is required to fill the valley, load level for raw profiles even falls below the original load level. This happens due to the non-linearity of the profiles, which reduces the uniformity, and smoothness of the re-shaped system profile. The alternative profiles, on the other hand, are more convex in nature, their constraints tend to be linear, and thus the solution achieved using alternative profiles is uniform and smoother as compared to raw profiles. Therefore, alternative profiles produce a better solution without using computationally intensive and complex non-linear techniques to get an improved solution.

6.4 Summary

A novel holistic approach for DSM application at the power distribution systems level is proposed. Cost savings for all scenarios of case studies demonstrate the higher savings emanating from the proposed approach at an effective DSM planning in a power distribution system. Overall results suggest that the inherent non-linearity of the raw profiles is likely to produce suboptimal solutions for the DSM application. The increased convexity of raw profiles by modelling them as alternative profiles can potentially achieve the optimal solution. Higher cost savings for electricity consumers, higher reductions in load peaks at lower demand flexibility levels, reduced processing time and relative uniformity and smoothness of the re-shaped system profiles validate that the proposed approach provides extended opportunities for DSM planning in a power distribution system.

CHAPTER 7

7 Conclusion and Future Works

Smart meters data of individual consumers can be used to increase visibility in the low voltage networks for better understanding of the variations in load and generation profiles. However, a large number of smart meters present challenges in terms of data handling and size. Load patterns at lower aggregation level tend to have noise, high variability, volatility, non-linearity which makes forecasting such loads a challenge. Moreover, the scalability of a large number of consumers can have undesired impacts of DSM application. Thus, this research handled the issues of smart meter data size by reducing it with an innovative clustering algorithm that can benefit load forecasting and DSM with its application-oriented validity indices. A new mathematical framework is proposed to refine these profiles to increase their linearity for reduction of noise, volatility and variability. A new cluster selection index is proposed to mitigate the scalability issue for DSM. Applications in load forecasting and DSM showed significant improvement in load forecast accuracy and increased monetary and technical benefits for DSM. The conclusions drawn from the research presented in this thesis by outlining the key findings based on the proposed model are given below.

7.1 Conclusion

The traditional load profiling methods used in the power industry are outdated and incapable of dealing with big data of smart meters and the operational requirements of future smart grids [6]. For example, load profiles currently used in the UK for the power industry were developed in the 1990s [6, 94]. Typically, load profiles are developed at high aggregation level of load. However, these profiles are unable to reflect the load condition at the distribution networks and lack the insight into granular energy consumption behaviour for individual consumers and groups of consumers [6]. Particularly, the addition of renewable energy system (RES) at the consumer level has changed the shape of load and current methods of load profiling are unable to mirror the impact of RES generation in the load profiles [23]. Emergence of smart grids has enabled the utility to extract detailed information about the energy consumption behaviour of consumers with smart meters. The granular data from the smart meters can be used to develop load profiles of individual consumers or group of consumers. Future smart grid applications like load forecasting at lower aggregation level and demand side management (DSM) in a distribution network, require new load profiling methods that can provide detailed information about the load in distribution network as well as improve the load forecast and benefit DSM to provide the system operator with the flexibility to enhance system reliability and security. Moreover, with increased visibility in the network and need for localized studies for the implementation of distributed control, more system studies at distribution network level are required. However, the size, variability and volatility of smart meter data can add to the complexity and computational burden for such studies that often require linear processing of non-linear data. These features of smart meter data also pose great challenges for applications like load forecasting using smart meter data and need to be dealt with approaches tailored to the need of the applications.

First and the most important problem in the development of load profiles using the smart meter data is the amount of smart meter data. Due to enormity and volatility of data, it is challenging to directly use the big data of smart meters to support power system applications like load forecast, DSM, tariff design etc. An innovative data-driven clustering algorithm for clustering profiles of consumers using non-conventional application-based cluster validity criteria are proposed. The algorithm automatically selects the appropriate level of aggregation based on the number of consumers and produces clusters with high intra-cluster pattern similarity. High intra-cluster pattern similarity can benefit in many power system applications including load forecasting and DSM. Applicability, stability and robustness of the algorithm is validated by applying it on real world smart meter data of more than 5,000 houses. The extended k-means clustering algorithm not only reduced the big data of smart meter to a manageable size, but it explored the patterns embedded within deeper levels of data to select the optimum number of clusters based on the intra-cluster pattern similarity index that can potentially benefit applications, such as load forecasting and DSM. Therefore, the proposed innovative clustering algorithm can provide significant benefits in applications where data reduction and high intra-cluster pattern similarity is desired such as load profiling, load forecasting, DSM etc.

