INTERNATIONAL HELLENIC UNIVERSITY

SCHOOL OF ENGINEERING
DEPARTMENT OF INFORMATION AND ELECTRONIC ENGINEERING

BSc THESIS
BIG DATA ANALYTICS FOR ELECTRIC VEHICLES IN THE SMART GRID

Of student
Angelos Koutsopoulos
Reg. Number: 123834

Supervisor
Dimitrios Dervos
Professor

Thessaloniki 2020

BSc title: Big Data analytics for electric vehicles in the smart grid
BSc code: 19036
Student's name: Angelos Koutsopoulos
Supervisor's name: Dimitros Dervos
BSc's assumption date: 13/5/2019
BSc completion date: 17/6/2020

I hereby declare that I am the original author of this thesis and that every assistance that I received in the course of the project work reported is acknowledged and (where possible) cited in the "References/Bibliography" section of the present manuscript. All sources of information and/or data, ideas, pictures, graphs, figures and text excerpts used (be it "as is" quotations, or re-phrased), they are properly referenced in footnotes or directly in text. I also declare that I myself have conducted the project work herewith reported and compiled the present manuscript, along the lines of fulfilling the BSc thesis project requirement of my study with the Department of Information and Electronic Systems Engineering.

Preface

This BSc thesis titled “Big Data analytics for electric vehicles in the smart grid” was prepared in the context of fulfilling the required conditions for obtaining my degree from the department of information and electronic engineering of the International Hellenic University based in Thessaloniki. It was undertaken on April 2019 and it was completed on June 2020, within the prescribed limits.

The research was carried out under the supervision of Prof. Dimitrios Dervos and in collaboration with Dr. Chrysovalantou Ziogou, senior research associate at the Centre For Research and Technology Hellas (CERTH).

Abstract

Global warming and the increasing fossil fuel prices favor the mass production of electric vehicles that are environmentally friendly due to their zero or minimum greenhouse gas emissions. However, their increase in number and uncoordinated charging policies can cause serious problems in the smart grid which is the evolved version of the older electrical grid. Therefore, proper electrical power management during peak hours is a critical issue. To address this issue, demand side management, which is an important aspect of smart grids, offers several strategic methods to improve the load on the network and provide benefits both to the consumers and to energy providers. In the present thesis project, a demand side management strategy is proposed as a method to tackle and control the issue in question, by utilizing analysis and descriptive statistics to manipulate data and visualize results. The aim is to migrate energy demands in the smart grid from peak hours to off-peak hours. The results show that the proposed strategy can lower energy demands relating electric vehicle charging during peak hours to satisfactory levels.

Περίληψη

Η υπερθέρμανση του πλανήτη και οι αυξανόμενες τιμές των ορυκτών καυσίμων οδηγούν στη μαζική παραγωγή ηλεκτρικών οχημάτων τα οποία είναι φιλικά προς το περιβάλλον λόγω των μηδενικών ή ελάχιστων εκπομπών βλαβερών αερίων του θερμοκηπίου. Ωστόσο, η αύξηση του αριθμού τους και η μη συντονισμένη φόρτισή τους μπορούν να προκαλέσουν σοβαρά προβλήματα στο ευφυές δίκτυο ενέργειας (smart grid), το οποίο αποτελεί την εξελιγμένη εκδοχή του παλαιότερου ηλεκτρικού δικτύου. Επομένως, η σωστή διαχείριση της ηλεκτρικής ενέργειας κατά τις ώρες αιχμής είναι ένα κρίσιμο ζήτημα. Για την αντιμετώπιση αυτού του ζητήματος, το demand side management, το οποίο είναι μία σημαντική πτυχή των smart grids, προσφέρει διάφορες στρατηγικές μεθόδους για τη βελτίωση του φόρτου στο δίκτυο και την παροχή οφελών τόσο στους καταναλωτές όσο και στους παρόχους ενέργειας. Στην παρούσα πτυχιακή εργασία προτείνεται μια στρατηγική του demand side management ως μέθοδος για να ξεπεραστεί το τρέχον ζήτημα, χρησιμοποιώντας εργαλεία αναλυτικής των δεδομένων (data analytics tools) και περιγραφική στατιστικής (descriptive statistics) για τη διαχείριση των δεδομένων και την για οπτικοποίηση των αποτελεσμάτων. Εφαρμόζοντας την εν λόγω στρατηγική, οι ενεργειακές απαιτήσεις του smart grid μεταφέρονται από τις ώρες αιχμής σε ώρες μη-αιχμής, στη διάρκεια της ημέρας. Τα αποτελέσματα δείχνουν ότι η προτεινόμενη μέθοδος μπορεί να μειώσει την ζήτηση ενέργειας για την φόρτιση των ηλεκτρικών οχημάτων κατά τις ώρες αιχμής σε ικανοποιητικά επίπεδα.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Dimitrios Dervos, and to Dr. Chrysovalantou Ziogou, senior research associate at the Centre For Research and Technology Hellas (CERTH), for their unwavering guidance through each stage of this project and for their continuous support and encouragement. The achieved result could not have been possible without their invaluable assistance and co-ordination.

I would also like to thank my family and friends for their unparalleled support and encouragement throughout the duration of my thesis project study and work.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Greenhouse gas emissions, mainly coming from factory and petrol-powered vehicle exhausts, are the main cause of global warming, posing an immense threat to the environment and therefore to the human race. Moreover, fossil fuel prices in the world market are becoming a wide concern to both economical and industrial levels. As a result, the focus is shifting towards the production of more Plug-in electric vehicles (PEVs) which are thought to reduce these problems due to their zero or minimal gas emissions [1].

PEV is a broad term which refers to vehicles that depend at least partially on a rechargeable battery to run. There are two main types of PEVs: battery electric vehicle (BEV) and plug-in hybrid electric vehicle (PHEV). BEVs run exclusively on electricity stored in high-capacity battery packs and operate through an electric motor. They have zero emissions but 4-6 hours are required for a complete battery charge. PHEVs are equipped with both an electric motor and a gasoline engine. They have low gas emissions and a complete charge ends in about 1-3 hours due to its smaller capacity battery [2] [3]. Table 1.1 provides the differences between the two previously mentioned PEV types.

Table 1.1 "BEV - PHEV comparison"

Battery Electric Vehicle (BEV)	Plug-in Hybrid Electric Vehicle (PHEV)
Electric motor	Electric motor & gasoline engine
High-capacity battery	Low-capacity battery
4-6 hours to charge	1-3 hours to charge
Zero gas emissions	Low gas emissions

The battery is the place where the PEV’s energy is stored. The electric motor uses this energy to drive the vehicle’s wheels. PHEVs run on the electric motor until the battery is depleted and then they switch over to the gasoline engine. On average, a battery lasts for 8 years or 160.000km. The most common battery type used in PEVs is the lithium-ion battery, which is also found in most portable appliances, like cell phones and computers. Moreover, most of this battery type parts are fully recyclable, making them a good choice for the protection of the environment. Figure 1.1 shows the Tesla model S lithium-ion battery chassis [4].

Figure 1.1 "Tesla Model S battery chassis [5]"

A PEV's battery can be charged by plugging the vehicle's cord into an electric power source at a public charging station, or into a power outlet conveniently located at the PEV owner's home. Three charging levels are currently available in the market: Level 1, Level 2 and Level 3 (also called Direct Current or simply DC).

Level 1 is the slowest and cheapest of the three. It uses 120V chargers, consumes 1.92kW and its power is equivalent to that of a toaster. It can fully charge a depleted BEV's battery in approximately 20 hours, adding 5 to 8 km of travel per hour of charging. On the other hand, Level 2 is twice as fast and fully charges a depleted battery in around 7 hours, adding 32 to 40 km of travel per hour of charging. It uses 240V chargers, consumes 6.6kW and its power consumption rate is equivalent to that of a tumble dryer. Lastly, DC is the fastest type of charging, using very high voltage chargers which range from 200V to 500V, consuming up to 40kW. It recharges up to 80 percent of the battery's capacity within 30 or less minutes. However, it is available only at selective public charging stations, and not all PEVs are compatible with it, it is more expensive than Level 1 and Level 2 chargers and it may damage the vehicle's battery [6] [7] [8].

When it comes to charging PEVs at home, Level 1 is the simplest because it does not require the purchase of additional charging equipment. Owners can charge their vehicle by simply plugging it directly into a standard 120V household power outlet. However, due to its slow charging capabilities, one may choose to pursue the Level 2 charging option. For this purpose, special equipment which allows the vehicle to be plugged into a 240V outlet has to be installed [9]. An in-home Level 2 charger is shown in Figure 1.2.

Figure 1.2 "Car charging with a Level 2 charger [10]"

Table 1.2 compares the Level 1 and Level 2 charge types. Level 1 charging is mostly suited for PHEVs due to their small capacity batteries, whereas Level 2 charging is preferred for BEVs because their battery capacity is larger.

Table 1.2 "Level 1 - Level 2 charge type comparison"

Level 1 charging	**Level 2 charging**
Uses 120V chargers	Uses 240V chargers
Power: 1.92kW	Power: 6.6kW
Power consumption equivalent: toaster	Power consumption equivalent: tumble dryer
Replenishes 5-8 km per hour	Replenishes 32-40 km per hour
Up to 20 hours to recharge a BEV	Up to 7 hours to recharge a BEV
No additional equipment required	Requires additional equipment
Most suitable for PHEVs	Most suitable for BEVs

The new 2020 Tesla model S is shown in Figure 1.3. This luxurious BEV comes with a 100kW lithium-ion battery pack feeding its two electric motors, thus granting it the ability to drive up to approximately 560-600 km within a single charge. A typical 240V charger can fully charge the car at around eight to ten hours but if charged at a Tesla supercharger station (DC charging) with a V3 supercharger (which a number of them is also shown in Figure 1.3, behind the car), charging time drops to 75 minutes [11].

Figure 1.3 "2020 Tesla Model S [12]"

The smart grid is a complex electricity network connecting numerous power generating units and power consumption units, altogether. Unlike the traditional electrical grid, the smart grid, allows two-way communication of electrical data and other types of data in real time between the providers and the consumers. The communication can be both wired (through cables) and wireless (through wireless communication technologies such as WiMAX) [13] [14]. A typical smart grid architecture is displayed in Figure 1.4.

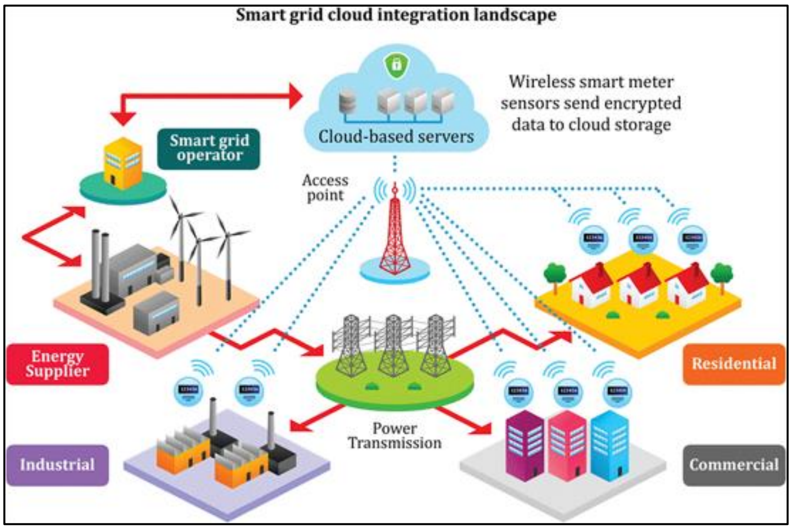

Figure 1.4 "smart grid architecture [15]"

The smart grid utilizes computer technology to automate, monitor and control the flow of data in order to improve the flexibility, reliability, efficiency, security and safety of the network. Additional benefits include the provision of information to consumers with regard to their energy usage and the incorporation of vehicle charging [13] [16].

Despite its many advantages, the smart grid possesses a number of challenges, such as power losses, overloads and voltage fluctuations caused by the uncontrolled charging of PEVs during peak hours (usually when the PEV owners return home from work) where the demand is higher [17].

Time-of-Use allows energy providers to set the electricity price rates according to different time periods. There are three Time-of-Use periods; Off-Peak, Shoulder (also called Mid-Peak), and Peak. The rates are low during off-peak hours (i.e. at night, on weekends, and on holidays), higher during shoulder hours and highest during peak hours (afternoon and early evening hours, on weekdays). Figure 1.5 presents the Time-of-Use periods [18].

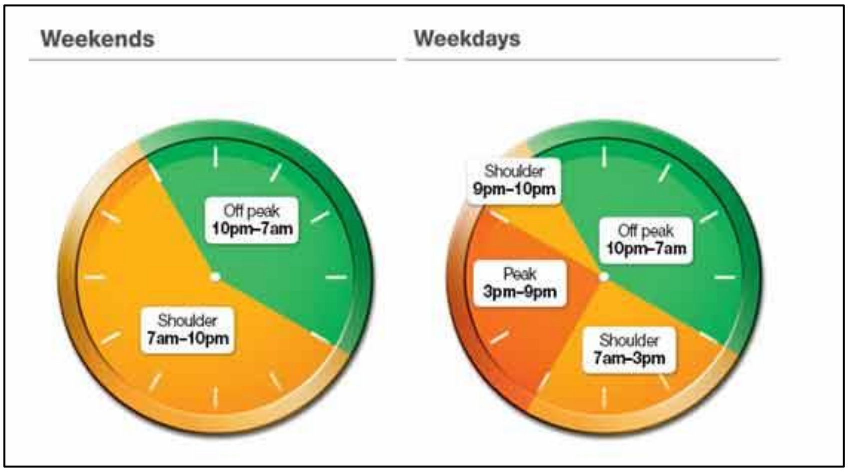

Figure 1.5 "Time-of-Use periods [19]"

Demand side management (DSM) is an important component of the smart grid which can improve the load and reduce peak demand [17]. Its various strategies can motivate the electricity consumers to modify their energy usage such that both the customers and the providers gain benefits [20]. The strategies include Peak Clipping, Valley Filling, Load Shifting, Strategic Conservation, Strategic Load Growth, and Flexible Load Shape.

Peak clipping refers to the reduction of peak demand by applying the Time-of-Use rates mentioned in the previous paragraph. Valley filling focuses on increasing the off-peak loads. Both peak clipping and valley filling aim to achieve balance between peak load levels and valley load levels. Load shifting which is considered as the most effective DSM strategy, refers to shifting the load from peak hours to off-peak hours [21] [22]. Strategic conservation aims to optimize the load shape by applying demand reduction methods directly at consumer facilities. Strategic load growth optimizes the load shape when the demand increases beyond the valley filling strategy. Flexible load shape is mostly related to the reliability of the Smart Grid. The load shape can become flexible for consumers who are willing to allow their loads to be controlled in exchange for a number of incentives [21] [22].

The basic concept of the mentioned DSM strategies is shown diagrammatically in Figure 1.6.

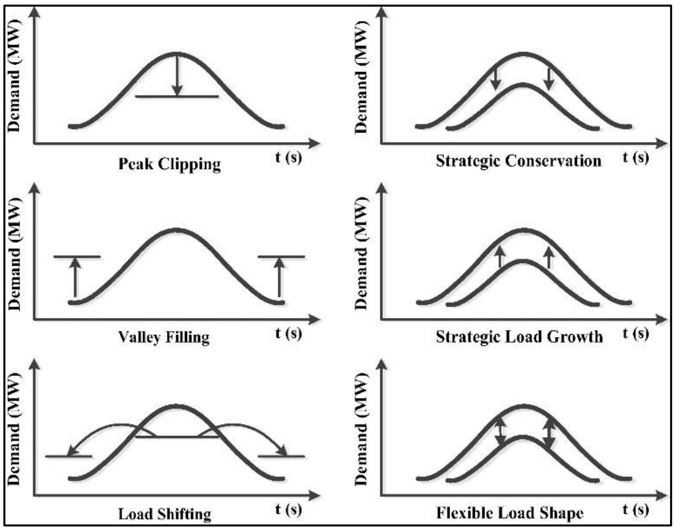

Figure 1.6 “Demand Side Management strategies [23]”

Data analytics is a broad field of data science. It involves the processing of raw data to extract information by discovering patterns in the data. Data visualization comprises an important aspect of data analytics as it leads to a better understanding of the extracted information. Data analytics enjoys applicability in healthcare, crime prevention and environmental protection. The resulting information facilitates decision making, the improvement of customer services, consequently: improved marketing strategies [24] [25].

The data collected and processed by a smart grid originates from a variety of sources, such as distribution stations, distribution switch stations and electricity meters. Analysis of these data using data analytics techniques can lead to better power plants scheduling policies [26].

The main challenge in smart grids is to manage the large electricity demand of the consumers during peak hours, especially those owning a PEV since more and more people are beginning to use electric vehicles. The main objective of the present thesis is to investigate the impact uncoordinated PEV charging has on a power distribution network and to suggest a DSM strategy for decreasing energy demand during peak hours. The results presented comprise the output of Descriptive Statistics and Exploratory Data Analytics type data processing operations conducted in the R/RStudio environment.

In Chapter 2, previous studies and their solutions are discussed. Furthermore, a diagram appearing on [27] with the data used in [28] is re-created to ensure harmonization and sequentiality with the previous studies.

Chapter 3 explains the concepts of descriptive statistics and exploratory data analysis. Then, it briefly reviews the most popular tools used for such purposes. Lastly, it presents the various packages and R functions used during the development of the code for the manipulation of the data and for the diagrammatic representation of the results obtained.

Chapter 4 and the chapters that follow report on the contribution made by the current thesis project. By extending the data analytics conducted in [28], additional information is revealed and presented via a number of new diagrams.

In Chapter 5, data are organized in a new R structure/object facilitating the establishment of a new perspective that makes possible the incorporation of time zones in data exploration and analysis.

In Chapter 6, a load shifting scenario with energy saving peak shaving strategies is proposed, consisting of two cases which are thoroughly investigated. Two energy saving approaches are also investigated in a separate section of this chapter.

In Chapter 7, which is the final chapter of the thesis, the conclusions and possible plans for future work are discussed.

In the Appendix, three of the most representative algorithms implemented in R are presented and they are considered in detail.

Chapter 2: Related Work

Introduction

The bibliographic retrospective of relevant research is presented in this chapter. The related work includes studying the charging behavior of plug-in electric vehicles and how they affect the energy supply network.

The final section of this chapter analyzes an R diagram which was created in RStudio and is based on the diagrams of [27] by using the data found in [28].

2.1 Impact of uncoordinated PEV charging [27]

In [27], the impact of uncoordinated PEV charging over the smart grid is studied.

To better assess the impact, an energy consumption simulation was created where no control or coordination strategies are implemented and it is modeled after real-world residential and vehicle power demand with a 10-minute resolution. It features 200 households with 502 individuals and 348 vehicles located in the United States.

Two alternative charging power levels are adopted for residential PEV charging; Level 1 (L1) and Level 2 (L2). L1 PEV charges operate at 1.92kW (or 1920W) and L2 PEV charges operate at 6.6kW (or 6600W).

Figure 2.1 shows the aggregate energy demand over the course of a sample week for 50% PEV market share with L1 PEV charging (first diagram) and L2 PEV charging (second diagram), where 174 out of the 348 vehicles are considered as PEV. The areas in blue color refer to the household consumption only while the red areas indicate the additional demand caused by PEV charging. PEV charging causes the overall energy demand to be significantly higher than that of the household only consumption, especially in the case of the L2 PEV charging where the curve is more abrupt.

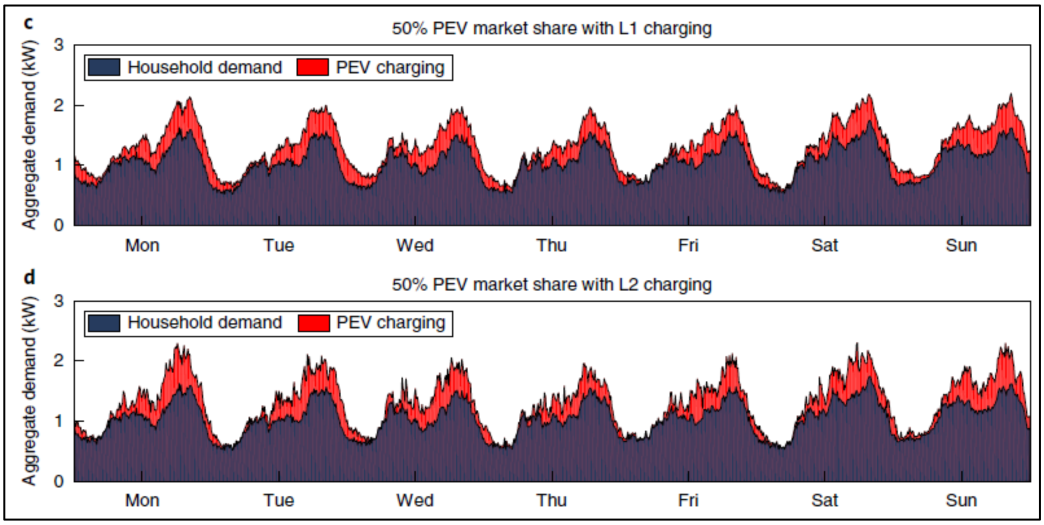

Figure 2.1 "50% PEV market share with L1 and L2 charging [27]"

2.2 Impact of uncoordinated PHEV charging [29]

In [29], the impact of PHEV charging on the smart grid of China is analyzed, assuming that they are charging at home through being plugged into an outlet. The paper considers peak hours to extend from 6:00pm to 10:00pm and the off-peak hours to extend from 10:00pm to 7:00am.

According to Chinese standards where the development of PHEVs is slower than that of Europe, the typical levels for residential charging are 3.0kW and 5.0kW. In a real-life scenario, it is unlikely that the same type of battery will be used for all vehicles.

The simplest charging scenario is the uncoordinated charging where PHEV owners can charge their vehicles at any time without taking any factors such as electricity price into consideration. People tend to charge their vehicles as soon as a power outlet is available, such as when they return home from work or during their work at the office.

A number of serious problems would arise if a lot of PHEVs are charging at the same time, especially during peak hours where the smart grid’s power might be thrusted above its limits.

Coordinated charging could improve the condition of the smart grid’s load and cost. The paper states that the best time for PHEVs to charge is at night, during off-peak hours, where the power demand is lower.

Figure 2.2 displays the load profiles for uncoordinated charging with 5%, 10%, and 20% market share.

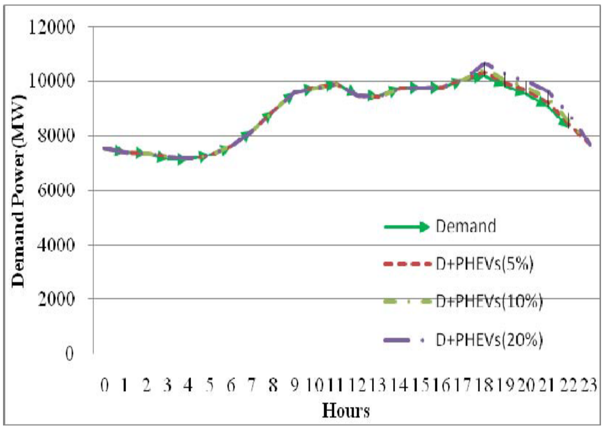

Figure 2.2 "Load profile for uncoordinated charging [29]"

2.3 Peak shaving for coordinated charging and discharging [17]

In [17], peak shaving (or else peak clipping) is proposed as a method to overcome possible overloads caused by the uncoordinated charging of PEVs. In the designated strategy, PEV owners are allowed to choose their preferred charging and discharging time zones according to their priority.

