MONASH University

WP1.2 Technical modelling of electrification of transport profiles

Milestone Report: EV charging profile extraction methodology and results

DISCLAIMER AND ACKNOWLEDGEMENT

Acknowledgement and Disclaimer:

This research is part of a program of fourteen projects, which have variously been funded and supported under C4NET's Enhanced System Planning collaborative research project. While conducted independently, the authors acknowledge the funding and in-kind support of C4NET, its members and supporters, and its ESP project partners.

MONASH TEAM

Jinhao Li, Leo Li, Chenghao Huang, Hao Wang (Lead)

Department of Data Science and AI

Monash University

In addition to the current team members, we would like to acknowledge Ryan Wang and Xiangrui Li for their contributions to an earlier version of this report.

Contact: Dr Hao Wang

Email: hao.wang2@monash.edu

EXECUTIVE SUMMARY

This report presents a framework for analysing electric vehicle (EV) charging patterns using data from smart meters. Through this project, we extracted EV charging profiles, to serve as a foundation for subsequent EV modelling efforts. Our EV charging disaggregation/extraction approach allows for a detailed understanding of how to identify and extract EV charging profiles from overall energy consumption in the meter data. In this report, we focus on extracting EV charging profiles for non-PV customers as they are not affected by PV generation leading to likely degraded performance, especially during the daytime.

We would like to highlight that disaggregating EV charging profiles without the ground truth is a very challenging task. The task can be affected by various factors, such as the possible impacts of COVID on the residential load (as we rely on load data in 2020 and 2022) and lack of information regarding whether customers have new appliances or batteries installed at home during the period 2020-2022. Moreover, the presence of batteries is not considered in this report due to a lack of information, nor are changes in EV ownership and charging locations. Key assumptions underpinning the analysis include consistent non-EV energy consumption patterns over time. These assumptions are critical to our developed method in identifying the EV charging load from other household energy uses. Statistical filtering techniques were employed to address potential outliers and variability in customer behaviour, enhancing the reliability and accuracy of the analysis.

The technical methodology involves several critical steps: preprocessing smart meter data, identifying EV customers, differentiating energy consumption, applying travel behaviour filters, and carefully examining the impacts of extreme weathers on the disaggregated EV charging profile. Each step is meticulously designed to refine the data to be used in the methodology. For instance, preprocessing smart meter data ensures accuracy from the beginning, while identifying EV customers involves cross-referencing with VIC Government EV rebate programs and vehicle registration data. Differentiating energy consumption separates EV charging load from other uses, travel behaviour filters ensure that charging patterns align with realistic driving habits, and the extreme weather check component ensures that the extracted charging patterns will not be affected by the extreme weather conditions, such as extreme hot and cold days.

In the implementation phase, we applied our methodology on all the meter data provided to us. During the initial data processing, the sample size was reduced by eliminating incomplete or inconsistent data. Identifying EV customers further reduced the dataset by filtering out non-PV customers. Differentiating energy consumption helps in identifying likely EV charging load, while travel behaviour filters ensure that only realistic and consistent charging patterns are considered. The sampled weekly EV charging profiles along with meter load, as well as the verification on extreme weather days, demonstrate the effectiveness and robustness of our framework. Further analysis on EV charging on weekdays and weekends provide insights into individuals' charging behaviours. More importantly, it is worth noting that, after classifying each customer to a specific type of charger, the majority of customers identified from the dataset use the 2.3kW charger at home, while a small number of customers use 3.7kW or 7.4kW chargers. Therefore, we acknowledge that the disaggregated EV charging profiles for customers with these two kinds of chargers may not represent the general EV charging profiles for 3.7kW or 7.4kW charger due to this small number of customers, which may lead to bias.

This approach ensures that the extracted data is reliable, which is crucial for subsequent modelling and analysis efforts. Through detailed data integration, statistical analysis, and modelling techniques, this report presents the framework for extracting and understanding EV charging patterns from smart meter data. The developed EV charging extraction framework/tool, along with the insights gained, can inform future EV charging modelling and their potential impacts, thereby supporting necessary network management and planning in the era of transport electrification.

GLOSSARY OF TERMS/ABBREVIATIONS

PV	Photovoltaic
EV	Electric vehicle
DNSP	Distributed network service provider
NMI	National meter identifier

1. Introduction

According to the previous deliverable report (on literature review), existing literature highlights various approaches to modelling EV charging patterns, ranging from stochastic models to machine learning techniques. Most existing approaches depend on labelled data, i.e., charger data together with meter data. But the available data in our report does not contain the corresponding label information (i.e., ground truth) about customer charger data, thus we take a rule-based approach, using day-to-day comparisons of meter data for the same customer over two different years to capture EV charging activities hidden in the meter readings.