The clusters resulting from the extended k-means clustering process were used to extract typical load profiles for the clusters. The typical profiles of the smart meter consumers were highly non-linear, variable and volatile. In order to address the variability and non-linearity of the profiles, this thesis proposed a novel load profiling approach. A new mathematical model for a new load profiling approach is proposed. The approach linearizes and optimizes the profiles to reduce non-linearity, noise, variability and volatility of profiles. The approach generates refined load profiles that have lower noise, variability and volatility with increased linearity. The approach simplifies the complexity of data while maintaining the precision of representation close to the original data as compared to conventional linear approximation techniques. Alternative profiles are significantly accurate representation of the raw load profiles. The proposed load profiling approach is extended to reduce volatility and variability of intermittent RES generation profiles. The alternate RES generation profiles show higher accuracy as compared to the load profiles due to lesser variability. Application of alternative load and generation profiles in power system reliability assessment shows high level of accuracy with reduced computation time as reported in the published paper [14]. Linearizing the non-linear profiles which were high-order non-linear differential equations into sets of linear equations by eliminating the noise terms and leaving only imminent variations in load, not only reduced the complexity of data functions but also reduced the variability of the load profiles while maintaining the granularity of data. It enhanced the computational efficiency to process the data for the applications in power system that required linear processing of the non-linear data. The reduced complexity of data functions and reduced variability of the profiles generated using proposed novel load profiling approach can benefit power system studies that require linear processing of non-linear data, are sensitive to variability in data and energy optimization applications by increasing the convexity of optimization function.

Load forecasting using smart meter data faces challenges of the high variability of load profiles. At the distribution network level, aggregation of the load does significantly cancel the noise in the profiles and thus, forecasting load becomes challenging with highly variable and irregular load patterns. The pattern similarity index used in the proposed extended k-means clustering algorithm and reduced variability of the proposed alternative profiles make them ideal candidates for improvement in load forecasting. Therefore, the load forecast using the alternative profiles was simulated to observe the impact of linearization of alternative profiles. The load forecasting using alternative profiles presented two-fold benefits, firstly it significantly enhanced the forecast accuracy and secondly it reduced the training time for neural networks. Average improvements of 1.97%, 1.98%, 2% and 2.2% for 0%, 10%, 20% and 30% PV penetration respectively are achieved. Maximum of 6.4% improvement in forecast accuracy is observed which is due to small number of consumers in the cluster. It has been observed that forecast accuracy using the raw profiles are more sensitive to the number of consumers as compared to the alternate profiles.

Selection of appropriate consumers/clusters for DSM is highly reliant on the accuracy of the forecast. An error in the forecast should be considered while planning a DSM activity. A novel cluster selection index for DSM is proposed. The index benefits in selection of appropriate clusters thus reduces the number of consumers required to participate in the DSM program. Moreover, with detailed information about energy consumption behaviour of individual clusters and consumers, incentives for DSM can be tailored to needs of specific consumers. Better forecast accuracy of alternative profiles makes the forecast generated using alternative profiles suitable for DSM application. Use of alternative forecast showed improvements in cost savings and higher peak reduction. Benefits from using the proposed alternative forecast include increased certainty in DSM application with reduced forecast error and improved cluster selection approach. Moreover, due to the linear nature of alternative profiles improved convexity of optimization problem provided more freedom to find global optima which resulted in better optimization solutions i.e. higher cost savings and peak reduction.

The research presented in this thesis has shown that using the mean absolute percentage error (MAPE) as one of the criteria for cluster validity leads to the extraction of clusters having high intra-cluster pattern similarity. Moreover, linearization and optimization of the load profiles successfully the handle the challenges of the complexity of the profiles due to noise, high variability and volatility, and presents significantly accurate results when compared with the non-linear profiles. High intra-cluster pattern similarity of the clusters and refined alternative profiles benefit the power system applications, including but not limited to power system reliability assessment, load forecast and DSM, with significantly high accuracy, improved solutions and reduced processing time. Linear nature of profiles eliminates the need for conversion of non-linear problems into linear and improves the convexity of optimization function to reach global optima.

7.2 Future Works

The objectives of the research proposed in this thesis have been achieved and some of the possible avenues of research are given below.