An algorithm is proposed for managing the coordination of PEV charging, assuming that the PEVs will be controlled automatically through sending and receiving messages which will schedule their charging. However, the algorithm presents some issues, such as energy prices in different hours and PEV owners' time zone preferences for charging their vehicles.

Three distinct time zones for charging and discharging are proposed for the PEV owners to choose from. During the "Red time zone" (18:00-22:00), PEV owners who choose to charge their vehicle will pay a higher tariff rate, due to the zone's high energy demand. "Blue time zone" (22:00-1:00) charges at a lower tariff rate and "Green time zone" (1:00-8:00) is the cheapest of the three.

The paper's simulations derive from a modified IEEE test system which includes a number of feeders, transformers and a total of 1537 households where consumers can charge only one PEV. A residential load profile was implemented in the model, assuming that the peak load of a consumer is 1.5kW.

Figure 2.3 displays the simulation's results of four different scenarios; without PEV charging, with uncoordinated charging and discharging, without coordinated charging and with coordinated charging and discharging for the market share of 31%. The results reveal that coordinated charging and discharging can reduce peak demand in the smart grid in comparison to the other scenarios.

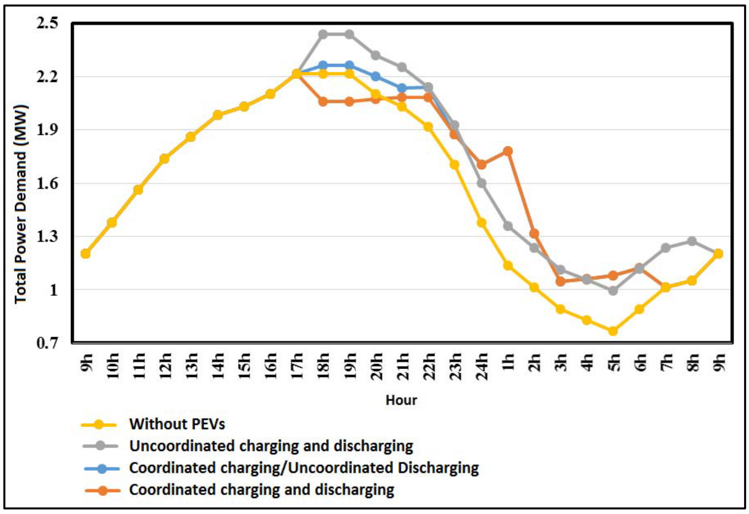

Figure 2.3 "Total power demand in different scenarios (31% market share) [17]"

2.4 Coordinated PEV charging with fuzzy logic control system [30]

[30] suggests a fuzzy logic control system which determines the charging and discharging priorities of the PEVs based on their battery level and the electricity tariff, aiming to keep the balance between the PEV owners' needs and the smart grid's needs. The system uses real-time communication between the smart grid and the PEVs to gain knowledge about the battery's state and to determine the electricity costs. A set of fuzzy rules designate the charging power level of each PEV.

By utilizing such a system, charging shifts to the off-peak hours and discharging shifts to peak hours, thus decreasing the load from the smart grid and preventing transformers and cables from overloading.

A 400V low voltage distribution network based on a real distribution network in Egypt which includes 96 households is selected for the execution of the simulations. The PEVs' battery capabilities are modeled after the Nissan Leaf lithium-ion battery which has a maximum capacity of 24kWh. The PEVs are charged at a rating of 6.6kW when connected to the charger.

Four different PEV charging scenarios are analyzed in the paper. The first one is called "base case", the distribution network supplies only residential consumers. In the second one, called "uncontrolled plug-in EVs charging", PEVs charge as soon as they return to the residence, without any restrictions. The fuzzy logic control system is implemented to operate the charging of the PEV vehicles in the third scenario called "controlled plug-in EVs charging" and to operate the charging and discharging in the fourth scenario called and "controlled plug-in EVs charging/discharging".

In Figure 2.4, the total power demand of the four scenarios is shown. In the "uncontrolled" scenario, the demand increased more than the "base case", especially at 20:00. In the "controlled charging" scenario, the PEV owners were motivated to charge their vehicle during the off-peak hours which come at a lower electricity cost, and stop charging when the electricity cost is higher. In the "controlled charging/discharging" scenario, vehicles charge at a low price and discharge during high electricity demand which led to its curve being lower that the "base case" during peak hours.

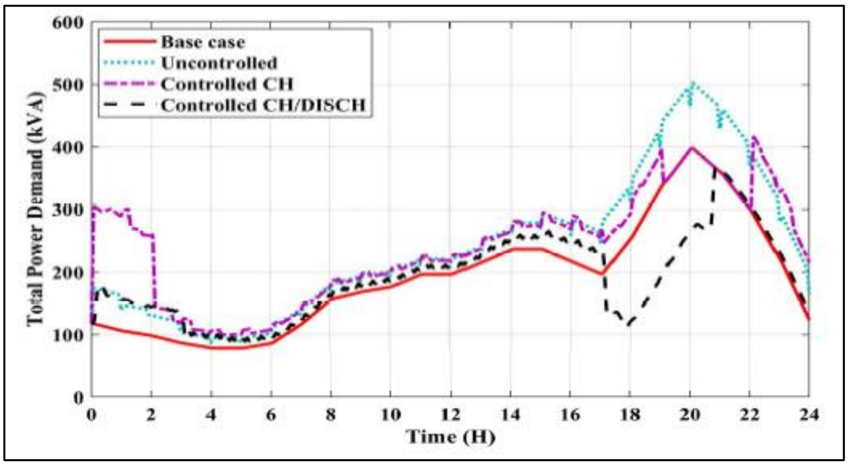

Figure 2.4 "Total Power Demand [30]"

2.5 Impact of PEV charging in low voltage smart grid [31]

Paper [31] studies the impact of PEV charging on a low voltage smart grid modeled after a typical Hungarian low voltage distribution network which supplies a total of 139 households. The smart grid also contains three main feeders and a transformer. A set of real-world data was used to model the household consumption profiles. The PEVs' batteries are modeled after the Nissan Leaf lithium-ion battery with a maximum capacity of 24kWh. The PEVs are charged at 3.3kW.

Two different scenarios are studied; uncoordinated charging and delayed charging. Both scenarios make use of different market shares; 20%, 40% and 60%. In the uncoordinated charging scenario, PEVs start charging as soon as the owners get back home from work. The arrival time is based on the National Household Travel Survey 2009. It was found that most of the vehicles return home during the hours 17:00-19:00 and almost no vehicles will be available for charging during the hours 1:00-7:00. In the delayed charging scenario, the vehicles will start charging at four different times of the off-peak period; 22:00, 23:00, 3:00 and 4:00 and the priority will be given to the vehicles a lower battery charge state, aiming to have all the vehicles completely charged before 7:00.

The PEV charging will split along the off-peak hours to avoid the creation of new peaks. By utilizing a pricing system where charging a PEV during the off-peak hours will cost less than charging it at any other time, owners will be motivated to charge their vehicle during the off-peak hours.

Figure 2.5 shows the effect of uncoordinated charging on the transformer with three different market share levels (20%, 40% and 60%) including a case where no PEVs are present. In the “No PEV” case, the transformer loads up to 66%. Since the PEV charging happens to occur during the network’s peak hours, the peaks become higher as the market share increases and the maximum load of the transformer reaches up to about 110% when the market share is at 60%. In Figure 2.6 which shows the transformer loading of the delayed charging scenario, the maximum load of the 60% market share level drops to 77% as opposed to the 110% of the previous scenario. Due to PEV charging during the off-peak hours, a valley is formed on the load profile during the rest of the hours.

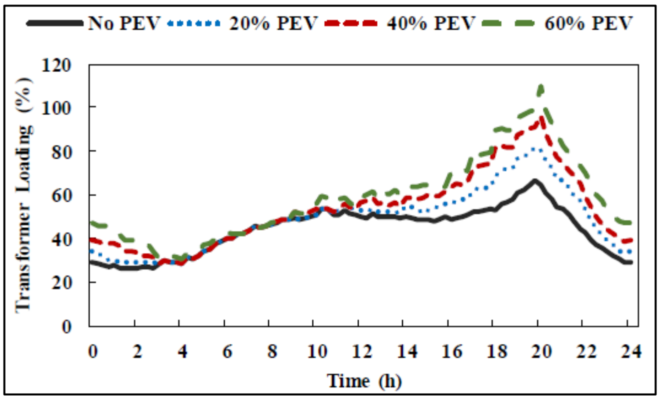

Figure 2.5 "Transformer loading with uncoordinated charging [31]"

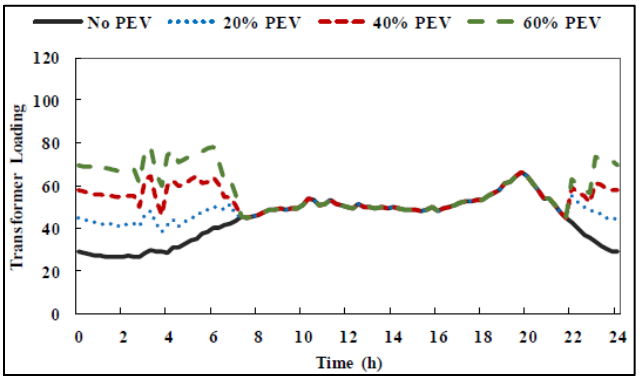

Figure 2.6 "Transformer loading with Delayed charging [31]"

2.6 The first diagram in RStudio

In this section, a diagram similar to those found in page 4 of [27] will be created in RStudio by using the data from [28]. Unlike the paper's diagrams which include market share cases during a sample week for the L1 and L2 PEV charges, the RStudio diagram will focus on the average yearly power demand by day for both charge types without constructing any PEV market share scenarios.

In order to successfully import the data into RStudio, the three *.xlsx* files had to be converted to *.csv* format and their contents had to be slightly modified. In the file containing the household consumption, the first row which contained the words "*Residential Electricity Demand [W]*" in the first cell, was entirely removed, whereas in the files containing the vehicles' energy demand when they are charged using the L1 or L2 charge types, the first row containing the words "*Residential PEV Charging [W]*" was removed and the information in the second row (the household ids) was combined with that of the third row (*vehicle ids*) to form a single household-vehicle id, from "H001-V001" up to "H200-V348". Appending zeros in front of single and double digit numbers is an effective method to sort the ids alphabetically in the correct order, should the circumstances require it.

The files were now ready to be imported into RStudio by calling the function *read.csv2()* that produces a data frame of the argument inside of the parenthesis which has to be a *.csv* file that uses the semicolon as the delimiter. The first row of the *.csv* file is used for naming the columns.

In order to have a first look at the data, a plot visualizing the average yearly demand by day for all households and vehicles was created (Figure 2.7, below).

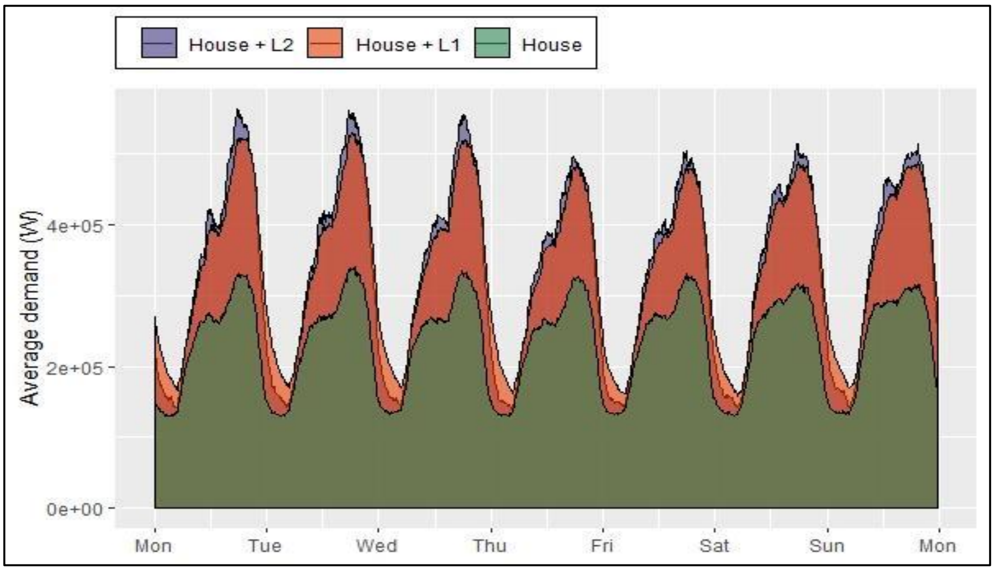

Figure 2.7 "Average yearly power demand by day"

In Figure 2.7, “*House*” refers to household consumption only, whereas “*House + L1*” and “*House + L2*” refer to household consumption plus the vehicles’ power demand during charging for L1 and L2 charge types, respectively.

The power demand is seen to be higher for the L2 PEV charging, but faster when compared to that of the L1 PEV charging. The concave area between two successive days indicates that the “*House + L1*” curve always lies above the “*House + L2*” curve during those hours.

It is also noted that the power demand is higher during peak hours of the day (14:00-20:00), and lower during off-peak hours (22:00-07:00) for both charging types.

It should be noted that at a first glance, it may seem peculiar that the *Monday*, *Tuesday* and *Wednesday* curves look very similar to each other. The same applies to the (*Thursday*, *Friday*) and (*Saturday*, *Sunday*) pairs. The reason behind this peculiar behavior of the data is discussed in the following two chapters.

Conclusion

Studies have proven that uncoordinated plug-in electric vehicle charging can cause serious problems to the smart grid such as overloads, power losses and voltage fluctuations.

Several solutions for improving the load profile and providing cost benefits to PEV owners were proposed. [29] suggests coordinated PEV charging during the night, where the electricity demand is lower, [17] suggest peak shaving for coordinated charging and discharging, [30] suggests the implementation of a fuzzy logic control system to optimize controlled charging and discharging and [31] suggests a delayed charging strategy.

Moreover, an R diagram analogous to the basic diagram in [27] is also constructed using the data from [28]. It is shown diagrammatically that the L2 PEV charging is shorter and of higher power demand when compared to the L1 PEV charging. The diagram also reveals a noticeable similarity of household plus PEV charging power demands in each and every day of the week.

Chapter 3: Exploratory data analysis and descriptive statistics

Introduction

Exploratory data analysis (EDA) refers to the investigation of the data in order to discover patterns, missing values and outliers, test hypothesis and check assumptions by using summary statistics and graphs (such as histograms and scatterplots) to visualize the results before making decisions about the next step in the data processing [32].

In descriptive statistics, the data is summarized and organized in a way that makes its understanding easier [33]. It provides useful information about the dataset's various measurements and indicate any possible relationships between the variables [34].

This chapter takes a look of some of the most well-known tools used in EDA and descriptive statistics. It also discusses the various packages and in-built functions of R which was the preferred data analytics programming environment.

3.1 Descriptive statistics and EDA tools

Apache spark is a powerful open source platform which focuses especially on data analytics involving large amounts of unstructured data [35]. It executes applications in Hadoop clusters [36] and it can work well with all major programming languages like SQL, Java, Python and R [37].

Python is an open source object-oriented programming language [36]. It has more than 800.000 available packages and can handle different types of data analysis [38]. Packages like numpy and scipy are used for data statistics [35] while matplotlib, plotly and seaborn are popular data visualization packages [38].

R is the most popular data analytics tool [35]. It is an open source programming language which focuses on statistics and data visualization with more than 15.000 packages available [38]. R also supports object-oriented features [39] and can handle large datasets well [35].

R was chosen as the data analytics tool of the current thesis, not only because of its popularity but also because it is easy to learn and use and it provides packages that make the data manipulation a lot simpler. The code was developed in RStudio; an open source IDE for R.

3.2 R packages and in-built functions

The packages that were installed in RStudio for the needs of the thesis are ggplot2, dplyr, lubridate and glue. Those packages are part of the tidyverse, which is a collection of R packages designed for data science [40]. Figure 3.1 shows some of tidyverse's most popular packages.

Figure 3.1 "Popular tidyverse packages [41]"

ggplot2 is a graphics creating system. The user provides the data, maps the variables to aesthetics, declares the graphical primitives and adds other optional layers and then ggplot2 handles the visualization of the data [42]. The graphical primitives that were used in this thesis are the following; *geom_bar()*, which creates a bar chart, *geom_boxplot()*, which creates a boxplot, *geom_line()*, which creates a line graph, *geom_point()*, which creates a scatterplot and *geom_ribbon()*, which creates a ribbon graph.

dplyr allows the user to perform useful data manipulation tasks as well as to make use of pipelines by adding the characters "%>%" between two functions [43]. Some of dplyr's functions that were used in this thesis are the following; *summarize()* with the argument *mean()* or *sum()* which calculates the mean or sum of a given table's column. *summarize()* can work well with *group_by()* which creates a grouped copy of the table. *filter()* is very useful in extracting rows that match one or more logical conditions. The final two functions of dplyr that were used are *intersect()* which takes two tables and returns the rows which appear in both tables and *setdiff()*, which also takes two tables but returns the rows that appear only in the first table.

In R, dates and times are stored as the number of seconds that have passed since January 1st,1970. Package lubridate can efficiently manage such dates and times [44]. The particular package provides fast parsing with functions such as *dmy()* and *dmy_hms()*. Other simple functions for extracting certain components from dates and times are *year()*, *month()*, *mday()*, *hour()*, *minute()* and *second()*. With lubridate, the user can also add and subtract dates.

The fourth and final package that will be discussed in this chapter is glue. This handy package allows the user to concatenate strings. Any expression enclosed in braces is considered as R code which is then evaluated and inserted into the string. Compared to *paste()* (an in-built function of R which also concatenates strings), glue is easier and less time consuming to write [45]. In this thesis, glue's function *glue()* was used solely to concatenate date components.

Other notable functions which are in-built RStudio are the following; *read.csv2()* loads semicolon delimited .csv files into RStudio's workspace as data frames. *str_replace_all()* replaces all matched pattern with another string and *sub()* which replaces only the first match. *cbind()* and *rbind()* combine multiple tables by column or by row, respectively. Function *match()* was used for finding the position where the first match of two vectors was found. *seq()*, which generates sequences, was used for creating time intervals. Lastly, the function *factor()* was used for defining the order of categorical data prior to their visualization, that would otherwise be sorted alphabetically by default. For example, it was used to order the days correctly before their visualization in bar charts and boxplots.

Conclusion

In this chapter, Apache Spark, Python and R which are popular data analytics tools were discussed. All different R packages that were used in this thesis were also briefly analyzed; ggplot2 creates elegant graphs, dplyr can manipulate data, lubridate handles dates and times, while glue is an alternate way of concatenating strings.

Chapter 4: Exploratory Data Analysis

Introduction

From this point onward, the contribution of the present thesis project begins. In this chapter, the three modified .csv files mentioned in section 2.6 are going to be further analyzed. These files are going to be used throughout the remainder of the thesis. To find out more about the data's structure located in the files mentioned above, three bar charts, two boxplots and two line plus ribbon graphs were created.

4.1 Total monthly power demand examination

Figure 4.1 displays the total power demand per month. The blue bars refer to the household consumption plus the L2 PEV charging. The red bars refer to the household consumption plus the L1 PEV charging. Lastly, the green bars refer to the household consumption only. January, July and December are the most energy-intensive months, followed by February, March and August.

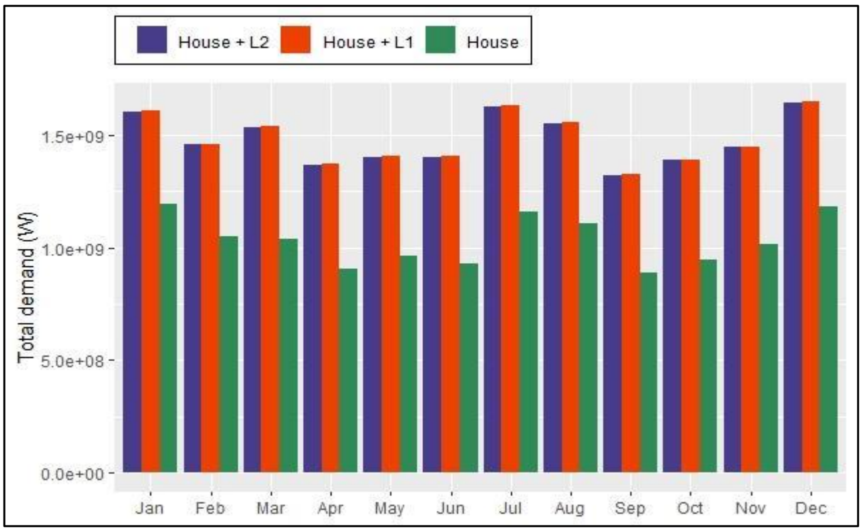

Figure 4.1 "Total power demand per month"

4.2 Average yearly power demand examination

In the next two Figures, the average yearly charging behavior of vehicles using L1 and L2 charge types are considered. They are separated from the average yearly household power consumption. Figure 4.2 presents the results relating to weekday time periods. Figure 4.3 presents the results for weekend days (Saturday and Sunday).