- Apart from the time series load profiling, spectral load profiling methods can be developed to process the smart meter data on its spectral features. Evaluating load profiling in the frequency domain by characterizing the load profiles by periodic spectral component and representing the big data of smart meters using a small number of spectral coefficients can be promising. Moreover, research into load profiling using hybrid approach of time domain and frequency domain can lead to better analytics.
- This research implemented the DSM on forecasted load, however, implementation of DSM necessarily means a change in future load profile. Consequently, the forecasted profiles need re-forecasting. Consideration of demand response as an input to forecast can potentially incorporate the expected changes in system profile. Therefore, in the presence of DSM, more adaptive load forecasting approaches will be required. Research on this issue has been limited due to the uncertainty in demand response. Consumer identification process presented in this research can potentially reduce the uncertainty in demand response. Tailored DSM strategies for the consumer identified using cluster

selection index can reduce the uncertainty in demand response. Thus, with reduced uncertainty in demand response, the adaptive forecast approach can be developed with demand response as an active input.

References

[1] U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability, "Advanced metering infrastructure and customer systems," 2016.Available: https://www.energy.gov/sites/prod/files/2016/12/f34/AMI%20Summary%20Report_09-26-16.pdf. Accessed on:03/01/2019

[2] C. Ramsay, G. Strbac, A. Badelin, and C. Srikandam, "Impact analysis of increasing (intermittent) RES and DG penetration in the electricity system," 2007.Available: https://pdfs.semanticscholar.org/8ddc/9e5557af86b285a0fefb8ca916a71f01aeb9.pdf?_ga=2.237847365.412614599.1572814573-772953225.1571659406. Accessed on:05/02/2019

[3] Department for Business Energy & Industrial Strategy, "Smart Meter Statistics," 2019.Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/804767/2019_Q1_Smart_Meters_Report.pdf. Accessed on:12-12-2018

[4] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Transactions on Smart Grid,* vol. 10, pp. 3125-3148, 2018.

[5] M. Hommelberg, C. Warmer, I. Kamphuis, J. Kok, and G. Schaeffer, "Distributed control concepts using multi-agent technology and automatic markets: An indispensable feature of smart power grids," in *2007 IEEE Power Engineering Society General Meeting*, 2007, pp. 1-7.

[6] R. Li, "Load profiling on time and spectral domain: from big data to smart data," University of Bath, 2015.Avilable: https://purehost.bath.ac.uk/ws/portalfiles/portal/187958239/Thesis_with_copyright_material_removed.pdf. Accessed on:20/03/2019

[7] K.-J. Park and S.-Y. Son, "A Novel Load Image Profile-Based Electricity Load Clustering Methodology," *IEEE Access*, vol. 7, pp. 59048-59058, 2019.

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

[9] B. Yildiz, J. I. Bilbao, J. Dore, and A. Sproul, "Household electricity load forecasting using historical smart meter data with clustering and classification techniques," in *2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia),* 2018, pp. 873-879.

[10] Z. Khan, D. Jayaweera, and H. Gunduz, "Smart meter data taxonomy for demand side management in smart grids," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 International Conference on*, 2016, pp. 1-8.

[11] Z. A. Khan and D. Jayaweera, "Approach for smart meter load profiling in Monte Carlo simulation applications," *IET Generation, Transmission & Distribution*, vol. 11, pp. 1856-1864, 2017.

[12] Z. A. Khan and D. Jayaweera, "Approach for forecasting smart customer demand with significant energy demand variability," in *2018 1st International Conference on Power, Energy and Smart Grid (ICPESG)*, 2018, pp. 1-5.

[13] Z. A. Khan, D. Jayaweera, and M. S. Alvarez-Alvarado, "A novel approach for load profiling in smart power grids using smart meter data," *Electric Power Systems Research*, vol. 165, pp. 191-198, 2018.

[14] H. Gunduz, Z. A. Khan, A. Altamimi, and D. Jayaweera, "An Innovative Methodology for Load and Generation Modelling in a Reliability Assessment with PV and Smart Meter Readings," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, 2018, pp. 1-5.

[15] Z. A. Khan and D. Jayaweera, "Planning and Operational Challenges in a Smart Grid," in *Smart Power Systems and Renewable Energy System Integration*, ed: Springer, 2016, pp. 153-177.

[16] Z. A. Khan and D. Jayaweera, "Efficient Management of Demand in a Power Distribution System with Smart Meter Data," in *2019 IEEE Milan PowerTech*, 2019, pp. 1-6.

[17] D. Bakken, *Smart Grids: Clouds, Communications, Open Source, and Automation*: CRC Press, 2014.

[18] T. Hong, J. Wilson, and J. Xie, "Long term probabilistic load forecasting and normalization with hourly information," *IEEE Transactions on Smart Grid,* vol. 5, pp. 456-462, 2014.

[19] 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,* vol. 5, pp. 402-410, 2014.

[20] J. I. Maletic and A. Marcus, "Data cleansing: A prelude to knowledge discovery," in *Data Mining and Knowledge Discovery Handbook*, ed: Springer, 2009, pp. 19-32.

[21] K. Orr, "Data quality and systems theory," *Communications of the ACM,* vol. 41, pp. 66-71, 1998.

[22] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*: Elsevier, 2011.

[23] F. L. Quilumba-Gudino, "Using advanced metering infrastructure data for smart grid development," 2014.Avilable: https://rc.library.uta.edu/uta-ir/bitstream/handle/10106/24446/QuilumbaGudino_uta_2502D_12542.pdf?sequence=1. Accessed on:20/08/2018

[24] J. Shishido, "Smart Meter Data Quality Insights," EnerNOC Utility Solutions.Available: https://aceee.org/files/proceedings/2012/data/papers/0193-000375.pdf. Accessed on:21/02/2016

[25] J. S. Lopes, "VALIDATION, EDITING AND EXPANSION IN A DEREGULATED ENVIRONMENT " 2000.Available: http://www.lopesenergyconsulting.com/wp-content/uploads/2010/08/Lopes-Valeditaeic2k.pdf. Accessed on:25/06/2019

[26] Elhub, " Standard for Validation, Estimation and Editing (VEE) of AMS metering values," Elhub, Oslo, Norway2014.Available: https://docplayer.net/62949858-Standard-for-validation-estimation-and-editing-vee-of-ams-metering-values.html. Accessed on:04/05/2019

[27] V. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial intelligence review,* vol. 22, pp. 85-126, 2004.

[28] L. Rokach and O. Maimon, "Clustering methods," in *Data mining and knowledge discovery handbook,* ed: Springer, 2005, pp. 321-352.

[29] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: algorithms and applications," *The VLDB Journal—The International Journal on Very Large Data Bases,* vol. 8, pp. 237-253, 2000.

[30] V. Barnett and T. Lewis, "Outliers in Statistical Data (Probability & Mathematical Statistics)," ed: Wiley Chichester, 1994.

[31] C. Chen and D. J. Cook, "Energy Outlier Detection in Smart Environments," *Artificial Intelligence and Smarter Living,* vol. 11, 2011.

[32] B. Rossi, S. Chren, B. Buhnova, and T. Pitner, "Anomaly detection in smart grid data: An experience report," in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, 2016, pp. 002313-002318.

[33] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Transactions on Smart Grid,* 2019.

[34] F. L. Quilumba, W.-J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, "Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behavior Similarities," *IEEE Trans. Smart Grid*, vol. 6, pp. 911-918, 2015.

[35] H. L. Willis and J. E. Northcote-Green, "Spatial electric load forecasting: a tutorial review," *Proceedings of the IEEE,* vol. 71, pp. 232-253, 1983.

[36] H. Willis, A. Schauer, J. Northcote-Green, and T. Vismor, "Forecasting distribution system loads using curve shape clustering," *IEEE Transactions on power apparatus and systems*, pp. 893-901, 1983.

[37] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *IEEE Transactions on Power Systems,* vol. 16, pp. 498-505, 2001.

[38] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on power systems,* vol. 16, pp. 44-55, 2001.

[39] S. Valero, M. Ortiz, C. Senabre, C. Alvarez, F. Franco, and A. Gabald, "Methods for customer and demand response policies selection in new electricity markets," *IET generation, transmission & distribution,* vol. 1, pp. 104-110, 2007.

[40] M. R. Anderberg, *Cluster analysis for applications*: Academic press, 1973.

[41] B. Everitt, S. Landau, M. Leese, and D. Stahl, "Cluster Analysis. –John Wiley & Sons," *Ltd., New York*, p. 330, 2011.

[42] C. C. Aggarwal and C. K. Reddy, *Data clustering: algorithms and applications*: CRC press, 2013.

[43] R. Xu and D. Wunsch, *Clustering* vol. 10: John Wiley & Sons, 2008.

[44] B. Mirkin, *Clustering for data mining: a data recovery approach*: Chapman and Hall/CRC, 2005.

[45] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern recognition letters,* vol. 31, pp. 651-666, 2010.

[46] B. Pitt and D. Kirschen, "Application of data mining techniques to load profiling," in *Power Industry Computer Applications, 1999. PICA'99. Proceedings of the 21st 1999 IEEE International Conference*, 1999, pp. 131-136.