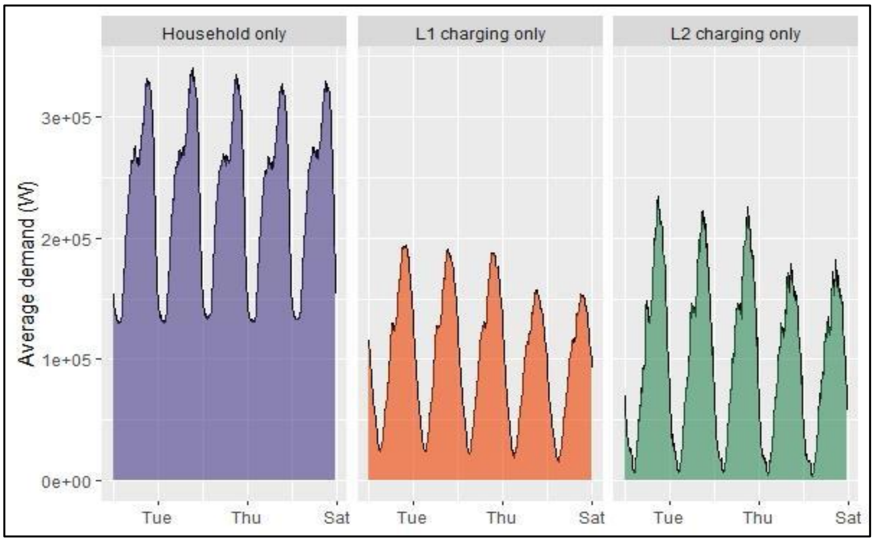

Figure 4.2 "Average yearly power demand on working days"

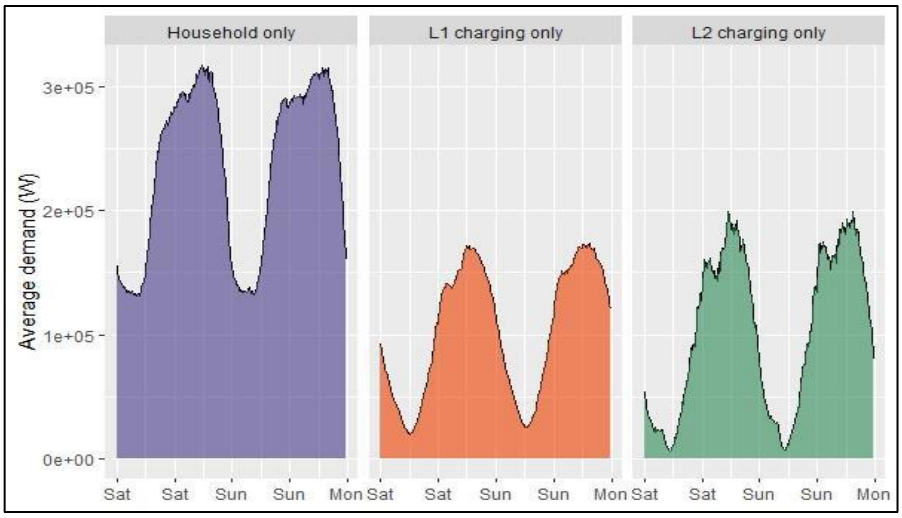

Figure 4.3 "Average yearly power demand on weekend days"

Houses have appliances such as deep freezers and refrigerators which operate on a 24/7 basis. Devices that are plugged into sockets can still consume an amount of fixed energy, even if they remain on standby mode for hours [46]. This explains the rectangular box shape in the bottom of the diagram. It depicts the power demand for household appliances only (the leftmost graphs in Figure 4.2 and Figure 4.3).

It was of high concern that in the same two Figures the curves' shape is almost identical to each other for each day of the week. To be more precise, *Monday's* curve looks identical to *Tuesday's* and *Wednesday's*. Even the peaks have about the same height. The same applies to the (*Thursday*, *Friday*) and (*Saturday*, *Sunday*) pairs.

To examine these observations in more detail, bar charts and boxplots were created. In Figure 4.4 and Figure 4.5 the bar chart and boxplot of the average yearly power demand on working days is depicted. In Figure 4.6 and Figure 4.7 the bar chart and boxplot of the average yearly power demand on weekend days is depicted.

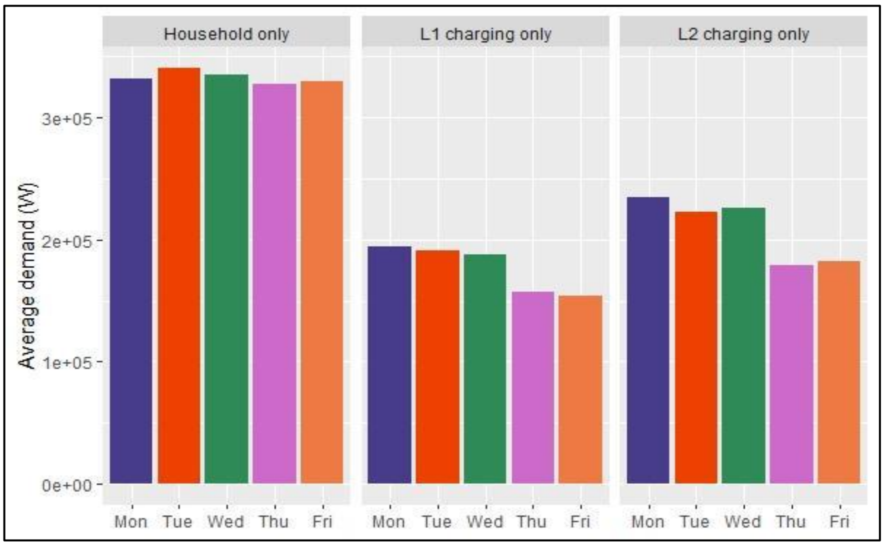

Figure 4.4 "Average yearly power demand on working days (bar chart)"

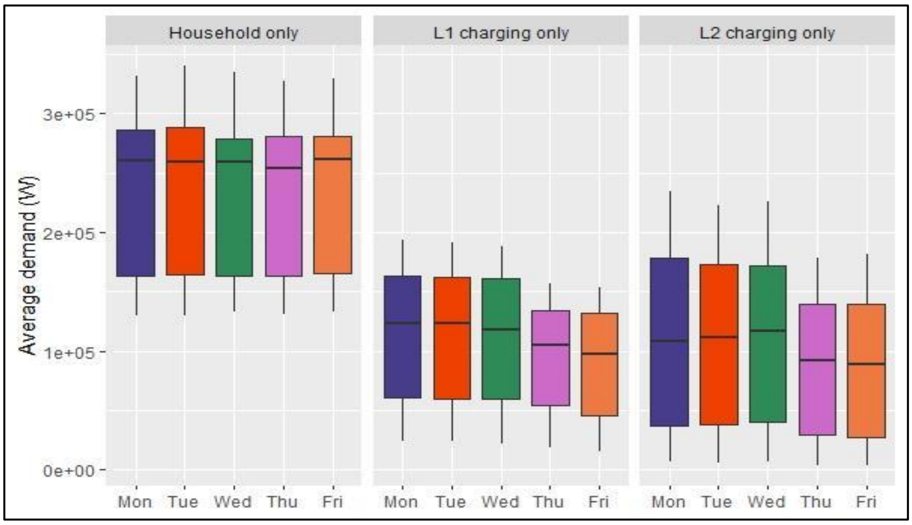

Figure 4.5 "Average yearly power demand on working days (boxplot)"

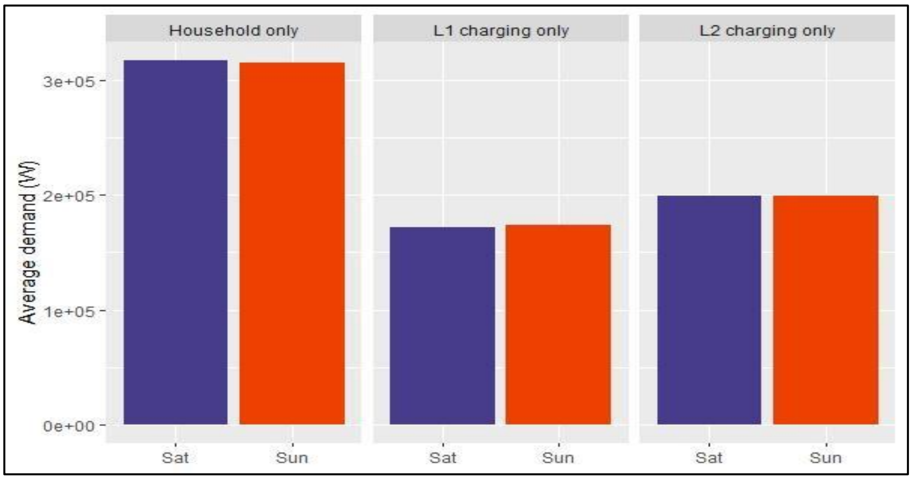

Figure 4.6 "Average yearly power demand on weekend days (bar chart)"

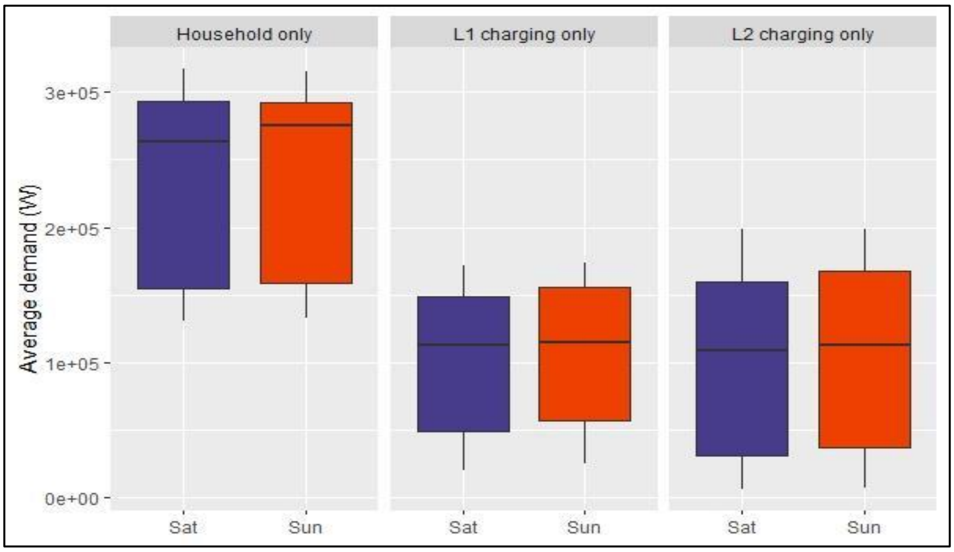

Figure 4.7 "Average yearly power demand on weekend days (boxplot)"

It is now clearly visible that each day holds similar power demand values in the *"Household only"* bar charts of Figure 4.4 and Figure 4.6. In the *"L1 charging only"* bar charts, the power demand of *Monday* through *Wednesday* has similar values and the same applies to the (*Thursday*, *Friday*) and to the (Saturday, Sunday) pair. The same observation occurs in the *"L2 charging only"* bar charts.

As for the boxplots of Figure 4.5 and Figure 4.7, the whiskers, the first and third quartiles and even the median of each day in the *"Household only"* diagrams are located in similar positions. In the *"L1 charging only"* diagrams, the boxplots of *Monday* through *Wednesday*, the pairs (*Thursday*, *Friday*) and (*Saturday*, *Sunday*) look identical. The same behavior occurs in the *"L2 charging only"* boxplots. There are no outliers in any of these diagrams.

Conclusion

In this chapter, the contents of the three .csv files were examined after the creation of various bar charts, boxplots and line plus ribbon diagrams. It was found that January, July and December were the highest in total power demand months. By examining the average yearly power demand on working days and weekend days, it was found that the curves (Figure 4.2 and Figure 4.3) and the values (Figure 4.4, Figure 4.5, Figure 4.6 and Figure 4.7) of the *"Household only"*, *"L1 charging only"* and *"L2 charging only"* diagrams are similar to each other. This could lead to the assumption that the data may present uniformity, which is discussed in-depth in the next chapter.

Chapter 5: The Time Zones structure

Introduction

The purpose of this chapter is to introduce a new data structure which is named "Time Zones". It derives from processing the data located in the two .csv files which contain the PEV charging profiles. The examination of the diagrams that were created using this structure will determine whether or not the data truly presents uniformity. As the name implies, the purpose of "Time Zones" is to categorize the power demand into its appropriate time zone.

Figure 5.1 shows the actual time limits of each time zone. A few details of these time zones have been adjusted to match the needs of the new structure. During working days (or else weekdays), shoulder has been split into two zones which are independent of each other; shoulder 1 and shoulder 2. Shoulder 1 begins at 7:00 and ends at 13:50. Peak begins at 14:00 and ends at 19:50. Shoulder 2 begins at 20:00 and ends at 21:50. Lastly, off-peak begins at 22:00 and ends at 6:50. As for the weekends, shoulder begins at 7:00 and ends at 21:50 and off-peak begins at 22:00 and ends at 6:50. It is also necessary to note that none of the weekdays will be considered as public holidays (for example, Friday 1/1/2010 will not be considered as a public holiday) in order to avoid any further programming complexities. All of the adjustments that were made are presented in Table 5.1 and Table 5.2.

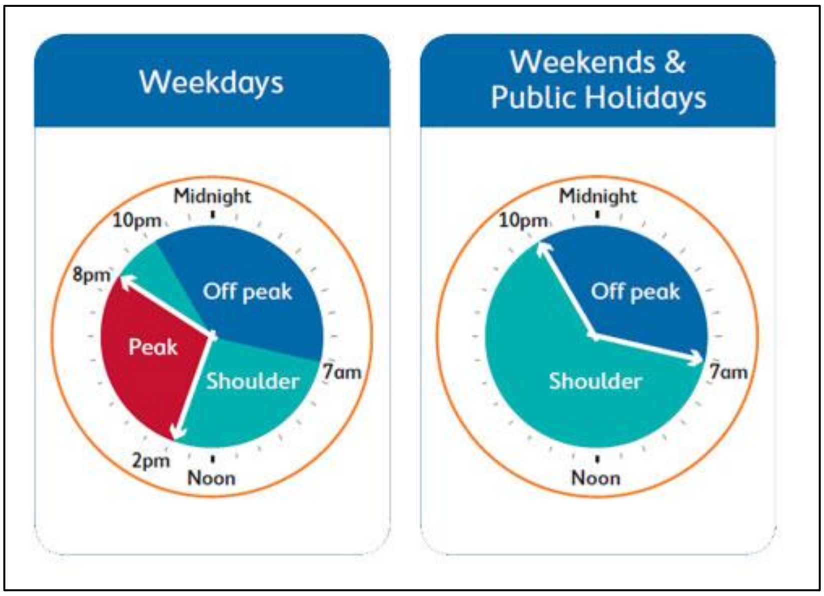

Figure 5.1 "Time zones on weekdays and on weekends [47]"

Table 5.1 ”Time zones adjustments on working days"

Time Zone	Start Time	Stop Time
Shoulder 1	7:00	13:50
Peak	14:00	19:50
Shoulder 2	20:00	21:50
Off Peak	22:00	6:50

Table 5.2 “Time zones adjustments on weekend days"

Time Zone	Start Time	Stop Time
Shoulder 1	7:00	21:50
Off Peak	22:00	6:50

5.1 The structure’s contents

In order to discover more details about the data of the two .csv files which contain the PEV charging profiles, they were examined from another perspective; they were converted into a new data frame which holds information about which time zone represents each of the charges. The duration of each charge is determined by the number of consecutive cells that do not contain a zero (1920 for the L1 PEV charging and 6600 for the L2 PEV charging). This new data frame was given the name “Time Zones” and it was exported as a .csv file for further examination.

In this new structure, the first column contains the charging type (L1 or L2), the second one contains the date (day/month/year) when the charge started. The third one contains the household-vehicle id. The fourth one contains the duration of each charge; it counts how many 10-minutes it lasts for. The fifth column contains the start time (hour and minute) of the charge while the sixth column contains its stop time (hour and minute). The seventh column contains the name of the time zone to which the charge belongs to. The various values of this column are as follows; *“Shoulder 1”*, *“Peak”*, *“Shoulder 2”* and *“Off Peak”* for the working days. Here, the shoulder time zone was split into two distinct time zones; shoulder 1, which is placed before peak zone and shoulder 2, which is placed after peak zone. Since shoulder 1 and shoulder 2 have different names, they are treated as separate time zones. As for the weekend days, the two available values are *“Shoulder”* and *“Off Peak”*. The eighth column of the new structure distinguishes the days into working days with the letter “W” and weekend days with the letter “H”. The ninth column contains the kWh of the charge.

Should a charge exceed the limits of the time zone at which the vehicle began charging, it will split into charges of shorter duration so that for each row of the new data frame, every charge will belong to a single time zone. When a charge-split has to occur, the charge’s stop time is set as the stop time of the time zone at which the vehicle began charging. Then, a new row will be created, having its start time set as the start time of the next time zone. If the charge still exceeds the limits of this time zone, the charge will split again, following the same rules.

5.2 kWh distribution examination

After the construction of the new structure, some bar charts were created to visualize its contents. Figure 5.2 and Figure 5.3 depict each time zone's number of charges per day for the L1 and L2 PEV charging respectively. As it was mentioned in the previous section, shoulder, shoulder 1 and shoulder 2 are all treated as different values, therefore each one is represented by a different color.

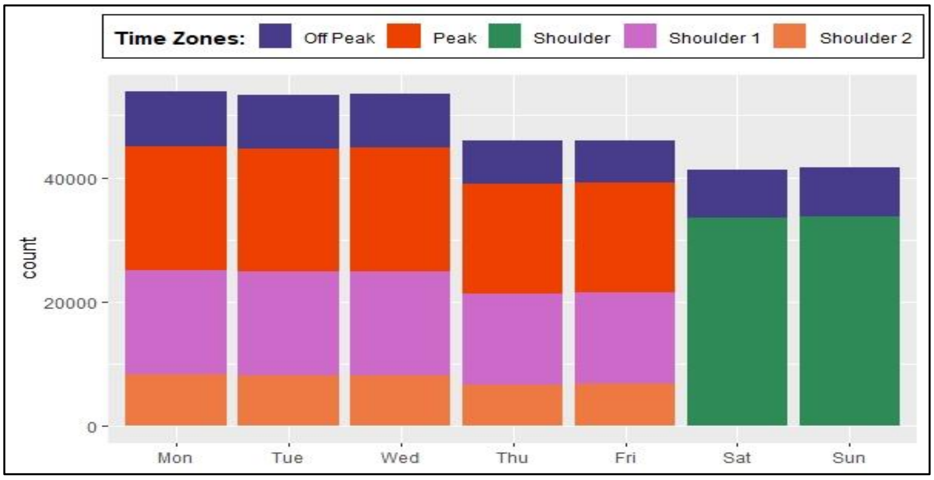

Figure 5.2 "Number of charges per day with L1 charging grouped by time zone"

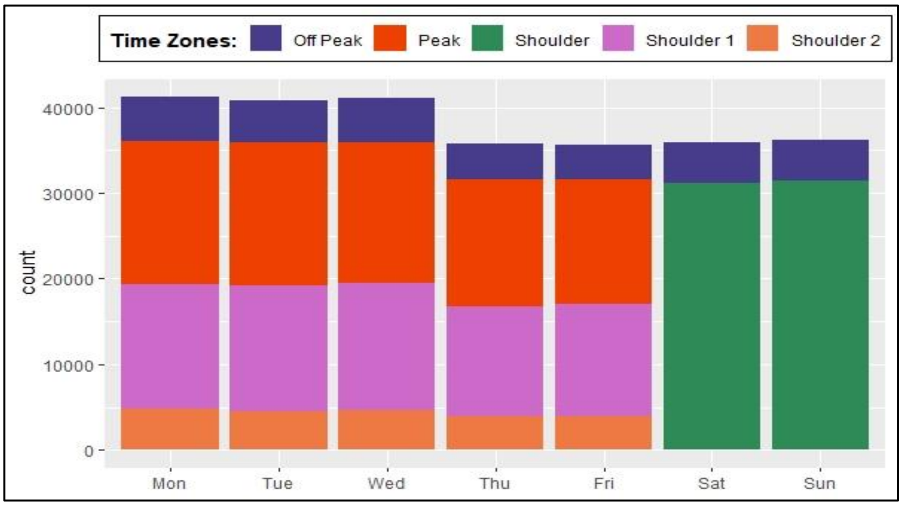

Figure 5.3 "Number of charges per day with L2 charging grouped by time zone"

It is clearly visible that each time zone contains almost the same number of charges every day. For the working days, most charges belong to peak and shoulder 1 followed by off-peak and lastly by shoulder 2. This applies to both the L1 and L2 PEV charging bar charts, which also look almost identical. The difference lies in the fact that the L1 PEV charges are larger in number than the L2 PEV charges. Since the L1 PEV charges are longer in duration, they also have a higher possibility of being split into smaller in duration charges, so that each charge will belong to only a single time zone, thus increasing the number of charges for the particular charge type.

The same phenomenon occurred when the number of the charges by time zone was examined per week for three months belonging to different seasons; January (a winter month), March (a spring month) and August (a summer month). The purpose of this examination was not only to determine the presence of uniformity, but also to estimate whether or not the seasons can affect the number of charges by time zone.

The results are presented in Figure 5.4, Figure 5.5, Figure 5.6 and Figure 5.7, without and with proportion for the L1 and L2 PEV charging. *"Week 5"*'s bar height is much shorter than that of the rest because it represents the last few remaining days of the month. In March's, *"Week 5"*, the shoulder time zone is completely absent because this month's last few days do not include a Saturday or a Sunday. Each week's off-peak includes charges of both working and weekend days.

The seasons do not seem to affect the number of charges. It also does not differ much from that of Figure 5.2 and Figure 5.3. Excluding the Shoulder time zone, most of the charges belong to peak and to shoulder 1, less belong to off-peak and the least belong to shoulder 2. Furthermore, in Figure 5.4 and Figure 5.6, each month's bars have about the same height, except *"Week 5"*, since it contains fewer days and therefore fewer charges. However, even in this case, January's *"Week 5"* bar has about the same height as March's and August's *"Week 5"* bar.

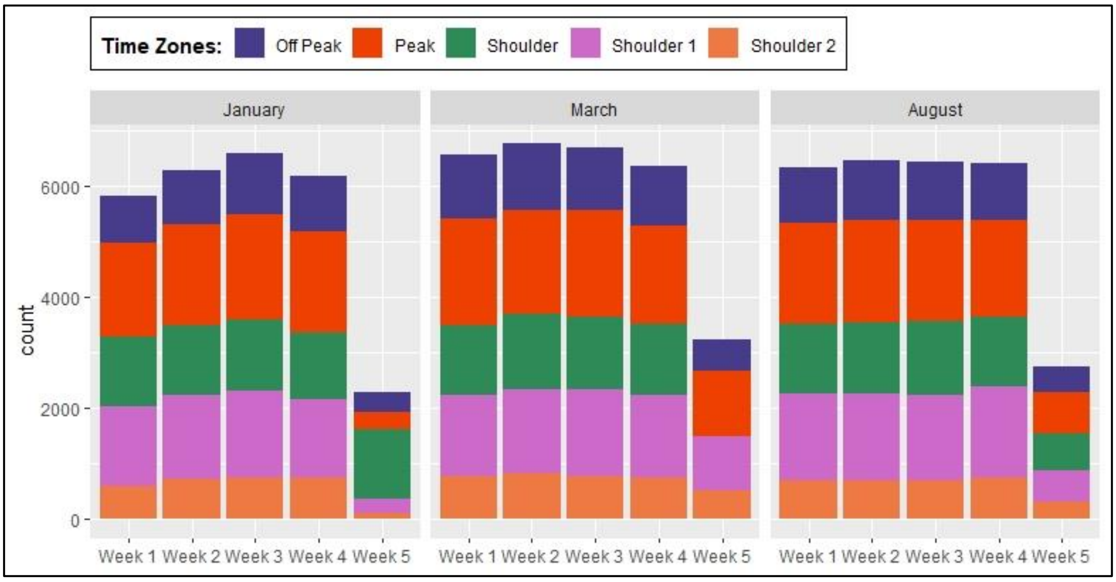

Figure 5.4 "Number of charges per week with L1 charging grouped by time zone"

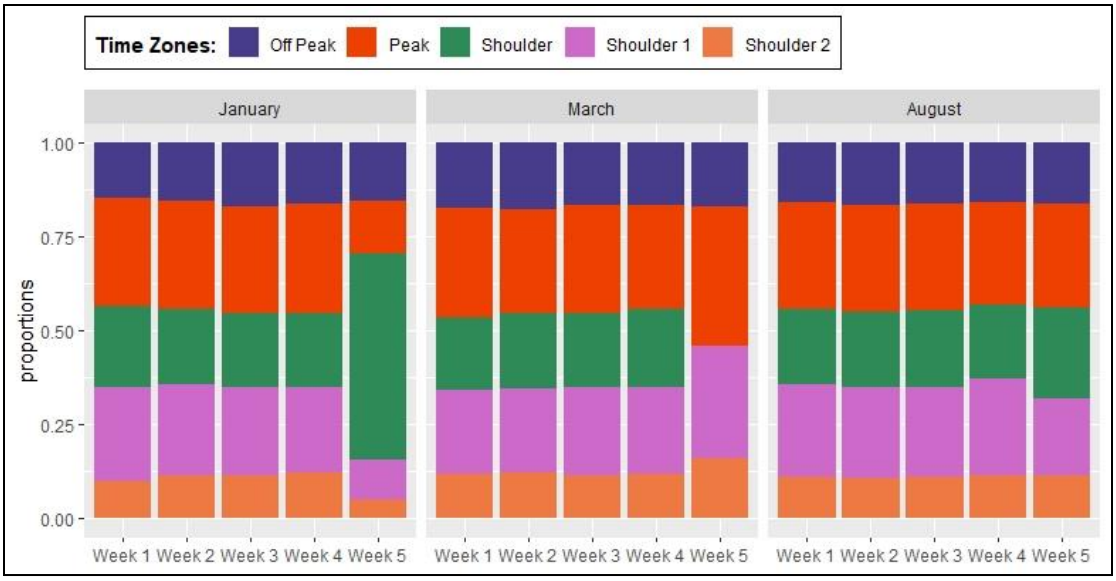

Figure 5.5 "Number of charges per week with L1 charging grouped by time zone (proportions)"

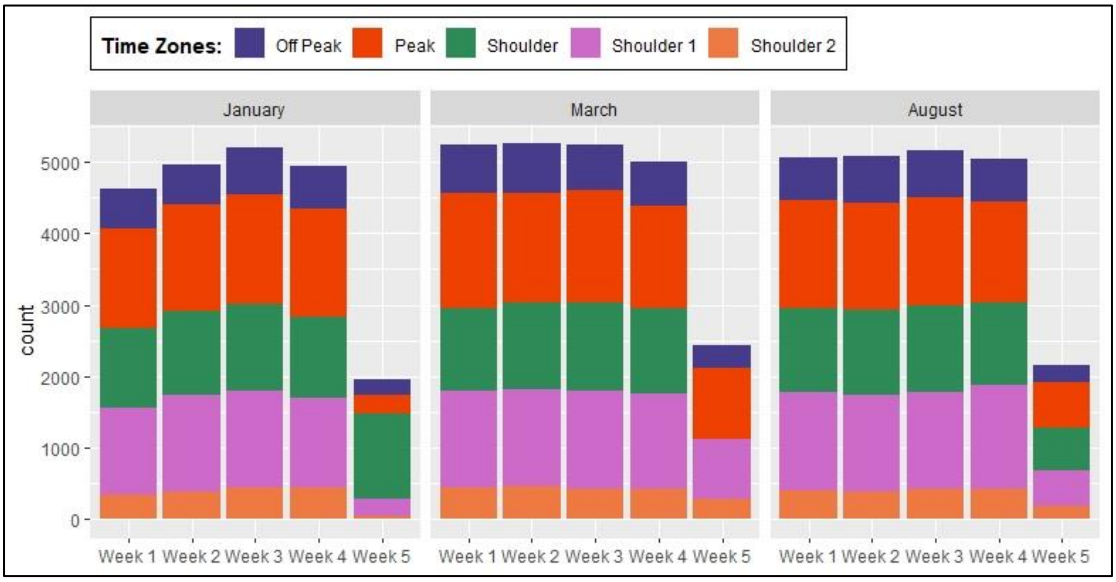

Figure 5.6 "Number of charges per week with L2 charging grouped by time zone"

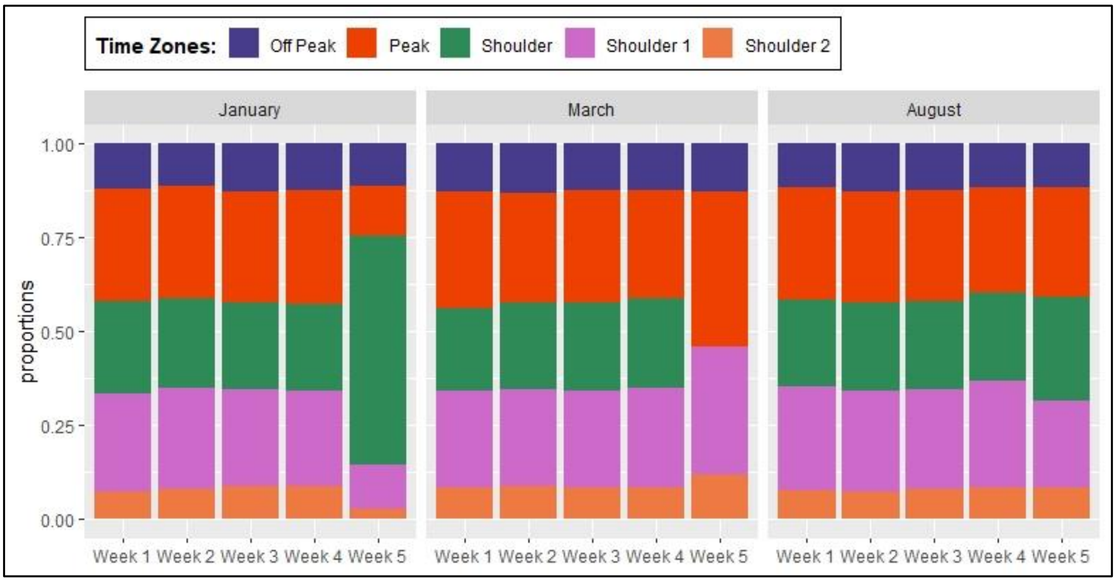

Figure 5.7 "Number of charges per week with L2 charging grouped by time zone (proportions)"

Since the number of charges by time zone is the same per day and also per week, it is undoubtedly true that the data is uniform. It could be possible that the simulation which produced the original data does not successfully represent a realistic PEV charging scenario. This will not be a major drawback for the load shifting research, however, as valuable information can still be obtained.

5.3 Week selection for the load shifting research

Since the data is uniform, there is no need to perform load shifting over the course of the entire year. Instead, the research will elaborate on the span of one week consisting of working days only, as there is no peak zone during the weekends. The chosen week begins on Monday 4/1/2010 0:00 and ends on Friday 8/1/2010 23:50. The reason that a week of January was chosen is because as it was found in Chapter 4, January is one of the thee energy-intensive months. The .csv file containing the Time Zones structure was modified to contain only the five days that were needed so as to minimize its size and to make processing easier and faster.

Conclusion

By examining the new Time Zones structure, it was found that the data does present uniformity and even the diagrams depicting the L1 PEV charging (Figure 5.2, Figure 5.4 and Figure 5.5) look similar to those depicting the L2 PEV charging (Figure 5.3, Figure 5.6 and Figure 5.7). However, this will not negatively affect the load shifting process which is explained in the next chapter. The days Monday 4/1/2010 through Friday 8/1/2010 were chosen for this process.

Chapter 6: Load Shifting with energy saving

Introduction

Peak shaving and load shifting are two major demand side management (DSM) strategies for reducing peak demand and keeping the smart grid's energy cost at a low level. Peak shaving refers to the brief minimization of energy consumption by the consumer while load shifting refers to a brief decrease in energy consumption followed by an increase in demand at a later time, when the smart grid's energy demand is lower [48].

A comparison between peak shaving and load shifting is shown diagrammatically in Figure 6.1. One can see that in the left diagram which corresponds to the load shifting strategy, the peak demand has migrated to the off-peak zones. On the right diagram which corresponds to the peak shaving strategy, the peak demand has been erased.

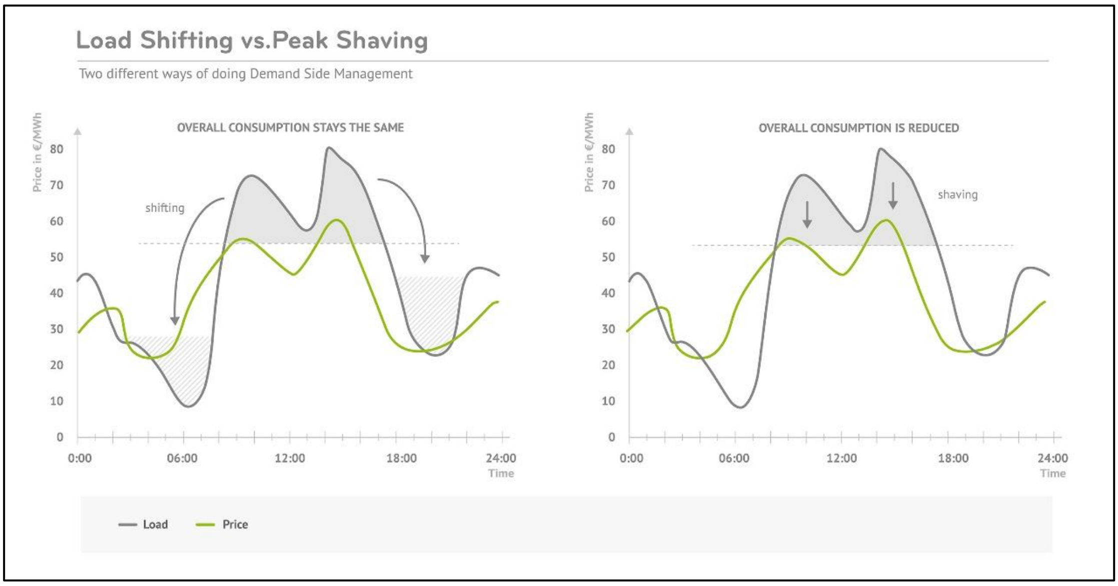

Figure 6.1 "Load Shifting vs. Peak Shaving [48]"

In this chapter, two different load shifting cases will be studied. Each case is divided into five subcases. A diagram was created for each subcase of the two charge types (L1 and L2). It depicts the power demand before and after load shifting. This chapter also analyzes two different energy saving strategies for each of the two load shifting cases.

6.1 Excluded charges from the load shifting process

It was found that 32% of the L1 PEV charges' duration exceed the limit of the zone in which the vehicle began charging. That percentage for the L2 PEV charges is 9%. These charges were marked with the logical "TRUE" in a new column in the Time Zones data frame named and they will not be included in the load shifting procedure.

As it was mentioned in the previous Chapter, when the duration of a charge exceeds the limits of the time zone in which the vehicle began charging, it splits into two or more smaller in duration charges so that for each row of the data frame, every charge will belong to a single time zone. If such a row was to be chosen for the load shifting process, the rest of the other small charges which form the complete charge would not migrate together with it (or they might not even be chosen to migrate at all) because they are treated as separate charges. It would be more realistic if the whole charge rather than parts of it is migrated, but since this is not implemented in the code of this thesis to avoid further complexity, these charges will be excluded from the load shifting process.

6.2 Load shifting Cases

Since percentages of kWh are going to be migrated, the data frame was also updated with a new column which contains the number of kWh for each charge.

Two load shifting cases are going to be examined. Case 1 includes energy demand migration from peak, shoulder 1 and shoulder 2. Case 2 includes energy demand migration from peak zone only.

In both Case 1 and Case 2, a percentage of kWh will be randomly selected to migrate from peak zone. This percentage will be called "a1". From the selected charges, 50% of them will migrate to off-peak, 20% will migrate to shoulder 1 and 30% will migrate to shoulder 2, so that 50% plus 20% plus 30% equals 100%. Moreover, Case 1 also includes a percentage of kWh that will be randomly selected to migrate from shoulder 1. This percentage is called "b1". From those selected charges, 50% of them will migrate to off-peak and the other 50% will migrate to shoulder 2, so that 50% plus 50% equals 100%. Finally, a percentage of kWh will be randomly selected to migrate from shoulder 2 to off-peak, in Case 1 only. This percentage is called "c1".

Five sub cases will be created for Case 1. In the first sub case, a1, b1 and c1 are equal to 10% while the rest of the percentages that were mentioned in the previous paragraph maintain their values. In the second sub case, a1, b1 and c1 are all equal to 20%. In the third sub case they are equal to 30%. In the fourth one they are equal to 40% and in the fifth and final one they are equal to 50%.

Five sub cases will also be created for Case 2. In the first sub case, a1 is equal to 10% while the rest of the percentages maintain their values. In the second sub case, a1 is equal to 20%. In the third sub case it is equal to 30%. In the fourth one it is equal to 40% and in the fifth one it is equal to 50%.

Table 6.1 and Table 6.2 provide a better visualization of the percentages mentioned in the previous three paragraphs.

Table 6.1 "Load shifting percentages for Case 1"

a1) Percentage of kWh that will migrate from peak	Subcase 1: 10% Subcase 2: 20% Subcase 3: 30% Subcase 4: 40% Subcase 5: 50%
- d1) Percentage of a1) that will migrate to off-peak	All subcases: 50%
- e1) Percentage of a1) that will migrate to shoulder 1	All subcases: 20%
- f1) Percentage of a1) that will migrate to shoulder 2	All subcases: 30%
b1) Percentage of kWh that will migrate from shoulder 1	Subcase 1: 10% Subcase 2: 20% Subcase 3: 30% Subcase 4: 40% Subcase 5: 50%
- g1) Percentage of b1) that will migrate to off-peak	All subcases: 50%
- h1) Percentage of b1) that will migrate to shoulder 2	All subcases: 50%
c1) Percentage of kWh that will migrate from shoulder 2	Subcase 1: 10% Subcase 2: 20% Subcase 3: 30% Subcase 4: 40% Subcase 5: 50%

Table 6.2 "Load shifting percentages for Case 2"

a1) Percentage of kWh that will migrate from peak	Subcase 1: 10% Subcase 2: 20% Subcase 3: 30% Subcase 4: 40% Subcase 5: 50%
- d1) Percentage of a1) that will migrate to off-peak	All subcases: 50%
- e1) Percentage of a1) that will migrate to shoulder 1	All subcases: 20%
- f1) Percentage of a1) that will migrate to shoulder 2	All subcases: 30%

The percentages of Table 6.1 and Table 6.2 apply to each and every day. This could be better understood with an explanation about how the algorithm works. First of all, another new column is added to the Time Zones structure containing the index number of each row. Then, the algorithm creates a vector containing all the indexes of a single day, of a given time zone and charge type. From those indexes, a percentage of them (a1, b1 or c1) is selected for the load shifting process while the rest are discarded. By using these indexes, the algorithm gains full access to the corresponding rows of the Time Zones structure, thus altering the time zone of the charges that were chosen to migrate. The same steps are repeated for each one of the remaining days.

Apart from the random assignment of a new zone to the migrated charge, a new start time whose value lies within the limits of the new time zone will be chosen automatically by a random number generator. There is always a possibility that a charge could migrate to a time zone where another charge of the same vehicle already exists. The algorithm will try to locate an empty space where no other charges take place. Sometimes, no such space can be found and it is impossible to completely eradicate the overlap. For this purpose, a new column named “Noise” was created in the data frame that will host the data after load shifting. This column contains the overlapped kWh.

There are cases where the migrated charge’s duration is too big to fit entirely in the new time zone. If something like that occurs, the charge’s start time is set as the start time of the new time zone and any overlapped kWh are recorded in the “Noise” column. However, if the charge does fit in the new time zone then the algorithm enters a “while” loop which tries up to 100 times to find an empty space where no other charges take place, as mentioned in the previous paragraph and if no such space is found then the “Noise” column is updated with the overlapped kWh. A one-element vector named “try”, initialized to zero, is used to count the number of these tries in the “while” loop.

It is not necessary to update the migrated charges’ stop time as this column is no longer going to be needed. It could either be deleted or filled with NAs.

After the completion of the load shifting process for all sub cases of the two main cases, the produced data frames were exported to .csv files.

6.3 Energy saving

Any row of the data frames that were produced after load shifting and have their “Noise” value above zero will be considered as energy saving. Upon the creation of these data frames, the percentage of kWh that is considered as energy saving was calculated for each subcase.

It would be interesting to see what kind of results are produced by the algorithm when the “while” loop which tries to find an empty space where no overlaps occur is not iterated again. This means that the algorithm will not keep looking for an empty space provided that in the first try an overlap occurred. These results are displayed in the four tables below. “try = 1” refers to the maximum number of tries being equal to 1 and “try = 100” refers to the maximum number of tries being equal to 100, while “a1”, “b1” and “c1” refer to the percentages that were presented in Table 6.1 and Table 6.2.

Table 6.3 "Energy saving for Case 1 when try = 1"

a1, b1, c1	L1 charge type	L2 charge type
10%	0.79%	0.29%
20%	1.30%	0.60%
30%	2.00%	0.72%
40%	2.78%	1.04%
50%	3.89%	1.68%

Table 6.4 "Energy saving for Case 1 when try = 100"

a1, b1, c1	L1 charge type	L2 charge type
10%	0.47%	0.1%
20%	0.58%	0%
30%	0.93%	0.13%
40%	1.26%	0.30%
50%	1.45%	0.13%

Table 6.5 "Energy saving for Case 2 when try = 1"

a1	L1 charge type	L2 charge type
10%	0.23%	0.10%
20%	0.53%	0.23%
30%	0.89%	0.53%
40%	1.07%	0.65%
50%	1.19%	0.8%

Table 6.6 "Energy saving for Case 2 when try = 100"

a1	L1 charge type	L2 charge type
10%	0.03%	0%
20%	0.25%	0.14%
30%	0.18%	0.05%
40%	0.34%	0.27%
50%	0.48%	0.14%

The Tables' results are also visualized with line plots in the two Figures below. Figure 6.2 refers to Case 1 and Figure 6.3 refers to Case 2.

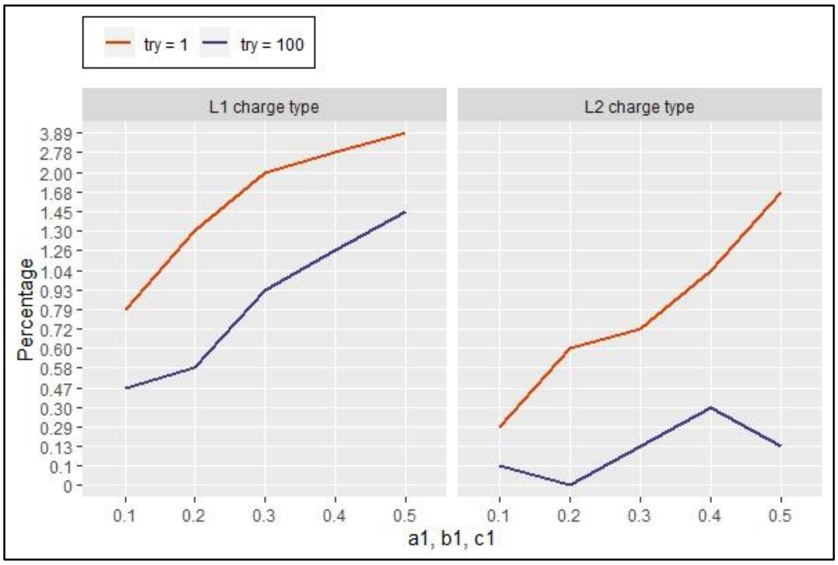

Figure 6.2 "Energy saving for Case 1"

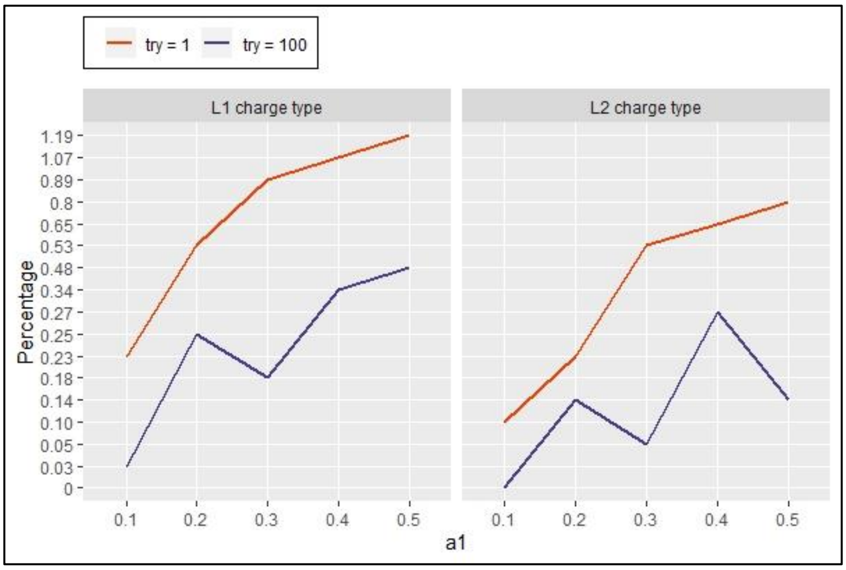

Figure 6.3 "Energy saving for Case 2"

It is clearly visible that in the diagrams of Figure 6.2 and Figure 6.3, the red line which refers to “try = 1” is always above the blue line which refers to “try = 100”. As it was mentioned in a previous paragraph, when the “while” loop is not iterated, the migrated charge may end up in a position where another charge of the same vehicle exists because the algorithm does not try to find a suitable start time for the charge where no overlaps occur. This explains why the percentages of energy saving are higher.

The red and blue lines of the L2 charge type for both cases are at a noticeably lower level than those of the L1 charge type. Vehicles using the L2 type, charge faster because the power of this type is higher. This means that the charges' durations are shorter than those of type L1, so it is less likely that an overlap will occur.

Since the sampling is random, the results may vary for different executions of the algorithm but the two points mentioned in the previous two paragraphs apply to all of its executions.

Even though the "try = 1" approach saves more energy, the strategy that will be followed in this thesis is to maintain the kWh as closer as possible to those before load shifting and not to save as much energy as possible, so the "try = 100" approach was chosen.

6.4 Studying Case 1

Before moving on to study Case 1 in more depth, it would be useful to have an overview of what the kWh per time zone on a certain day looks like after load shifting compared to the kWh per time zone before load shifting. This information is displayed in Figure 6.4 which depicts the kWh per time zone on Tuesday for the L1 charges only. On the x axis, point "0.0" refers to the kWh per time zone before load shifting. The rest of the points ("0.1" through "0.5") refer to the kWh per time zone after load shifting. For example, "0.1" refers to subcase 1 where a1, b1 and c1 are 10%.

By looking at the Figure 6.4, one can see that before load shifting, peak is the highest in energy demand zone. It is followed by shoulder 1 and off-peak which have close energy demand values and lastly by shoulder 2. In the after load shifting subcases, the kWh in peak and shoulder 1 gradually drop. In addition, peak's energy demand drops more abruptly than that of shoulder 1. Off-peak's kWh gradually increase. They surpass peak's kWh in subcase 4 (point "0.4" of the same Figure). Shoulder 2's kWh increase too but without surpassing any of the other zones' kWh. What is peculiar about this particular Figure is that by connecting the dots of the corresponding time zone, a straight line is drawn. This is caused by the uniformity of the data. Since the data was uniform before load shifting, it will still be uniform, even after load shifting.