[47] L. Rokach, "A survey of clustering algorithms," in *Data mining and knowledge discovery handbook*, ed: Springer, 2009, pp. 269-298.

[48] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, pp. 264-323, 1999.

[49] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications* vol. 20: Siam, 2007.

[50] D. Pelleg and A. W. Moore, "X-means: Extending k-means with efficient estimation of the number of clusters," in *Icml*, 2000, pp. 727-734.

[51] J. Podani, *Introduction to the exploration of multivariate biological data*: Backhuys Publishers, 2000.

[52] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*: SIAM, 2007.

[53] F. Kovács, C. Legány, and A. Babos, "Cluster validity measurement techniques," in *6th International symposium of hungarian researchers on computational intelligence*, 2005.

[54] M. V. Maria Halkidi, "Quality Assessment Approaches in Data Mining," in *Data mining and knowledge discovery handbook*, ed: Springer, 2005, p. 623.

[55] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," *Journal of intelligent information systems,* vol. 17, pp. 107-145, 2001.

[56] G. W. Milligan and M. C. Cooper, "An examination of procedures for determining the number of clusters in a data set," *Psychometrika,* vol. 50, pp. 159-179, 1985.

[57] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods,* vol. 3, pp. 1-27, 1974.

[58] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *Journal of cybernetics,* vol. 4, pp. 95-104, 1974.

[59] T. Hong, "Short term electric load forecasting," North Carolina State University, 2010.Avilable: https://repository.lib.ncsu.edu/handle/1840.16/6457. Accessed on:05/10/2017

[60] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting,* vol. 32, pp. 914-938, 2016.

[61] L. Hernandez, C. Baladron, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, J. Lloret, *et al.*, "A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings," *Communications Surveys & Tutorials, IEEE,* vol. 16, pp. 1460-1495, 2014.

[62] Y. Loewenstern, L. Katzir, and D. Shmilovitz, "The effect of system characteristics on very-short-term load forecasting," in *2015 International School on Nonsinusoidal Currents and Compensation (ISNCC)*, 2015, pp. 1-6.

[63] I. S. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *Power Systems, IEEE Transactions on,* vol. 4, pp. 1484-1491, 1989.

[64] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," *Renewable and Sustainable Energy Reviews,* vol. 54, pp. 1311-1322, 2// 2016.

[65] W. Charytoniuk and C. Mo-Shing, "Very short-term load forecasting using artificial neural networks," *Power Systems, IEEE Transactions on,* vol. 15, pp. 263-268, 2000.

[66] L. Hernandez, C. Baladron, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, J. Lloret, *et al.*, "A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings," *IEEE Communications Surveys & Tutorials,* vol. 16, pp. 1460-1495, 2014.

[67] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, "Deep belief network based electricity load forecasting: An analysis of Macedonian case," *Energy,* vol. 115, pp. 1688-1700, 2016.

[68] L. C. M. de Andrade and I. Nunes da Silva, "Using intelligent system approach for very short-term load forecasting purposes," in *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, 2010, pp. 694-699.

[69] H. Daneshi and A. Daneshi, "Real time load forecast in power system," in *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, 2008, pp. 689-695.

[70] R. E. Edwards, J. New, and L. E. Parker, "Predicting future hourly residential electrical consumption: A machine learning case study," *Energy and Buildings,* vol. 49, pp. 591-603, 2012.

[71] J. Kwac, J. Flora, and R. Rajagopal, "Household energy consumption segmentation using hourly data," *IEEE Transactions on Smart Grid,* vol. 5, pp. 420-430, 2014.

[72] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, "Short-term residential load forecasting based on resident behaviour learning," *IEEE Transactions on Power Systems,* vol. 33, pp. 1087-1088, 2018.

[73] T. K. Wijaya, M. Vasirani, S. Humeau, and K. Aberer, "Cluster-based aggregate forecasting for residential electricity demand using smart meter data," in *Big Data (Big Data), 2015 IEEE International Conference on,* 2015, pp. 879-887.

[74] V. Kecman, *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*: MIT press, 2001.

[75] F. L. Quilumba, W.-J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities," *IEEE Transactions on Smart Grid,* vol. 6, pp. 911-918, 2015.

[76] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural networks,* vol. 12, pp. 145-151, 1999.

[77] R. Setiono and L. C. K. Hui, "Use of a quasi-Newton method in a feedforward neural network construction algorithm," *IEEE Transactions on Neural Networks,* vol. 6, pp. 273-277, 1995.