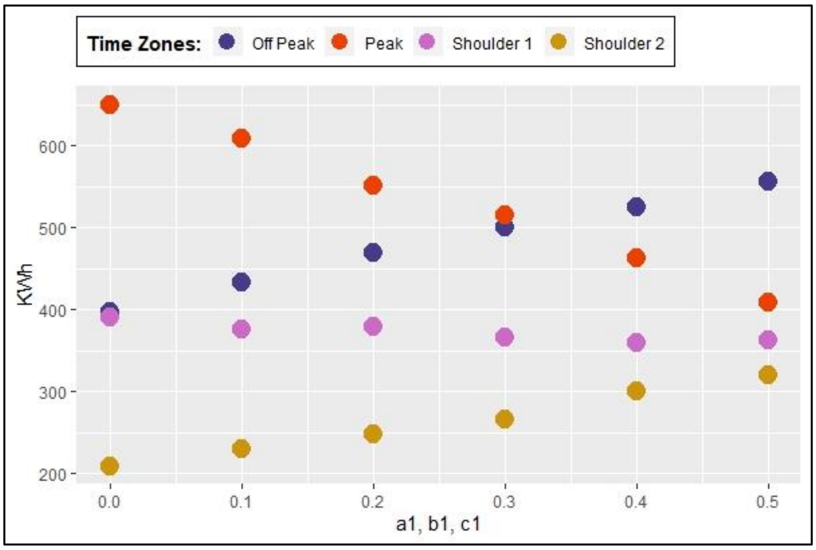

Figure 6.4 "Case 1: kWh per time zone for the L1 charges on Tuesday"

In order to continue the study and investigate the Case further, every Case 1 data frame was converted back into the L1 and L2 PEV charging tables in order to be compared to the ones before load shifting. For this process, rows whose “Noise” value was above zero were filtered out of the data frames since they are treated as energy saving. The new data frames that were created after this conversion were exported to .csv files.

As it was mentioned previously, according to Table 6.1, in the first subcase of Case 1, a1, b1 and c1 are equal to 10%. In the second subcase they are equal to 20%. In the third one they are equal to 30%. In the fourth one they are equal to 40% and in the last one they are equal to 50%.

The five Figures below display the power demand for the L1 PEV charging tables that were created after load shifting in comparison to the L1 PEV charging table before load shifting. Each of the five subcases start on Monday 4/1/2010 0:00 and end on Friday 8/1/2010 23:50. The power demand before load shifting is shown in blue color while the power demand after load shifting is shown in orange color. The areas where orange is darker are those where the before load shifting curve overlaps with the after load shifting curve.

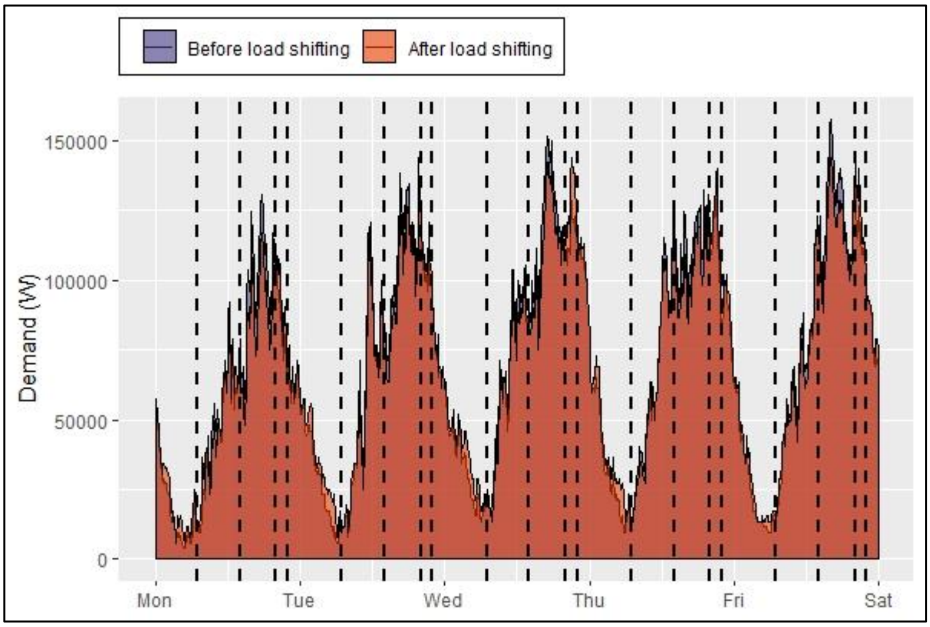

Figure 6.5 "Case 1 - Subcase 1: Weekly power demand with L1 charging"

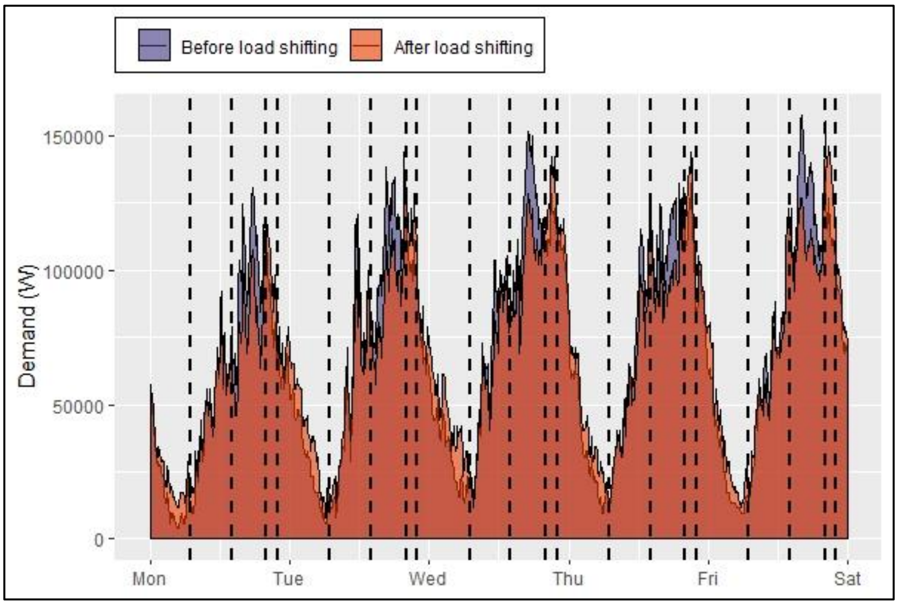

Figure 6.6 “Case 1 - Subcase 2: Weekly power demand with L1 charging”

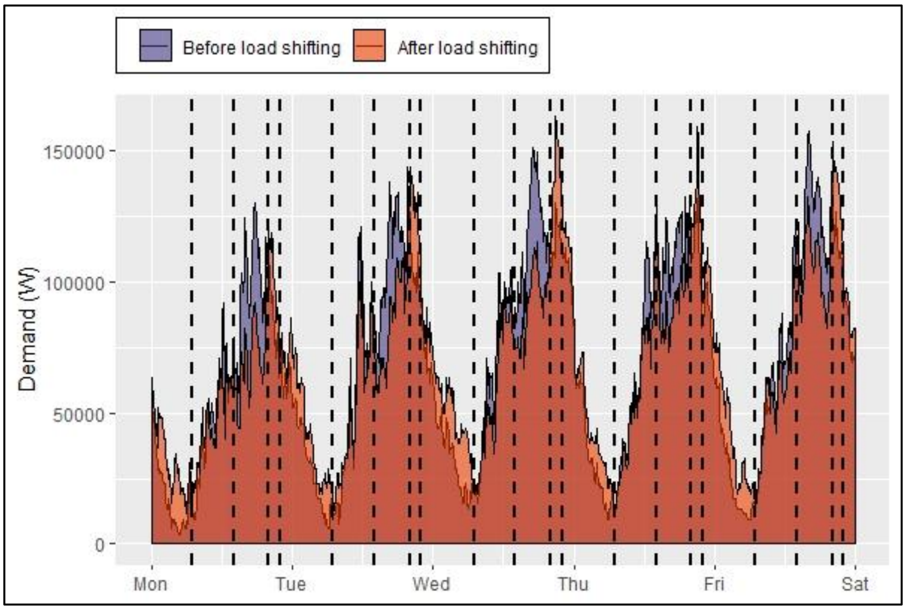

Figure 6.7 “Case 1 - Subcase 3: Weekly power demand with L1 charging”

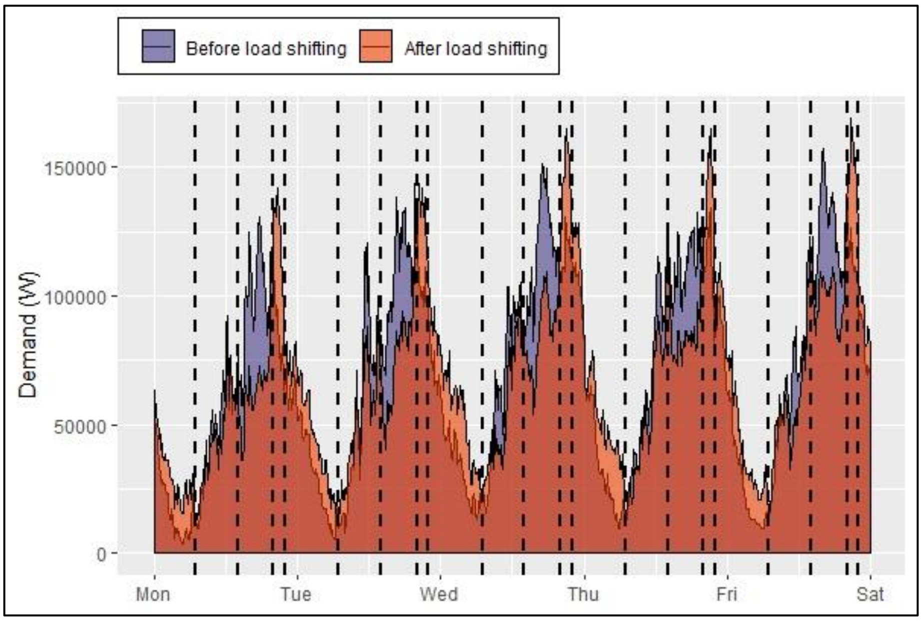

Figure 6.8 “Case 1 - Subcase 4: Weekly power demand with L1 charging”

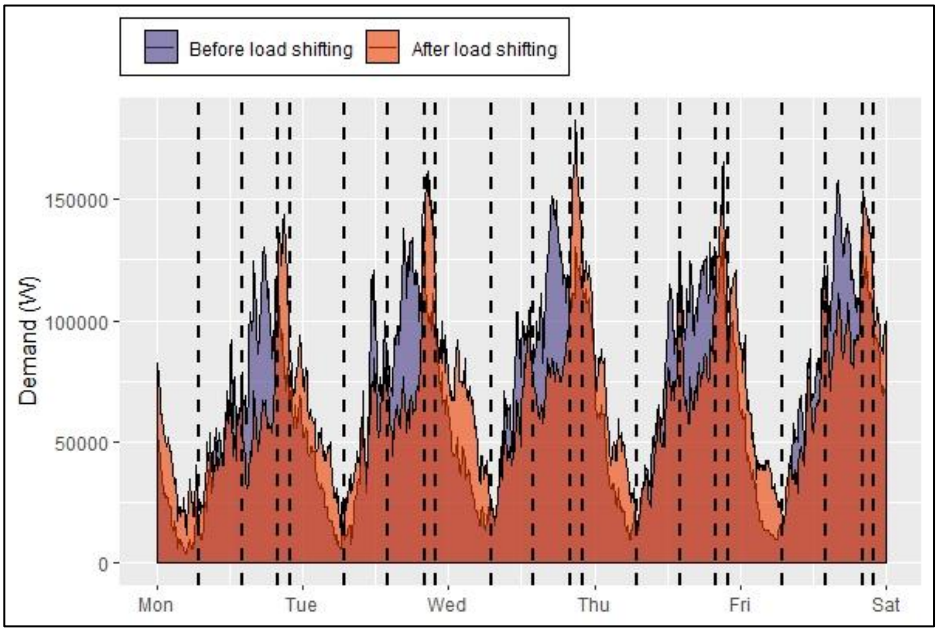

Figure 6.9 "Case 1 - Subcase 5: Weekly power demand with L1 charging"

As a1, b1 and c1 increase, more and more power demands migrate from peak zone. It is positioned on the left of the smallest area which is defined by the two vertical black dashed lines for each day in Figure 6.5, Figure 6.6, Figure 6.7, Figure 6.8 and Figure 6.9. One can also see that the power demand has also migrated from the shoulder 1 zone. It is located on the left of peak zone. The power demand seems to have migrated most to either shoulder 2 or off-peak. This last point was also confirmed in Figure 6.4.

Another noticeable detail is that the curve's height in shoulder 2 seems to be getting higher, especially in the last two subcases (Figure 6.8 and Figure 6.9). Due to its small length, this zone gets easily cramped with power demands, so it unwillingly transforms itself into the new peak zone, surpassing even the height of the original peak zone as it was before load shifting, in some cases. In a realistic scenario, this is bound to happen if the electricity providers convince people to charge their vehicles late in the afternoon. This is not a major drawback since there are other factors to be taken into consideration, such as the overall household consumption during those hours.

The five Figures below display the results for the L2 PEV charging, one for each subcase.

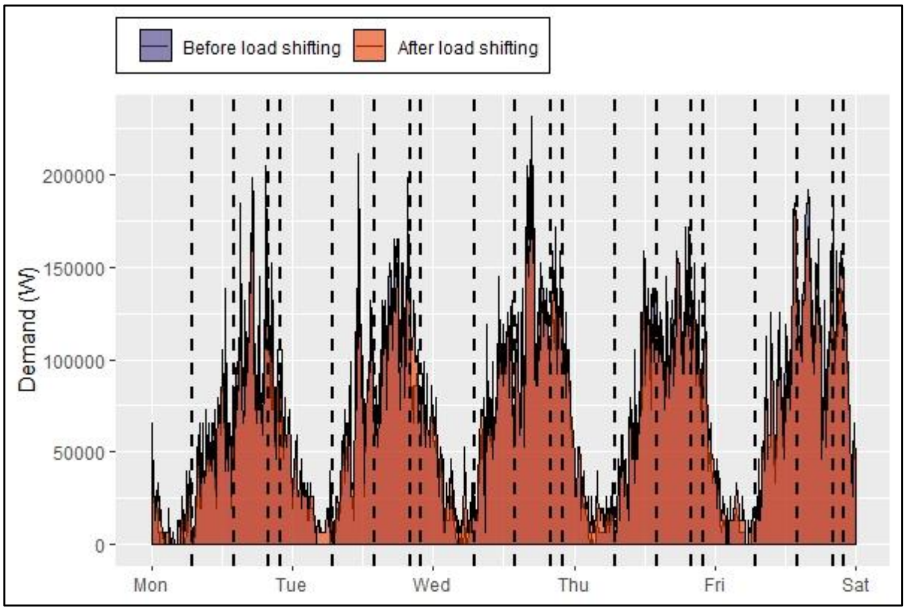

Figure 6.10 "Case 1 - Subcase 1: Weekly power demand with L2 charging"

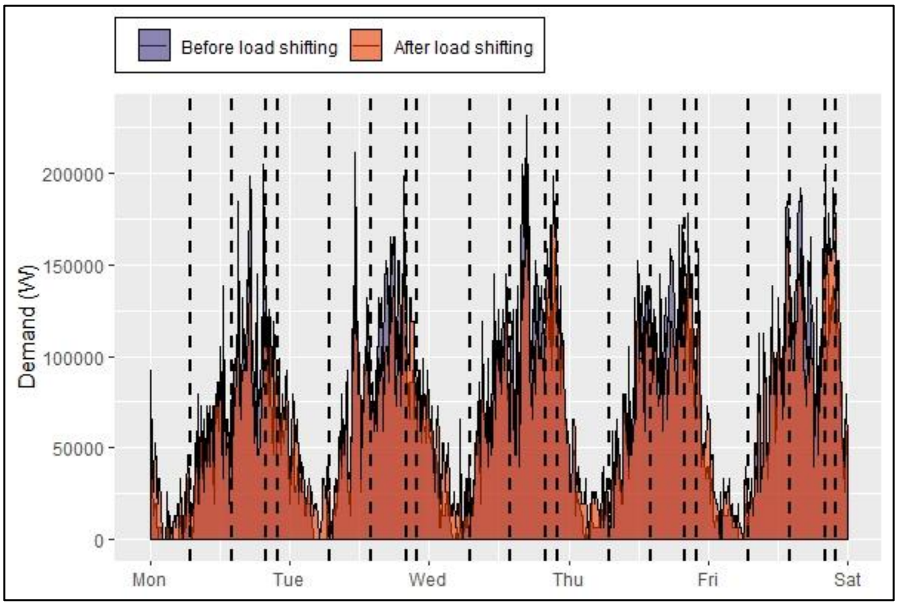

Figure 6.11 "Case 1 - Subcase 2: Weekly power demand with L2 charging"

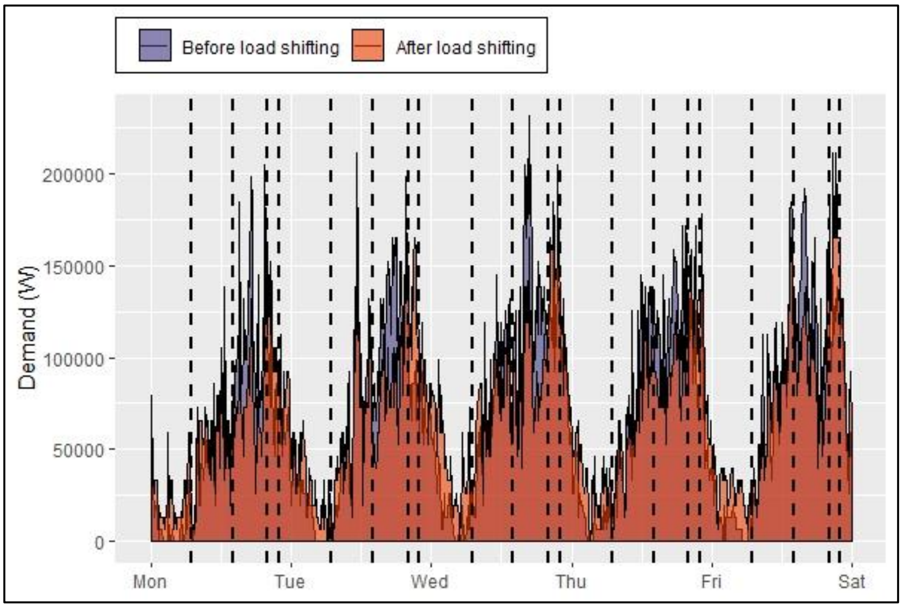

Figure 6.12 "Case 1 - Subcase 3: Weekly power demand with L2 charging"

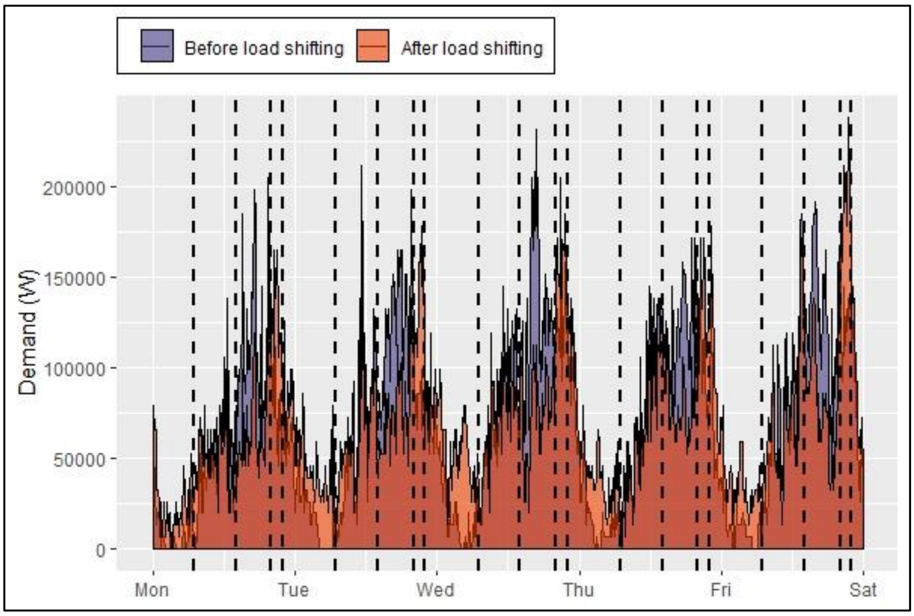

Figure 6.13 "Case 1 - Subcase 4: Weekly power demand with L2 charging"

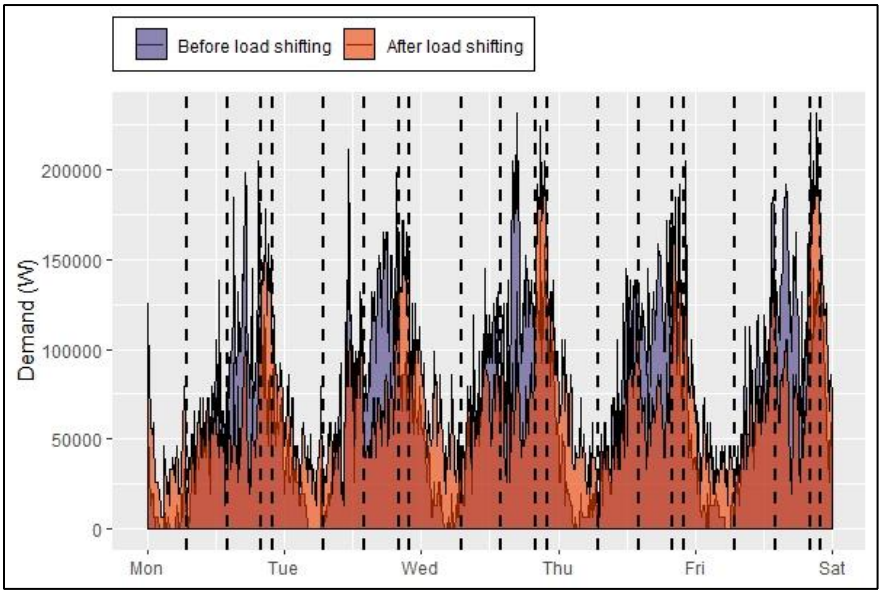

Figure 6.14 "Case 1 - Subcase 5: Weekly power demand with L2 charging"

It is noted that power demands migrate from peak zone to the off-peak zones, filling the concave areas between two successive days/peaks, similarly to the L1 PEV charging cases (Figure 6.5, Figure 6.6, Figure 6.7, Figure 6.8 and Figure 6.9). In an analogous way, power demand instances migrate from the shoulder 1 zone to either the shoulder 2 zone, or to the off-peak zone.

Even though shoulder 2’s power demand has gotten higher that the rest of the zones’ power demand, it is not high enough to surpass that of the original peak zone as it was before load shifting. To be more exact, in Figure 6.13 and in Figure 6.14, the power demand of shoulder 2 after load shifting is higher than that of peak before load shifting during Friday only. For the rest of the days as well as for the rest of the subcases (Figure 6.10, Figure 6.11 and Figure 6.12), it is always lower.

Since the previous ten Figures (Figure 6.5 through Figure 6.14) represent information about a week excluding the weekend, minor details cannot be easily identified. New line plots which focus on daily power demand were crated. They start from the off-peak zone of a day and ending on the off-peak zone of the next day. The initiation point on the x axis is Monday 4/1/2010 22:00 (first hour and minute of the off-peak zone) and its ending point is Wednesday 6/1/2010 6:50 (last hour and minute of next day’s off-peak zone). This is a time span of approximately 33 hours. Vertical black dashed lines are used to distinguish each time zone. The blue line refers to the power demand over time before load shifting. The red line refers to the power demand over time after load shifting.

The three Figures below display the results of the first, third and fifth subcases respectively, for the L1 PEV charging.

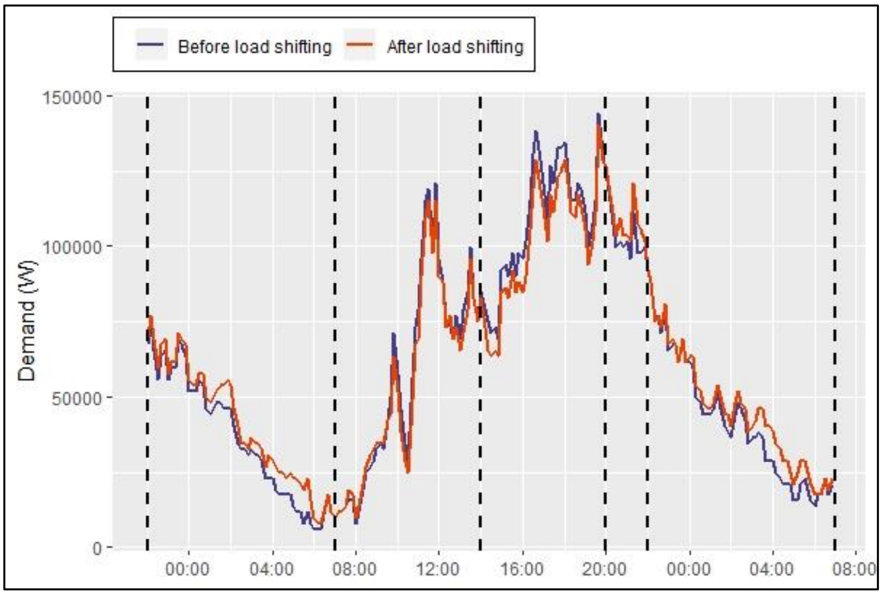

Figure 6.15 "Case 1 - Subcase 1: Daily power demand with L1 charging"

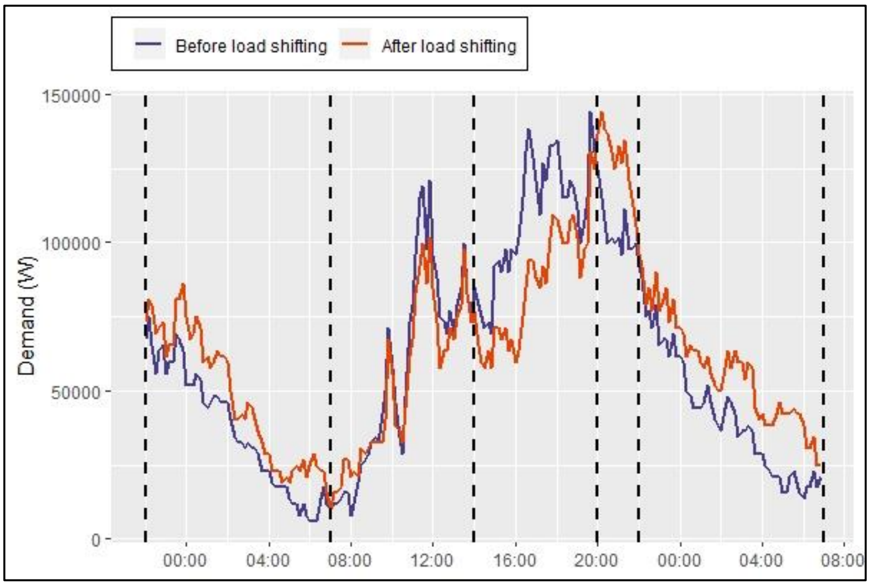

Figure 6.16 "Case 1 - Subcase 3: Daily power demand with L1 charging"

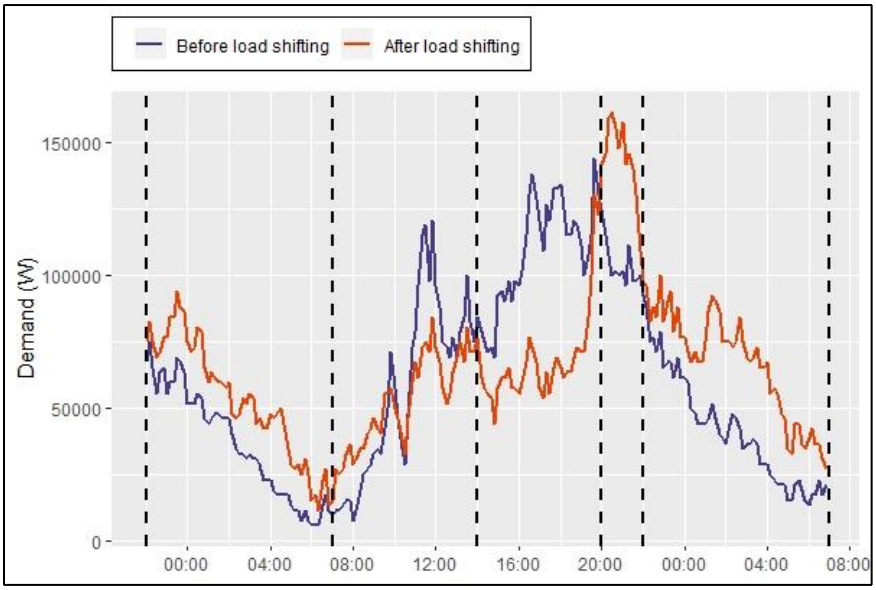

Figure 6.17 "Case 1 - Subcase 5: Daily power demand with L1 charging"

In Figure 6.15, most power demands have migrated from the pea” zone to the off-peak zone. No major improvements are spotted in shoulder 1 and Shoulder 2.

In Figure 6.16, more power demands have migrated to off-peak and also to shoulder 2. It is clearly visible that shoulder 2 is now the zone which contains the highest power demand. Its highest top is at the same height as the highest top of the peak zone before load shifting.

In Figure 6.17, peak zone’s power demand has decreased drastically. Another decrease in demand is also visible in shoulder 1, as even more charges have migrated to off-peak and shoulder 2. Moreover, shoulder 2’s, highest top has now surpassed that of peak zone as it was before load shifting.

The three Figures below show the results of the L2 PEV charging for subcases one, three and five.

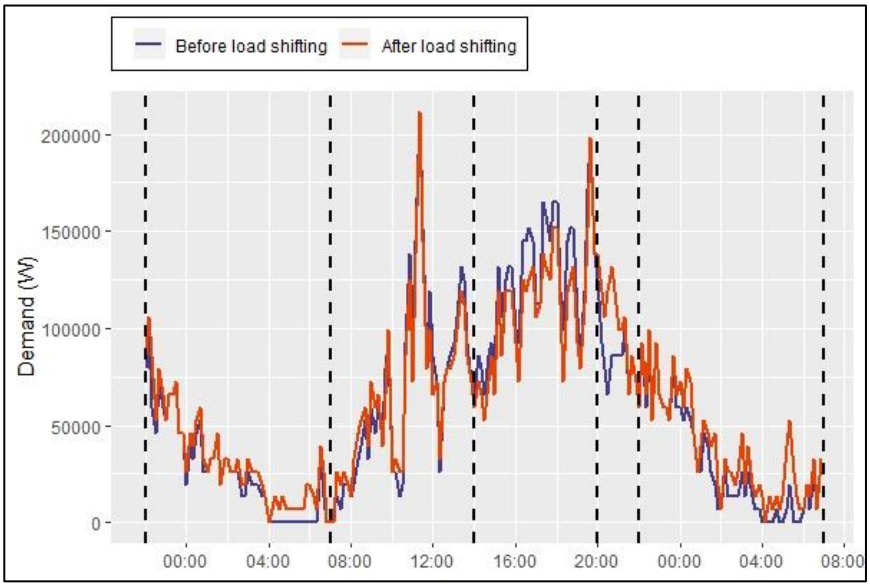

Figure 6.18 "Case 1 - Subcase 1: Daily power demand with L2 charging"

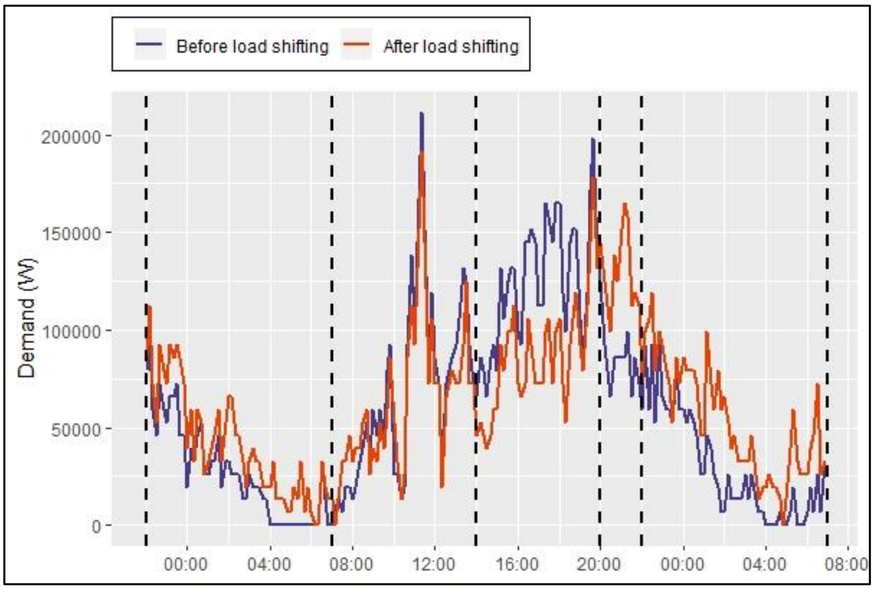

Figure 6.19 "Case 1 - Subcase 3: Daily power demand with L2 charging"

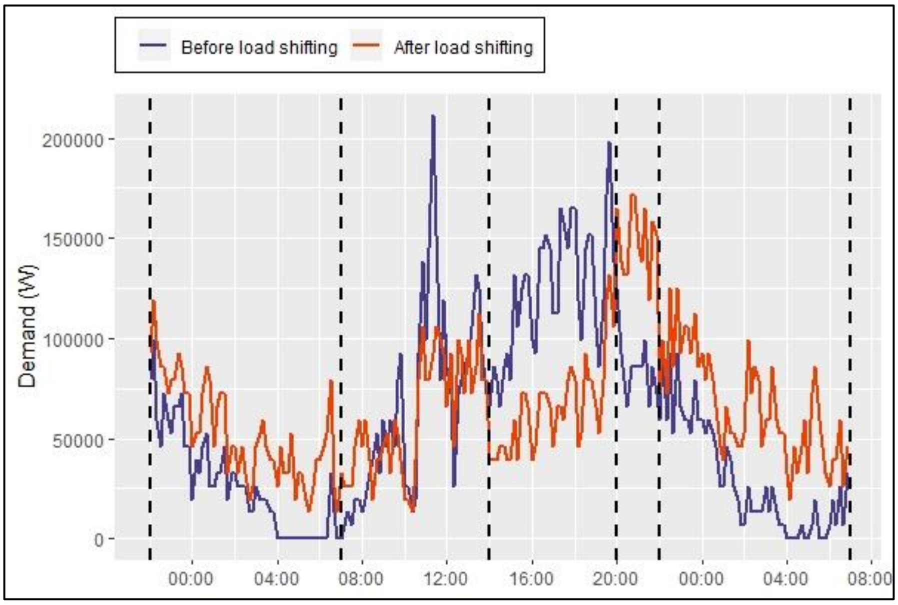

Figure 6.20 "Case 1 - Subcase 5: Daily power demand with L2 charging"

In Figure 6.18, most power demands have migrated from peak zone to the off-peak zone and there aren't any major improvements in shoulder 1, similarly to subcase 1 of the L1 PEV charging case (Figure 6.15). Apart from that, shoulder 2's power demand is higher than it was before load shifting. Another noticeable detail is that shoulder 1 is the zone with the highest power demand values.

In Figure 6.19, peak zone's power demand is lower as more charges have migrated to off-peak and to shoulder 2. Shoulder 1's power demand has dropped as well, but this zone together with peak zone still maintain the highest power demand.

In Figure 6.20, peak's and shoulder 1's power demand has significantly dropped. The exact opposite has happened to shoulder 2 which is now the zone with the highest power demand values. However, those values do not surpass the highest power demand values of neither peak nor shoulder 1 as they were before load shifting.

6.5 Studying Case 2

Case 2 will also begin with an overview of the kWh per time zone on Tuesday for the L1 PEV charging only, before and after load shifting. This is displayed in Figure 6.21. On the x axis, point "0.0" refers to the kWh before load shifting. The rest of the points ("0.1" through "0.5") refer to the kWh after load shifting. For example, "0.1" refers to subcase 1 where a1 is 10%.

By looking at the Figure 6.21, one can see that peak zone's kWh gradually drop as a1 decreases. The kWh of the other zones gradually increase. Peak's kWh drop below both shoulder 1 and off-peak in subcase 4. Similarly to Figure 6.4, if the points of the corresponding time zones are connected with imaginary lines, the resulting lines turn out to be almost straight because the data was uniform before load shifting, so it will still be uniform after load shifting.

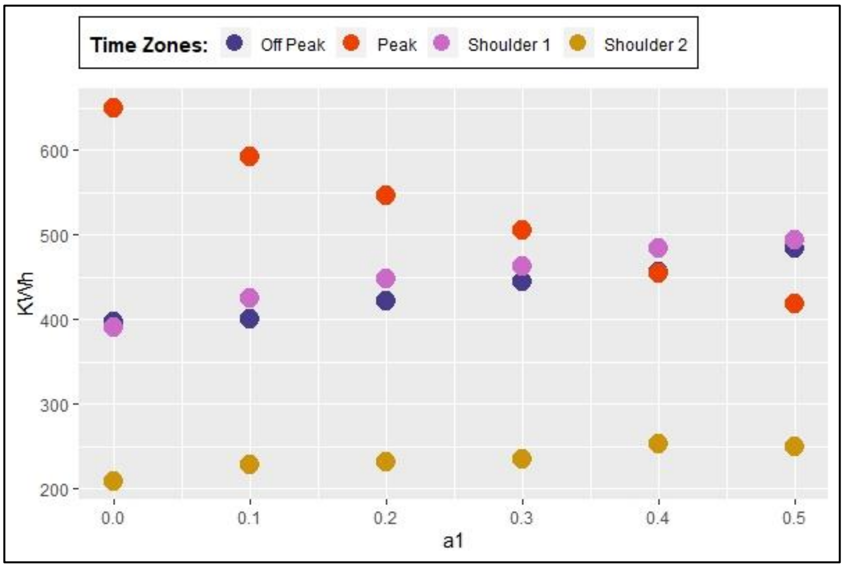

Figure 6.21 "Case 2: kWh per time zone for the L1 charges on Tuesday"

All the data frames that were created after load shifting were now ready to be converted back into the L1 and L2 PEV charging tables in order to be compared to the ones before load shifting. Any rows whose "Noise" value is above zero are filtered out of the data frame since they are treated as energy saving. The new data frames that were created from this process were exported as .csv files.

As it was mentioned previously, five subcases were created for Case 2. In the first subcase, a1 is equal to 10%. In the second subcase it is equal to 20%. In the third one it is equal to 30%. In the fourth one it is equal to 40% and in the last one it is equal to 50% (see Table 6.2).

The five Figures below display the power demand for the L1 PEV charging tables that were created after load shifting in comparison to the L1 PEV charging table before load shifting. Each of the five subcases start from Monday 4/1/2010 0:00 and end on Friday 8/1/2010 23:50. The power demand before load shifting is shown in blue color. The power demand after load shifting is shown in orange color. The areas where the orange is darker are those where the curve before load shifting overlaps with the curve after load shifting.

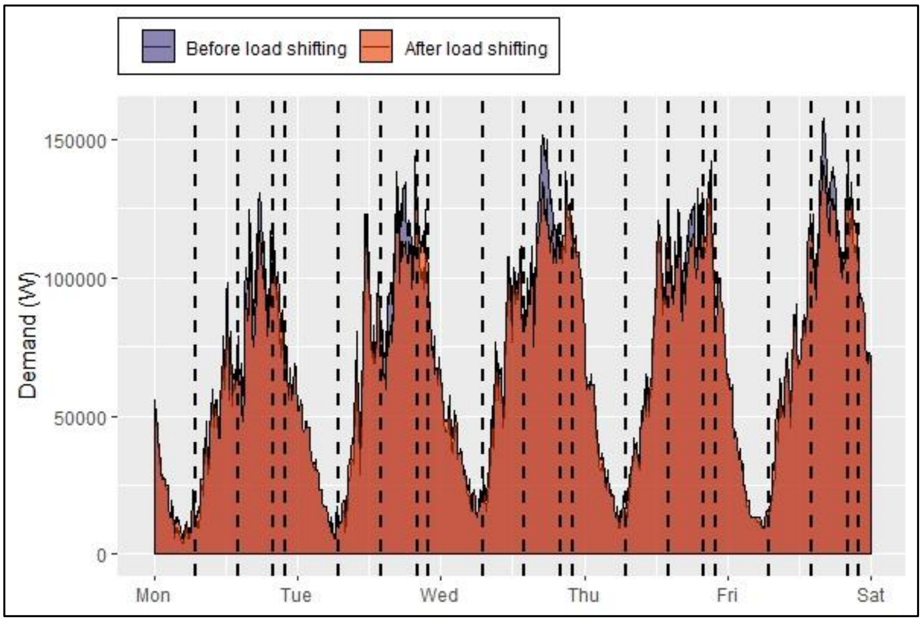

Figure 6.22 "Case 2 - Subcase 1: Weekly power demand with L1 charging"

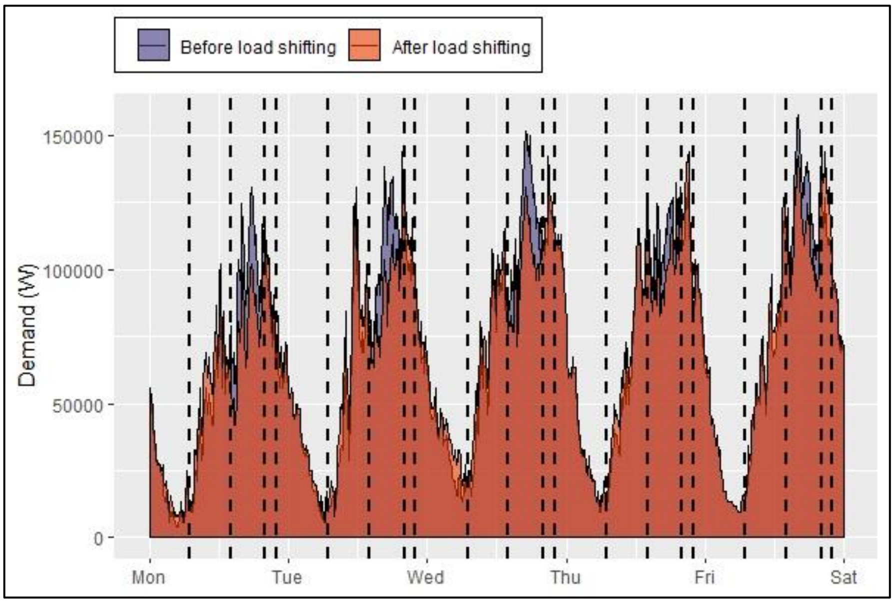

Figure 6.23 "Case 2 - Subcase 2: Weekly power demand with L1 charging"

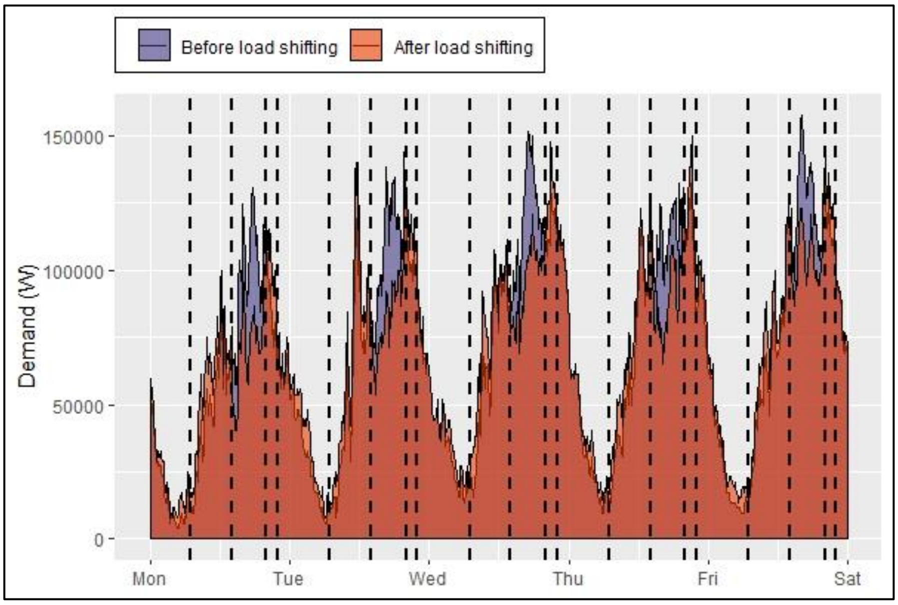

Figure 6.24 "Case 2 - Subcase 3: Weekly power demand with L1 charging"

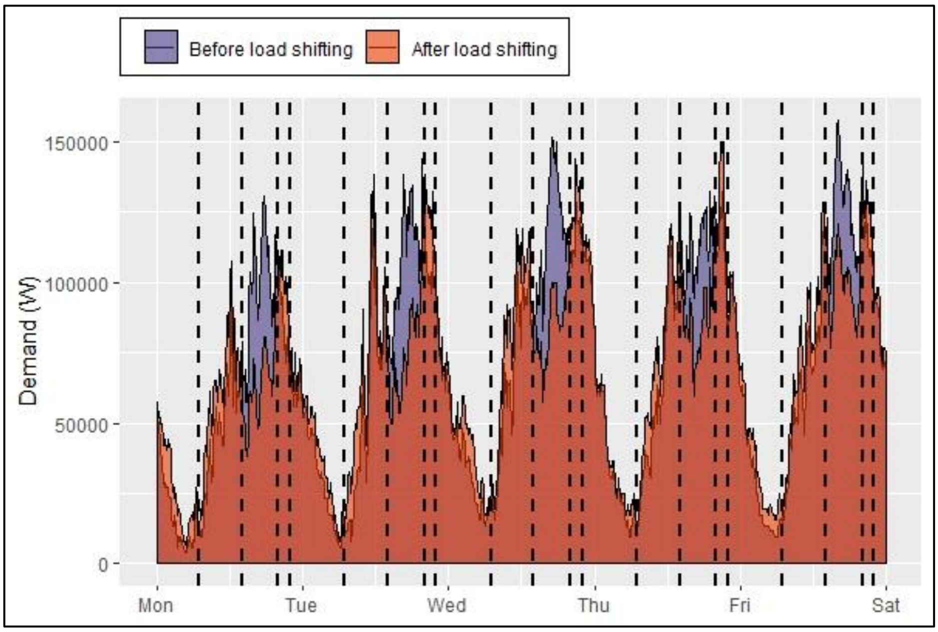

Figure 6.25 "Case 2 - Subcase 4: Weekly power demand with L1 charging"

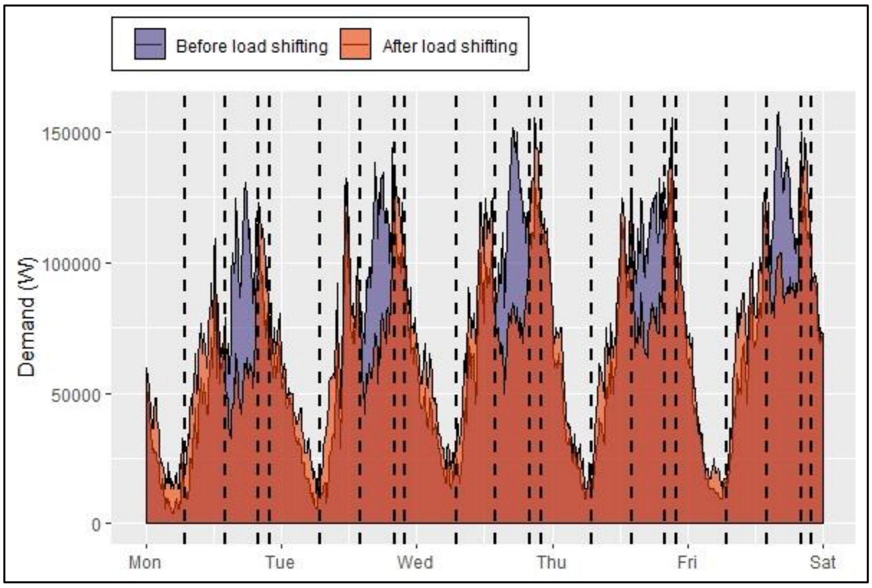

Figure 6.26 "Case 2 - Subcase 5: Weekly power demand with L1 charging"

In Figure 6.22 there are not any major improvements. The same circumstances prevail in Figure 6.23, more or less. The tops of peak zone have shrunk a bit more since 20% of the energy demand from this zone has migrated to the other zones.