[78] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis,* ed: Springer, 1978, pp. 105-116.

[79] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Transactions on Power Systems,* vol. 27, pp. 134-141, 2012.

[80] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE transactions on industrial informatics,* vol. 7, pp. 381-388, 2011.

[81] Z. Zhu, "Mathematical optimization techniques for demand management in smart grids," Loughborough University, 2014.Avilable: https://repository.lboro.ac.uk/articles/Mathematical_optimization_techniques_for_demand_management_in_smart_grids/9540347. Accessed on:09/09/2018

[82] A. Faruqui, D. Harris, and R. Hledik, "Unlocking the€ 53 billion savings from smart meters in the EU: How increasing the adoption of dynamic tariffs could make or break the EU's smart grid investment," *Energy Policy,* vol. 38, pp. 6222-6231, 2010.

[83] A. Mahmood, N. Javaid, and S. Razzaq, "A review of wireless communications for smart grid," *Renewable and sustainable energy reviews,* vol. 41, pp. 248-260, 2015.

[84] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, *et al.*, "A survey on smart grid potential applications and communication requirements," *IEEE Transactions on industrial informatics,* vol. 9, pp. 28-42, 2013.

[85] M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Communication network requirements for major smart grid applications in HAN, NAN and WAN," *Computer Networks,* vol. 67, pp. 74-88, 2014.

[86] A. Sinha and M. De, "Load shifting technique for reduction of peak generation capacity requirement in smart grid," in *Power Electronics, Intelligent Control and Energy Systems (ICPEICES), IEEE International Conference on,* 2016, pp. 1-5.

[87] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE transactions on smart grid*, vol. 3, pp. 1244-1252, 2012.

[88] T. M. Hansen, E. K. P. Chong, S. Suryanarayanan, A. A. Maciejewski, and H. J. Siegel, "A Partially Observable Markov Decision Process Approach to Residential Home Energy Management," *IEEE Transactions on Smart Grid*, vol. 9, pp. 1271-1281, 2018.

[89] M. Kojury-Naftchali, A. Fereidunian, and H. Lesani, "Identifying susceptible consumers for demand response and energy efficiency policies by time-series analysis and supplementary approaches," in *Electrical Engineering (ICEE), 2016 24th Iranian Conference on*, 2016, pp. 1130-1135.

[90] H.-A. Cao, C. Beckel, and T. Staake, "Are domestic load profiles stable over time? An attempt to identify target households for demand side management campaigns," in *Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE*, 2013, pp. 4733-4738.

[91] B. A. Smith, J. Wong, and R. Rajagopal, "A simple way to use interval data to segment residential customers for energy efficiency and demand response program targeting," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012, pp. 374-386.

[92] F. Javed, N. Arshad, F. Wallin, I. Vassileva, and E. Dahlquist, "Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting," *Applied Energy*, vol. 96, pp. 150-160, 2012.

[93] ISSDA. (2012). *CER Smart Meter Customer Behaviour Trials Data, accessed via the Irish Social Science Data Archive, CER Electricity*. Available: www.ucd.ie/issda Accessed on:20/11/2015

[94] *Electricity user load profiles by profile class (1997 ed.)*. Available: http://ukerc.rl.ac.uk/DC/cgi-bin/edc_search.pl?GoButton=Detail&WantComp=42&WantResult=&WantText=EDC0000041. Accessed on:05/12/2016

[95] ofgem. (2018). *Electricity Settlement Reform*. Available: https://www.ofgem.gov.uk/electricity/retail-market/market-review-and-reform/smarter-markets-programme/electricity-settlement. Accessed on:11-10-2018

[96] IBM, "Managing big data for smart grids and smart meters," USAMay 2012.Available: http://www-935.ibm.com/services/multimedia/Managing_big_data_for_smart_grids_and_smart_meters.pdf. Accessed on:13/05/2016

[97] C. Down, "Consultation on extending the existing smart meter framework for data access and privacy to Smart-Type Meters and Advanced Meters," ofgem2013.Available: https://www.ofgem.gov.uk/publications-and-updates/consultation-extending-existing-smart-meter-framework-data-access-and-privacy-smart-type-meters-and-advanced-meters. Accessed on:29/07/2017

[98] S. M. Bidoki, N. Mahmoudi-Kohan, and S. Gerami, "Comparison of several clustering methods in the case of electrical load curves classification," in *16th Electrical Power Distribution Conference*, 2011, pp. 1-7.

[99] L. Liu, G. Wang, and D.-h. Zhai, "Application of k-means clustering algorithm in load curve classification," *Power System Protection and Control*, vol. 39, pp. 65-68, 2011.

[100] T. Räsänen and M. Kolehmainen, "Feature-based clustering for electricity use time series data," in *International Conference on Adaptive and Natural Computing Algorithms*, 2009, pp. 401-412.

[101] V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, "An electric energy consumer characterization framework based on data mining techniques," *IEEE Transactions on power systems,* vol. 20, pp. 596-602, 2005.

[102] I. P. Panapakidis and G. C. Christoforidis, "Optimal Selection of Clustering Algorithm via Multi-Criteria Decision Analysis (MCDA) for Load Profiling Applications," *Applied Sciences,* vol. 8, p. 237, 2018.

[103] R. Damayanti, A. Abdullah, W. Purnama, and A. Nandiyanto, "Electrical Load Profile Analysis Using Clustering Techniques," in *IOP Conference Series: Materials Science and Engineering,* 2017, p. 012081.

[104] R. Granell, C. J. Axon, and D. C. Wallom, "Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles," *IEEE Transactions on Power Systems,* vol. 30, pp. 3217-3224, 2015.

[105] Y.-I. Kim, J.-M. Ko, and S.-H. Choi, "Methods for generating TLPs (typical load profiles) for smart grid-based energy programs," in *Computational Intelligence Applications In Smart Grid (CIASG), 2011 IEEE Symposium on,* 2011, pp. 1-6.

[106] P. ELIASSON and N. Rosen, "Efficient K-means clustering and the importanceof seeding," KTH Royal Institute of Technology 2013.Avilable: http://www.diva-portal.org/smash/get/diva2:668713/FULLTEXT01.pdf. Accessed on:17/05/2016

[107] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence,* pp. 224-227, 1979.

[108] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting,* vol. 32, pp. 669-679, 2016.

[109] W. Labeeuw and G. Deconinck, "Residential electrical load model based on mixture model clustering and Markov models," *IEEE Transactions on Industrial Informatics,* vol. 9, pp. 1561-1569, 2013.

[110] H. Kile and K. Uhlen, "Data reduction via clustering and averaging for contingency and reliability analysis," *International Journal of Electrical Power & Energy Systems,* vol. 43, pp. 1435-1442, 2012.

[111] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, "Data mining with big data," *IEEE transactions on knowledge and data engineering,* vol. 26, pp. 97-107, 2014.

[112] Y. Wang, Q. Chen, C. Kang, M. Zhang, K. Wang, and Y. Zhao, "Load profiling and its application to demand response: A review," *Tsinghua Science and Technology,* vol. 20, pp. 117-129, 2015.

[113] X. Liu, L. Golab, W. Golab, I. F. Ilyas, and S. Jin, "Smart meter data analytics: systems, algorithms, and benchmarking," *ACM Transactions on Database Systems (TODS),* vol. 42, p. 2, 2017.

[114] H. Niska, P. Koponen, and A. Mutanen, "Evolving smart meter data driven model for short-term forecasting of electric loads," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on,* 2015, pp. 1-6.

[115] W. Ko, J.-K. Park, M.-K. Kim, and J.-H. Heo, "A Multi-Energy System Expansion Planning Method Using a Linearized Load-Energy Curve: A Case Study in South Korea," *Energies,* vol. 10, p. 1663, 2017.

[116] T. Hong. (2014). *Energy Forecasting*. Available: http://blog.drhongtao.com/2014/09/load-demand-energy-power.html. Accessed on:26/05/2017

[117] A. Shahmohammadi, M. Moradi-Dalvand, H. Ghasemi, and M. Ghazizadeh, "Optimal design of multicarrier energy systems considering reliability constraints," *IEEE Transactions on Power Delivery,* vol. 30, pp. 878-886, 2015.

[118] E. A. M. Ceseña, T. Capuder, and P. Mancarella, "Flexible distributed multienergy generation system expansion planning under uncertainty," *IEEE Transactions on Smart Grid,* vol. 7, pp. 348-357, 2016.

[119] X. Zhang, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Optimal expansion planning of energy hub with multiple energy infrastructures," *IEEE Transactions on Smart Grid,* vol. 6, pp. 2302-2311, 2015.

[120] F. Domínguez-Muñoz, J. M. Cejudo-López, A. Carrillo-Andrés, and M. Gallardo-Salazar, "Selection of typical demand days for CHP optimization," *Energy and buildings,* vol. 43, pp. 3036-3043, 2011.