In Figure 6.24, Figure 6.25 and Figure 6.26, the power demand of peak zone has dropped to a satisfactory level. The concave areas between two successive days/peaks have been filled with the power demands that have migrated from peak zone. Since the power demands are migrating only from peak zone, two tops were created. They hold the highest power demands. The first top is located in shoulder 1 and the other in shoulder 2. Shoulder 2 has gotten cramped with power demands again due to its short length. However, unlike Case 1, shoulder 2's power demand does not exceed the power demand peak zone as it was before load shifting.

The five Figures below display the results for the L2 PEV charging, one for each subcase.

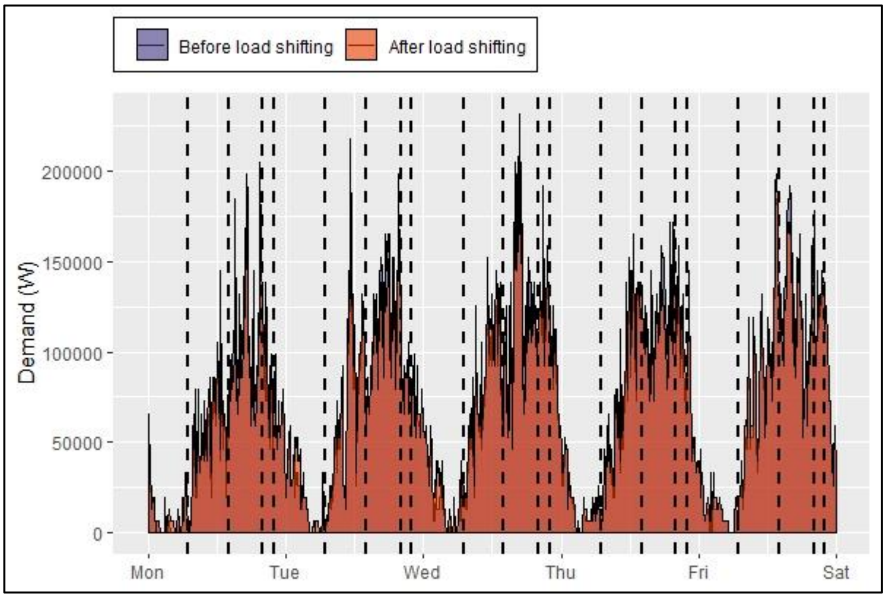

Figure 6.27 "Case 2 - Subcase 1: Weekly power demand with L2 charging"

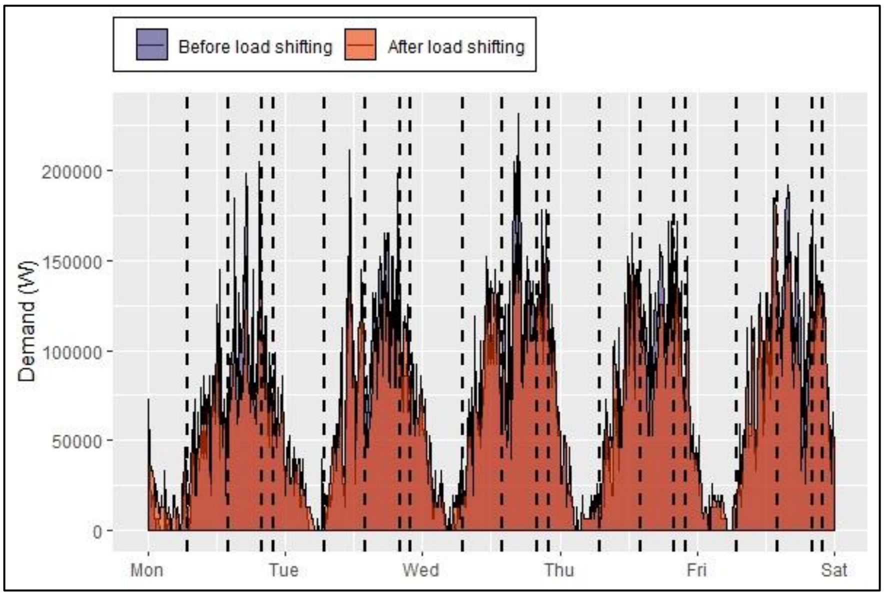

Figure 6.28 "Case 2 - Subcase 2: Weekly power demand with L2 charging"

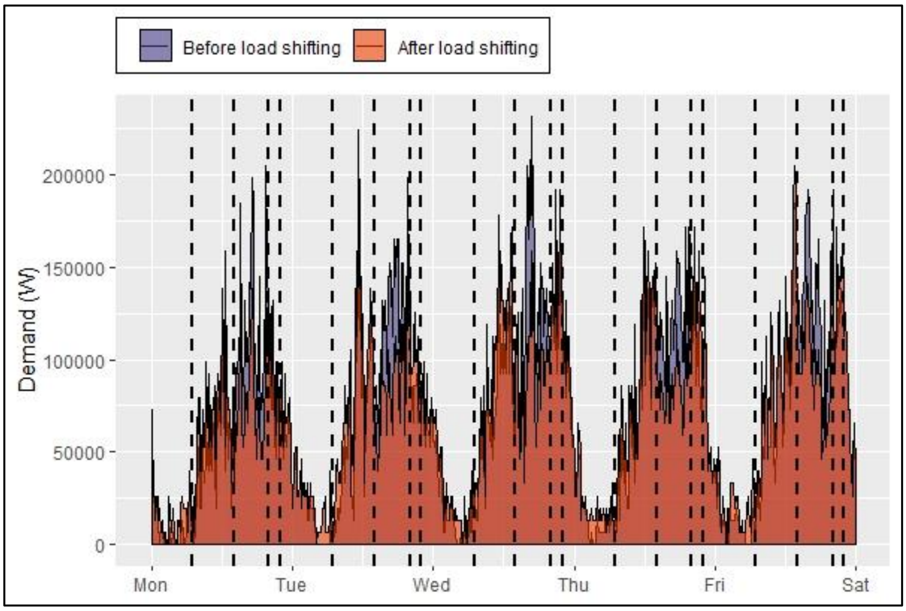

Figure 6.29 "Case 2 - Subcase 3: Weekly power demand with L2 charging"

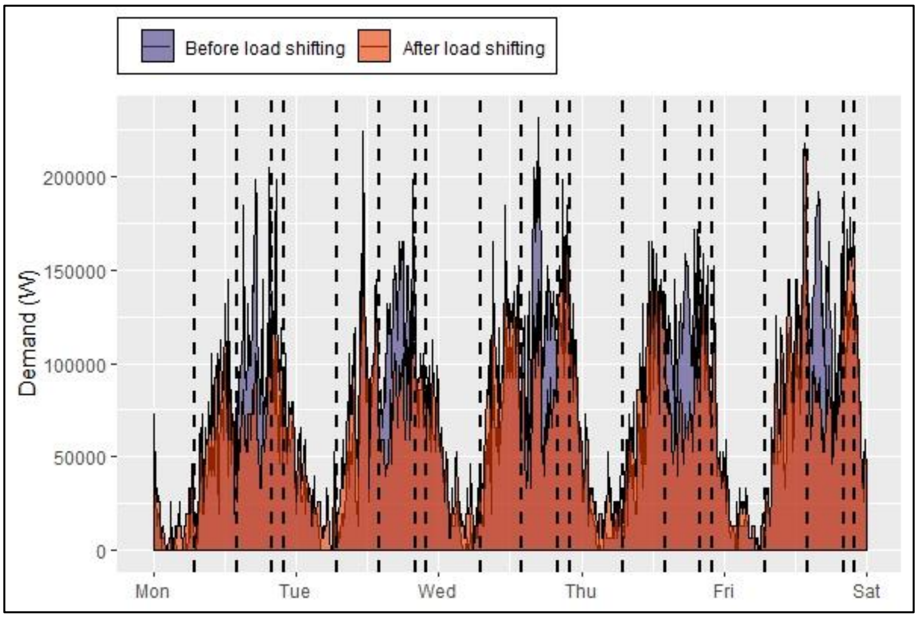

Figure 6.30 "Case 2 - Subcase 4: Weekly power demand with L2 charging"

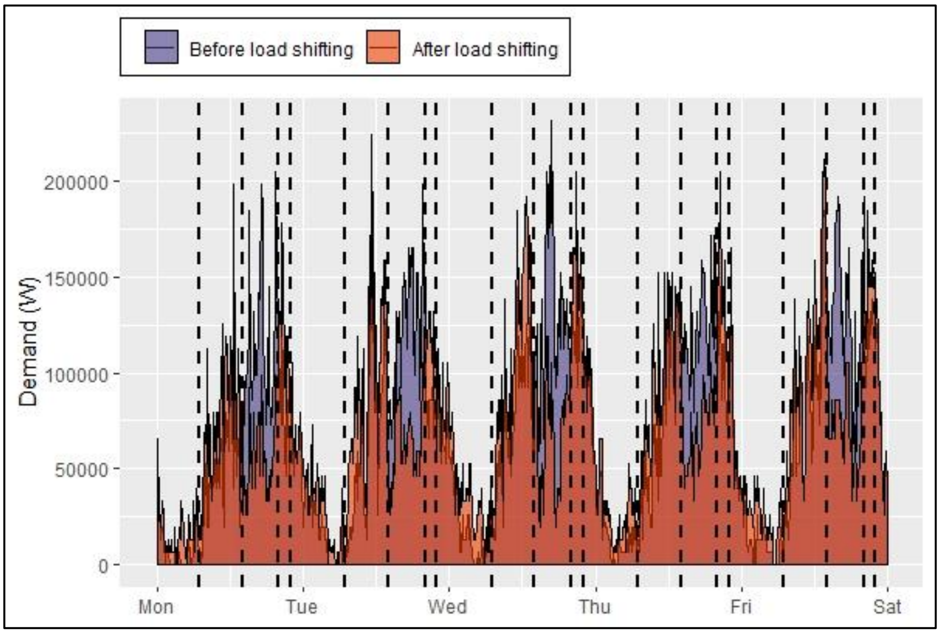

Figure 6.31 "Case 2 - Subcase 5: Weekly power demand with L2 charging"

The same phenomenon with a top forming on the left and right of peak zone is observed again (Figure 6.29, Figure 6.30 and Figure 6.31). It is hard to distinguish any clear changes occurring in Figure 6.27 and in Figure 6.28.

Similarly to Case 1, new line plots which focus on daily power demand were crated. They initiation point on the x axis is Monday 4/1/2010 22:00 and its ending point is Wednesday 6/1/2010 6:50. The blue line refers to the power demand over time before load shifting. The red line refers to the power demand over time after load shifting.

The three Figures below display the results of the first, third and fifth subcases respectively, for the L1 PEV charging.

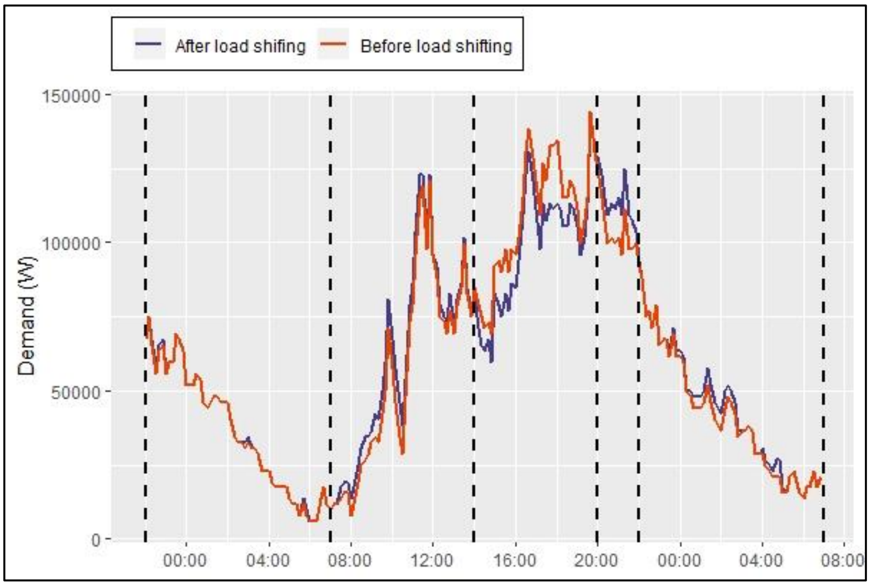

Figure 6.32 "Case 2 - Subcase 1: Daily power demand with L1 charging"

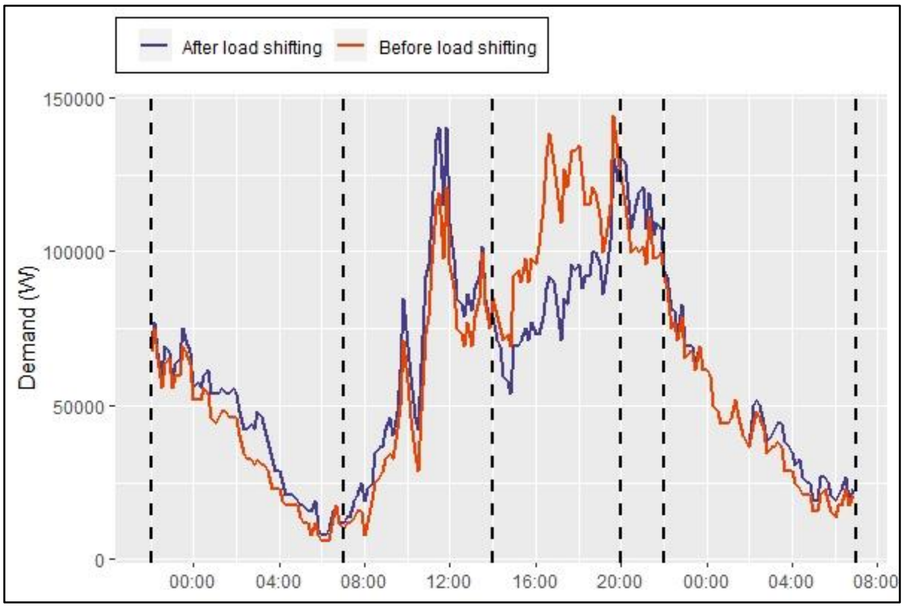

Figure 6.33 "Case 2 - Subcase 3: Daily power demand with L1 charging"

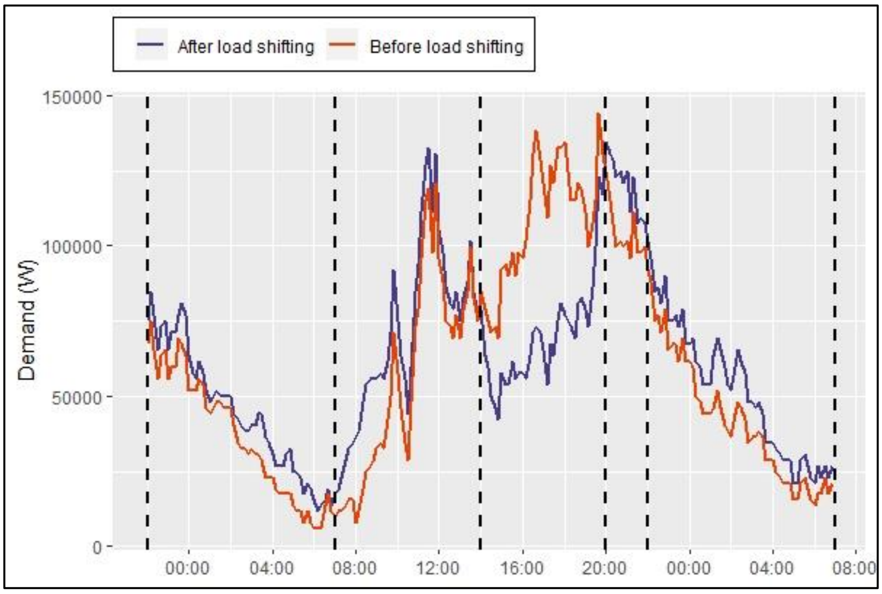

Figure 6.34 "Case 2 - Subcase 5: Daily power demand with L1 charging"

In Figure 6.32, one can see that most power demands have migrated to the off-peak zone. Peak zone's power demand has dropped a bit while shoulder 1's and shoulder 2's power demand seems to remain unchanged.

In Figure 6.33, peak zone's power demand has dropped more because 30% of it has migrated to the other zones. Shoulder 1 is now the zone which holds the highest power demand values, followed by shoulder 2. A concave area has been formed between those two zones.

In Figure 6.34, the height of the tops of shoulder 1 and shoulder 2 has increased. The height of the top of shoulder 2 has surpassed that of shoulder 1, but not that of peak zone before load shifting. Moreover, the demand of peak zone has decreased even more. The concave area between the tops of shoulder 1 and shoulder 2 has gotten deeper.

The three Figures below show the results of the L2 charge type for subcases one, three and five.

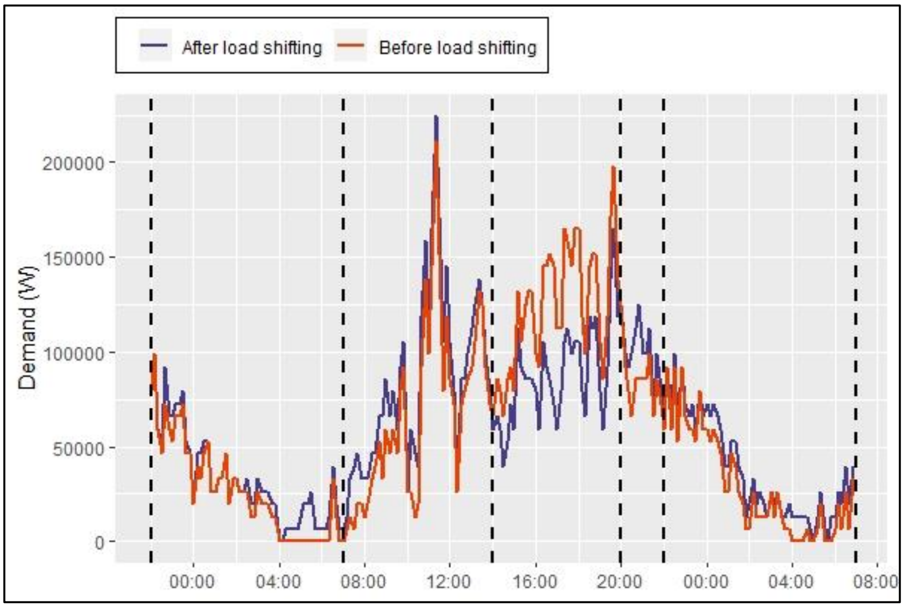

Figure 6.35 "Case 2 - Subcase 1: Daily power demand with L2 charging"

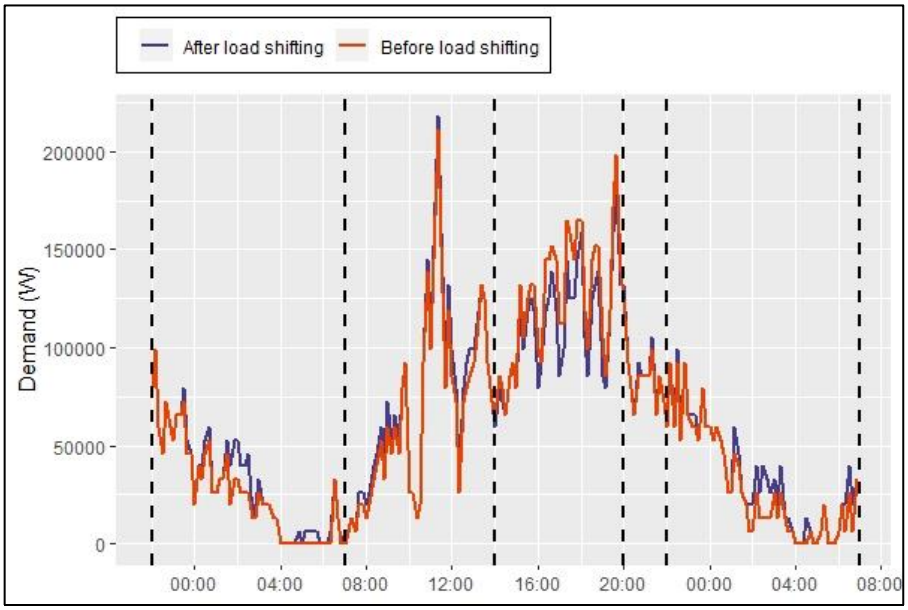

Figure 6.36 "Case 2 - Subcase 3: Daily power demand with L2 charging"

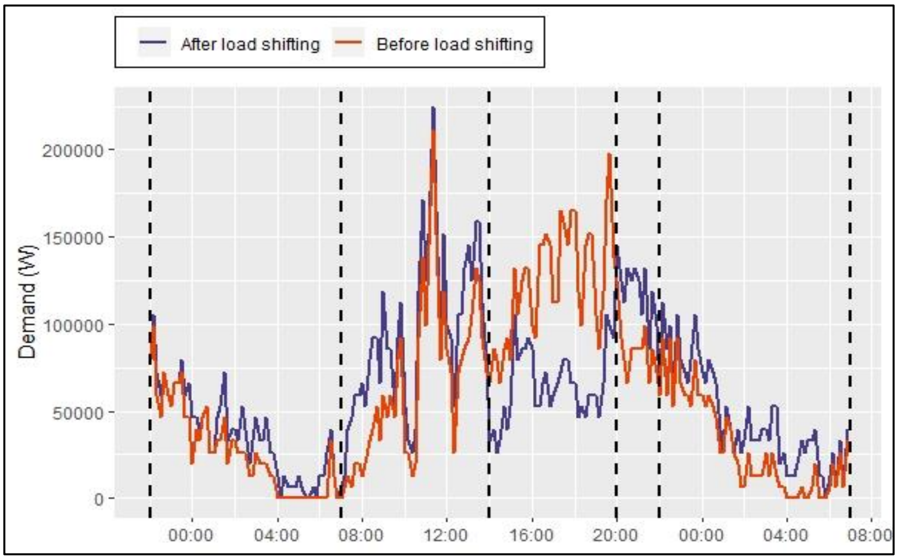

Figure 6.37 "Case 2 - Subcase 5: Daily power demand with L2 charging"

Similarly to Figure 6.32, there is not any visible power demand change in the shoulder 1 and shoulder 2 zones. The power demands from peak zone seem to have migrated mostly to the off-peak zone.

In Figure 6.36, peak zone's power demand has dropped. The power demand of the other zones has increased, forming a concave area between shoulder 1 and shoulder 2. The zone which holds the maximum demand value is shoulder 1.

In Figure 6.37, shoulder 1 still holds the maximum demand value. Peak zone's power demand has dropped significantly and the concave area has become deeper.

Conclusion

In this chapter, two different load shifting cases, each one divided into five subcases, were investigated. In the first case, various kWh percentages of peak, shoulder 1 and shoulder 2 were migrated. Although peak's and shoulder 1's power demand decreases, shoulder 2's power demand increases because it is the smallest time zone. In the second case, only kWh percentages of peak participated in the load shifting process. In this case, peak's demand drops and two new tops form; one in shoulder 1 and another in shoulder 2.

Due to energy demand overlaps of the same vehicle during load shifting, which was unavoidable most of the time, an energy saving strategy was explored. It was also found that since the data was uniform before load shifting, it will still be uniform even after load shifting.

Chapter 7: Conclusions and Future Work

This thesis considers load shifting strategies for achieving energy savings for owners of Plug-in Electric Vehicles (PEVs). An overview of the relation between the electric vehicles and the smart grid is presented, as well as the main challenges related to the smart grid's overloading during peak hours caused by electric vehicle charging. Moreover, various demand side management strategies are considered in the form of candidate solutions.

Previous related research is indicative of the large impact of uncoordinated PEV charging has to the grid. Proposed solutions include load shifting, peak shaving, coordinated charging with a fuzzy logic controller and delayed charging.

A number of tools for data analytics have been considered, leading to the adoption of R/RStudio as the preferred programming environment for the thesis' research, because it is an easy to learn language and its large collection of libraries makes data manipulation and graphical representation simple and effective.

The contribution of the thesis began with exploratory data analysis of the datasets used in [28]. The data were produced by a simulation model. Upon first observation there appeared to be a noticeable degree of homogeneity in them. In the original version the data comprised of power consumption measurements sampled every 10mins. We proceeded and organized them into the "Time Zones" structure [47], which examines the data from another perspective by categorizing the energy demand into time zones.

A random calendar week was selected for experimenting with and two different load shifting scenarios involving both L1 and L2 PEV charging.

A key issue was that during the load shifting process, a charge could migrate to a time zone where another charge of the same vehicle already exists. The problem was addressed with two energy saving approaches. In the first one, only one effort was made to place the charge in a position where no overlaps occur within the destination time zone. In the second approach, up to 100 tries were performed to initiate a charging cycle where no overlaps occur within the destination time zone. The second approach was chosen as it maintains the energy demand after load shifting as close as possible to the energy demand before load shifting.

After applying all necessary parameters into the simulated model and visualizing the results, it was concluded that load shifting with energy savings can improve the energy demand from the smart grid. This is due to a pre-specified percentage of peak demand instances being transferred to the off-peak demand zones.

The strategy did not have much of an effect for low energy demand percentages (in subcases 1 and 2). Changes in the demand curves started to show up for higher percentage values (subcases 3, 4 and 5).

In Case 1, as the percentages increase, more and more power demands are being built in the off-peak zones, but shoulder 2, which is the smallest in duration zone, gets built up with too many power demands and even surpasses the power demand peak zone had before load shifting in the L1 PEV charging. However, something like that does not necessarily put the smart grid at risk since shoulder 2 is still a low demand zone considering the fact that other household appliances might not be powered on during that time.

In Case 2, the energy demand migrates from peak zone only and thus a valley is formed in its place. As the percentage from peak zone increases, this valley becomes deeper and the tops that form to the left and the right of it grow taller but they never surpass the power demand peak zone had before load shifting.

The study showed that both cases succeed in decreasing peak zone’s power demand and shifting the load to the off-peak zones. Both cases provide their own benefits. It is up to the energy supplier to choose which of the two load shifting cases and energy saving approaches to implement in order to reduce the load, according to the smart grid’s needs.

Plans for future work could be considered as anything that could not be explored in the context of the current thesis. This includes allowing energy demands that expand in more than one zone to be migrated but also in the correct place which is the one being right after the other. Moreover, cases that represent the other energy saving approach that was not explored in the thesis could be created and investigated. It would also be interesting to model data for the case of Greece in a hypothetical scenario where the country switches to smart grid systems and replaces existing vehicles with electric ones and then apply the proposed load shifting method and see how it responds, so that the country is better prepared for the changes that are bound to come.

Appendix

The following algorithm creates a part of the Time Zones structure.

```
Algorithm 1: Creates a part of the Time Zones structure
Input: .csv file name, file; watt number, watt; charge type,
type;
Output: TZ_LX;
function convert.to.time.zones(file, watt, type)
  set df as the contents of file;
  initialize all necessary vectors for the new structure;
  set lower as the starting point of the loop;
  set upper as the ending point of the loop;
  z1 = ["7:00", "14:00", "20:00", "22:00"];
  z2 = ["7:00", "22:00"];
  duration = 0;
  for all df's columns except the first one do
    for upper to lower do
      if df's element == to watt and duration == 0 then
        update all necessary vectors;
        duration = 1;
        if (next df's element time is found in z1 and df's
element belongs to a working day) or (next df's element is
found in z2 and df's element belongs to a weekend day) then
          split the charge by updating the necessary vectors;
          duration = 0;
        end if
      else if df's element == watt and duration != 0 then
        duration = duration + 1;
        if (next df's element time is found in z1 and df's
element belongs to a working day) or (next df's element is
found in z2 and df's element belongs to a weekend day) then
          split the charge by updating the necessary vectors;
          duration = 0;
        end if
      else if df's element == 0 and duration != 0 then
        end the charge by updating the necessary vectors;
        duration = 0;
      end if
    end for
  end for
  set any other necessary vector to their appropriate values;
  set TZ_LX as a data frame containing all necessary vectors;
  return TZ_LX;
end function
```

The function takes three mandatory arguments; "file", which refers to the name of a .csv file about the vehicles' charging profiles where the first column needs to have the name "Time" and contain sequential dates and times divided by 10 minutes. The rest of the columns need to have a unique household-vehicle id as their names. When a vehicle is not charging, the appropriate cell needs to have been filled with a zero. Otherwise, it needs to have been filled with either 1920 in the case of the L1 PEV charging or 6600 in the case of the L2 PEV charging. “watt” refers to the power demand during charging (either 1920 or 6600) and “type”, which refers to the charge type (either “L1” or “L2”). The contents of “file” are then stored in the data frame named “df”.

The other worthy of mentioning vectors are the following; “lower” and “upper” are two one-element vectors which determine the lower and upper limits for the rows in the “for” loop statement. “duration”, which is initialized to zero, refers to the charge’s duration. “z1” holds every time zone’s start time on working days and “z2” holds the ones on weekend days. “z1” and “z2” are useful in determining whether or not a charge should split.

Once the loop is entered, “df” is scanned until an element equal to “watt” is found, provided that “duration” is zero. Then, the necessary vectors are updated accordingly. A possible split occurrence also has to be checked. If the time of the next element is equal to one of the elements of either “z1” (in case of working days) or “z2” (in case of weekend days), then the charge has to split and the other necessary vectors are updated accordingly with “duration” being set back to zero. If no split occurs, then “df” continues to be scanned and “duration” keeps increasing with each iteration until either a split occurs or the “df”’s element is zero, meaning that the vehicle has stopped charging for the time being. This procedure continues until all elements in “df” are checked.

Once the loop finishes, the function turns the necessary vectors into a data frame and returns it.

The algorithm below performs load shifting with energy saving.

```
Algorithm 2: Performs load shifting with energy saving
Input: the Time Zones structure, df; the index of the charge
that will migrate, ind; the destination zone, nZone;
Output: a part of the Time Zones structure after load
shifting, df;
function shifting(df, ind, nZone)
  if the charge does not fit to the destination zone then
    place the charge from the beginning till the end of the zone
and place its remainder to start from the beginning of the next
zone;
    note any overlapped kWh;
  else
    try = 0;
    flag = true;
```

```
    while try < 100 and flag == true do
      randomly generate a new position such that the charge will
always fit in the destination zone;
      place the charge in the generated position;
      try = try + 1;
      note any KWh overlaps with charges of the same vehicle;
      if no such overlaps occur then
        flag = false;
      return the updated data frame;
  end if
end function
```

The function takes three mandatory arguments; “df” which is the Time Zones structure, “ind”, which refers to the index of the charge that will migrate and “nZone” which contains the name of the charge’s destination zone.

The other notable variables are “try” which counts the number of tries in the “while” loop and “flag” which is a Boolean that becomes “true” when the algorithm has succeeded in placing the charge in an empty space where no overlaps of the same vehicle occur. If “try” reaches 100 and “flag” is still “true” (meaning that the “while” loop was unable to find a position without overlaps), the charge remains in its last generated position of its destination zone and accompanied by the overlapped kWh that had been calculated.

The algorithm below can convert any Time Zones structure that was created either before or after load shifting, back into the L1 and L2 PEV charging tables.

```
Algorithm 3: Converts Time Zones structures back to the L1/L2
tables
Input: a data frame with the first column named “Time”,
containing dates and times and the rest of the columns have
the unique household-vehicle id in alphabetical order and have
been pre-filled with zeros, x; the Time Zones structure, df;
watt number, watt;
Output: x;
function convert.to.lx(x, df, watt)
  counter = 1;
  k = 1;
  for all x’s columns except the first one do
    for all x’s rows do
      if x’s time == df’s time in position k then
        fill x’s element with watt;
        counter = counter + 1;
        if df’s charge duration in position k == 1 then
          counter = 1;
          k = k + 1;
```

```
        end if
      end if
      else if counter > 1 then
        if counter < df's charge duration in position k then
          fill x's element with watt;
          counter = counter + 1;
        end if
        else
          if counter == df's charge duration in position k
then
            fill x's element with watt;
            counter = 1;
            k = k + 1;
          end if
          else
            counter = 1;
            k = k + 1;
          end if
        end if
      end if
    end for
  end for
  return x;
end function
```

Function "convert.to.lx" takes three mandatory arguments. "x" is a pre-constructed data frame whose first column contains consecutive dates and times divided by ten minutes in the correct order while the rest of the columns (348 in number) need to have been filled with zeros. "df" is a Time Zones structure before or after load shifting and "watt" is the power demand (1920 for the L1 PEV charging or 6600 for the L2 PEV charging).

The other notable variables include "counter" which counts the number of "x's" cells that have to be filled with "watt's" value. This assignment ends when "counter" is equal to "df's" charge duration for the given charge. "k's" purpose is to navigate the "df" data frame.

References

[1] S. Panichtanakom, K. Chalermyanont and S. Thienmontri, "Study of Plug-in Electric Vehicles Charging by Using Load Shaving Method Based on TOU in Distribution System: A Case Study in Thailand," in *2018 International Electrical Engineering Congress (iEECON)*, IEEE, 2018, pp. 1-4.

[2] "Types of Vehicles," Monterey Bay Electric Vehicle Alliance, 2013. [Online]. Available: https://mbeva.org/pev-resources/types-of-vehicles/. [Accessed 16 June 2020].

[3] D. Hofstetter, "BEVs vs. PHEVs: Peaceful Coexistence or Life-or-Death Battle?," Greentech Medi, 21 October 2011. [Online]. Available: https://www.greentechmedia.com/articles/read/bevs-vs-phevs-peaceful-coexistence-or-life-or-death-battle. [Accessed 16 June 2020].

[4] "How do electric car batteries work?," EnergySage, 15 November 2018. [Online]. Available: https://www.energysage.com/electric-vehicles/101/how-do-electric-car-batteries-work/. [Accessed 16 June 2020].

[5] "Tesla Model S, lithium-ion battery chassis with suspension, powertrain, and wheels," GTcarlot.com, 6 February 2012. [Online]. Available: http://gtcarlot.com/news/photo.php?id=59858722. [Accessed 16 June 2020].

[6] "CHOOSING THE RIGHT CHARGING EQUIPMENT FOR YOUR PEV," Liberty Utilities Company, 18 July 2016. [Online]. Available: https://www.evolvewithliberty.com/fact_sheets/facts-about-evs/. [Accessed 16 June 2020].

[7] "Level 1 vs Level 2 EV Charging Stations," ClipperCreek, Inc., [Online]. Available: https://www.clippercreek.com/level-1-level-2-charging-stations/. [Accessed 16 June 2020].

[8] D. Kettles, Electric vehicle charging technology analysis and standards, Florida Solar Energy Center, FSEC Report Number: FSEC-CR-1996-15, 2015.

[9] "Charging Your Plug-in Electric Vehicle at Home," Energy Saver, 13 May 2013. [Online]. Available: https://www.energy.gov/energysaver/articles/charging-your-plug-electric-vehicle-home. [Accessed 16 June 2020].

[10] "In-Home Level 2 EV Charger," Green Mountain Power, 4 March 2018. [Online]. Available: https://greenmountainpower.com/product/home-level-2-ev-charger/. [Accessed 16 June 2020].

[11] Autotrader, "2020 Tesla Model S Review," Autotrader, 19 November 2019. [Online]. Available: https://www.autotrader.com/car-reviews/2020-tesla-model-s-review-281474979974731. [Accessed 16 June 2020].

[12] EVANNEX, "Tesla Model S Owner Test Drives 2020 Model S For Comparison," InsideEVs, 5 February 2020. [Online]. Available: https://insideevs.com/features/396934/2020-tesla-model-s-test-drive/. [Accessed 16 June 2020].

[13] A. Bari, J. Jiang, W. Saad and A. Jaekel, "Challenges in the smart grid applications: an overview," *International Journal of Distributed Sensor Networks,* vol. 10, no. 2, p. 974682, 2014.

[14] R. McKay, "What is a smart grid?," IBM Industries Blog, 17 April 2018. [Online]. Available: https://www.ibm.com/blogs/industries/what-is-the-smart-grid/. [Accessed 16 June 2020].

[15] M. Molnár and I. Vokony, "The Cyber-Physical Security of the Power Grid," IEEE Smart Grid, November 2019. [Online]. Available: https://smartgrid.ieee.org/newsletters/november-2019/the-cyber-physical-security-of-the-power-grid. [Accessed 16 June 2020].

[16] D. Brett, "Smart Grid," Student Energy, 18 May 2015. [Online]. Available: https://www.studentenergy.org/topics/smart-grid. [Accessed 16 June 2020].

[17] S. Khatiri-Doost and M. Amirahmadi, "Peak Shaving and Power Losses Minimization by Coordination of Plug-In Electric Vehicles Charging and Discharging in Smart Grids," in *2017 IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, IEEE, 2017, pp. 1-5.

[18] "TIME-OF-USE," Hydro Ottawa, 24 March 2020. [Online]. Available: https://hydroottawa.com/accounts-services/accounts/time-use. [Accessed 16 June 2020].

[19] "Electricity & Energy Tariffs in Victoria – Great Value," GloBird Energy, 5 December 2015. [Online]. Available: https://www.globirdenergy.com.au/offers-rates/all-tariffs/understanding-tariff/. [Accessed 16 June 2020].

[20] "Demand Side Management," Energy Market Authority of Singapore, 9 March 2020. [Online]. Available: https://www.ema.gov.sg/Demand_Side_Management.aspx. [Accessed 16 June 2020].

[21] F. Kreith and D. Y. Goswami, Energy management and conservation handbook, CRC Press, 2007.

[22] T. Logenthiran, D. Srinivasan and T. Z. Shun, "Demand Side Management in Smart Grid Using Heuristic Optimization," *IEEE transactions on smart grid,* vol. 3, no. 3, pp. 1244-1252, 2012.

[23] M. H. Alham, M. Elshahed, D. K. Ibrahim and E. E. D. A. El Zahab, "Optimal operation of power system incorporating wind energy with demand side management," *Ain Shams Engineering Journal,* vol. 8, no. 1, pp. 1-7, 2017.

[24] "What is Data Analytics?," Master's in Data Science, 16 May 2019. [Online]. Available: https://www.mastersindatascience.org/resources/what-is-data-analytics/. [Accessed 16 June 2020].

[25] "What Is Data Analytics?," Lotame Solutions, Inc., 23 September 2019. [Online]. Available: https://www.lotame.com/what-is-data-analytics/. [Accessed 16 June 2020].

[26] Y. Zhang, T. Huang and E. F. Bompard , "Big data analytics in smart grids: a review," *Energy Informatics,* vol. 1, no. 1, p. 8, 2018.

[27] M. Muratori, "Impact of uncoordinated plug-in electric vehicle charging on residential power demand," *Nature Energy,* vol. 3, no. 3, pp. 193-201, 2018.

[28] M. Muratori, "Impact of uncoordinated plug-in electric vehicle charging on residential power demand-supplementary data," *National Renewable Energy Laboratory-Data (NREL-DATA), Golden, CO (United States); National Renewable Energy Laboratory,* 2017.

[29] W. Jian, L. Zhizhen, W. Kuihua, W. Feng and Z. Yi, "Impact of Plug-in Hybrid Electric Vehicles on Power Distribution Networks," in *2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, IEEE, 2011, pp. 1618-1622.

[30] H. S. Salama, S. M. Said, I. Vokony and B. Hartmann, "Impact of Different Plug-in Electric Vehicle Categories on Distribution Systems," in *2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, IEEE, 2019, pp. 109-113.

[31] H. Ramadan, A. Ali and C. Farkas, "Assessment of Plug-in Electric Vehicles Charging Impacts on Residential Low Voltage Distribution Grid in Hungary," in *2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, IEEE, 2018, pp. 105-109.

[32] P. Patil, "What is Exploratory Data Analysis?," Towards Data Science, 23 March 2018. [Online]. Available: https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15. [Accessed 16 June 2020].

[33] S. Narkhede, "Understanding Descriptive Statistics," Towards Data Science, 6 June 2018. [Online]. Available: https://towardsdatascience.com/understanding-descriptive-statistics-c9c2b0641291. [Accessed 16 June 2020].

[34] "Descriptive Statistics," Research Connections, [Online]. Available: https://www.researchconnections.org/childcare/datamethods/descriptivestats.jsp. [Accessed 16 June 2020].

[35] J. Vadakkanmarveettil, "10 Most Popular Analytics Tools in Business," Analytics Training Blog, 15 January 2011. [Online]. Available: https://analyticstraining.com/10-most-popular-analytic-tools-in-business/. [Accessed 16 June 2020].

[36] "TOP 10 DATA ANALYTICS TOOLS," IMS Proschool, 30 October 2017. [Online]. Available: https://www.proschoolonline.com/blog/top-10-data-analytics-tools. [Accessed 16 June 2020].

[37] T. Foster, "11 Most Effective Data Analytics Tools For 2020," AltexSoft, 3 February 2020. [Online]. Available: https://www.altexsoft.com/blog/data-analytics-tools/. [Accessed 16 June 2020].

[38] "Top 25 tools for data analysis and how to decide between them," Stitch, 19 December 2019. [Online]. Available: https://www.stitchdata.com/resources/data-analysis-tools/. [Accessed 16 June 2020].

[39] "10 Best Data Analysis Tools For Perfect Data Management," Software Testing Help, 22 April 2020. [Online]. Available: https://www.softwaretestinghelp.com/data-analysis-tools/. [Accessed 16 June 2020].

[40] "Tidyverse," Tidyverse, 2017. [Online]. Available: https://www.tidyverse.org. [Accessed 16 June 2020].

[41] "Introduction to Tidyverse - A Collection of R Packages for Data Science," Analytics Vidhya, 13 May 2019. [Online]. Available: https://www.analyticsvidhya.com/blog/2019/05/beginner-guide-tidyverse-most-powerful-collection-r-packages-data-science/. [Accessed 16 June 2020].

[42] H. Wickham, W. Chang, L. Henry, T. L. Pedersen, K. Takahashi, C. Wilke, K. Woo, H. Yutani and D. Dunnington, "Create Elegant Data Visualisations Using the Grammar of Graphics • ggplot2," Tidyverse, 2016. [Online]. Available: https://ggplot2.tidyverse.org. [Accessed 16 June 2020].

[43] H. Wickham, R. François, L. Henry and K. Müller, "A Grammar of Data Manipulation • dplyr," Tidyverse, 2017. [Online]. Available: https://dplyr.tidyverse.org/index.html. [Accessed 16 June 2020].

[44] V. Spinu, G. Grolemund and H. Wickham, "Make Dealing with Dates a Little Easier • lubridate," Tidyverse, 2017. [Online]. Available: https://lubridate.tidyverse.org. [Accessed 16 June 2020].

[45] J. Hester, "glue 1.2.0," Tidyverse, 2017. [Online]. Available: https://www.tidyverse.org/blog/2017/10/glue-1.2.0/. [Accessed 16 June 2020].

[46] "How much does leaving electrical appliances on standby cost?," npower, 2 December 2013. [Online]. Available: https://www.npower.com/blog/2013/12/02/cost-of-appliance-standby/. [Accessed 16 June 2020].

[47] "Running your pool pump on off peak electricity," Grenum, [Online]. Available: https://grenum.com/au/running-pool-pump-on-off-peak-electricity/. [Accessed 16 June 2020].

[48] "Peak Shaving | What it is & how it works," Next Kraftwerke, 2019. [Online]. Available: https://www.next-kraftwerke.com/knowledge/what-is-peak-shaving. [Accessed 16 June 2020].