[121] P. M. Lara-Santillán, M. Mendoza-Villena, L. A. Fernández-Jiménez, and M. Mañana-Canteli, "A comparative study of electric load curve changes in an urban low-voltage substation in Spain during the economic crisis (2008–2013)," *The Scientific World Journal,* vol. 2014, 2014.

[122] L. N. Trefethen and J. A. C. Weideman, "Two results on polynomial interpolation in equally spaced points," *Journal of Approximation Theory,* vol. 65, pp. 247-260, 1991/06/01/ 1991.

[123] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical chemistry,* vol. 36, pp. 1627-1639, 1964.

[124] C. W. De Silva, *Modeling and control of engineering systems*: Crc Press, 2009.

[125] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, 1995, pp. 39-43.

[126] M. H. Kalos and P. A. Whitlock, *Monte carlo methods*: John Wiley & Sons, 2009.

[127] M. S. Alvarez-Alvarado and D. Jayaweera, "Reliability model for a Static Var Compensator," in *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, 2017, pp. 1-6.

[128] P. Jones, C. Kilsby, C. Harpham, V. Glenis, and A. Burton, "UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator," 2009.Available: http://cedadocs.ceda.ac.uk/1335/1/weather_generator_full_report.pdf. Accessed on:29/06/2016

[129] S. N. Fallah, M. Ganjkhani, S. Shamshirband, and K.-w. Chau, "Computational Intelligence on Short-Term Load Forecasting: A Methodological Overview," *Energies,* vol. 12, p. 393, 2019.

[130] T. Zufferey, A. Ulbig, S. Koch, and G. Hug, "Forecasting of smart meter time series based on neural networks," in *International Workshop on Data Analytics for Renewable Energy Integration*, 2016, pp. 10-21.

[131] K. Gajowniczek and T. Ząbkowski, "Short term electricity forecasting using individual smart meter data," *Procedia Computer Science,* vol. 35, pp. 589-597, 2014.

[132] C. C. Aggarwal, *Neural Networks and Deep Learning: A Textbook*: Springer, 2018.

[133] T. I. M. Service. Available: https://www.met.ie/climate/available-data. Accessed on:10/04/2016

[134] D. S. Kirschen, "Demand-side view of electricity markets," *IEEE Transactions on Power Systems,* vol. 18, pp. 520-527, 2003.

[135] G. Strbac, "Demand side management: Benefits and challenges," *Energy Policy*, vol. 36, pp. 4419-4426, 2008/12/01/ 2008.

[136] M. Kia, S. Mir Mohammad Reza, D. Esmaeil Abedini, and H. Seyed Hamid, "Simultaneous implementation of optimal Demand Response and security constrained Unit Commitment," in *16th Electrical Power Distribution Conference*, 2011, pp. 1-5.

[137] ofgem. (2019). *Ofgem making a positive difference for energy consumers*. Available: https://www.ofgem.gov.uk/. Accessed on:20/01/2018

Appendix A

The Smart Metering Electricity Customer Behaviour Trials (CBTs) took place during 2009 and 2010 with over 5,000 Irish homes and businesses participating [93]. The data set used in this study consists of 6 comma-separated values (CSV) files with 180 million rows of data for more than 5,000 consumers. The data consists of 25,728 time series half-hourly records of energy consumptions for each individual consumer including domestic and small businesses consumers for 1.5 years starting from the first of January 2009.

The final numbers selected for the field trials were as follows [93]:

• 5,800 single phase and 500 three phase meters throughout the country for the customers randomly selected for the CBT had GPRS based communications.

• 1,100 single phase meters for customers in 11 locations in Limerick and Ennis were installed for the powerline carrier trial. Eight of the locations chosen were urban and three were village areas.

• 2,281 meters comprising of 1591 meters installed in Cork City and 690 meters installed in the rural area of County Cork outside Bandon for the wireless mesh trial. Desktop studies were carried out on the remaining two technologies – PLC from Aclara and 868 MHz RF from Elster/Coronis.

The smart meters had following automatic functionalities [93];

- Automatic registration of the smart meter on the system

- Scheduled Daily Load Profile retrieval for 30 minute intervals
- Scheduled daily midnight register readings
- Scheduled events and alarms
- Event log management

Moreover, the on-demand functionalities included [93];

- De-energisation and re-energisation (sandbox)
- On demand profile reading
- On demand register reading
- Re-configuration of parameters on meter
- Firmware upgrade capability
- Power quality monitoring

The smart meter data considered in this study did not consider the additional functionalities and only daily load profiles are considered with the date and time stamp.