Consequently, this report dedicates substantial attention to data wrangling and preparation processes. It also covers detailed descriptions of our assumptions and methodologies. This report focuses on leveraging our rule-based model to extract EV charging profiles for customers without solar PV installation, as these customers are not affected by the solar PV generation likely leading to degraded performance. It is worth mentioning that some of the knowledge of EV charging power rates, the travel behaviour data, and the weather data used in our method settings were independently sourced from reports released by government agencies and previous projects, such as [1] - [6], [11]. Also, we utilise temperature data from the Bureau of Meteorology to facilitate our analysis of the impacts of extreme weathers on disaggregated EV charging profiles.

By meticulously addressing data quality and preparation, this report lays the groundwork for effective and reliable EV charging pattern analysis. It is worth mentioning disaggregating EV charging profiles without ground truth data is an interesting topic but challenging. The used method in this report and extracted EV profiles can be influenced by various other factors, such as potential impact of COVID-19 on residential load and customer devices, such as home batteries, that are invisible in the used data.

In the following, this report will present the dataset overview in Section 2 and Data wrangling and preprocessing in Section 3. The key assumptions used in the proposed methodology are presented in Section 4. The main body of the methodology, including the toolbox and rule-based model, is presented in Section 5. Then Section 6 illustrates the EV charging disaggregation results, and finally, Section 7 discusses the validation on disaggregated EV charging profiles.

2. Input dataset overview

Figure 1 DNSPs map in Greater Melbourne [7]

The input data contains two different types of datasets including smart meter data and information on customer participation in the EV rebate program. These datasets encompass five different DNSPs in Victoria: AusNet serving eastern Victoria, PowerCor covering western Victoria and the western suburbs of Melbourne, Citipower operating in Melbourne CBD and inner suburbs, Jemena serving northern and western Melbourne, and United Energy covering south-east Melbourne and the Mornington Peninsula, as shown in Figure 1.

2.1. Smart meter dataset

The smart meter dataset contains electricity usage data collected at regular intervals (i.e., every half hour, totalling 48 timesteps per day) from smart meters installed at customer premises. In general, the time-series data cover the period from January 2020 to the end of March 2023[1]. However, due to differences in the final update dates among different DNSPs, there are slight variations in the end dates for 2023. For instance, some DNSPs have data covering until mid-March, while others extend to early April. The data collected in 2020 are under the periods of Covid-19 and there was lockdown during this year in Melbourne areas. The lockdown can affect residents' electricity usage habits. For example, the average electricity consumption might increase during this period. However, considering that many residents will continue to follow the electricity usage habits they developed during the lockdown, such as working from home, even after the lockdown ends, analysing the impact of the lockdown is complex and beyond our ability to consider.

[1] In particular, United Energy only provides data for a two-year period, from January 2020 to April 2022.

Here are the key components of the smart meter dataset:

- **Identification of customers:** A random unique identifier (e.g., NMI[2], hash code of NMI) for each customer, ensuring data anonymity and integrity.
- **Classification of customers:** The random unique identifiers are classified to identify the type of customer, such as business, industry, or residential. This classification helps in understanding consumption patterns specific to a customer category.
- **Timestamp**: The precise date and time when the data was recorded, allowing for the analysis of consumption patterns over time.
- **Channel records**: Information on the energy imported from and exported to the upper grid[3], essential for understanding the bidirectional flow of energy and the impact of distributed generation (e.g., PV) for some customers.
- **Electricity usage**: The amount of electricity consumed corresponding to each interval, providing a detailed view of detailed time-series energy consumption or energy exported to the upper grid during recorded timestep, the unit is measured in kWh.

2.2. EV rebate program dataset

The EV rebate program dataset provides detailed information about customer participation in the EV rebate program. This dataset allows us to identify which customers are part of the EV rebate program, distinguishing EV customers from the overall customer base.

The key components of the EV rebate program dataset:

- **Identification of customers:** Similar to the smart meter dataset, each customer is identified by a random unique identifier, ensuring consistency and enabling the linkage between the datasets.
- **Status of the program participation:** Indicates whether the customer participates in the EV rebate program.

In addition to the key components of these two datasets, some DNSPs also provide geographic identifiers (i.e., 4-digit postcodes) for some customer locations.

2.3. Daily temperature dataset

To examine the impacts of extreme weathers, e.g., extreme hot or cold days, on the extracted EV charging patterns, we collect the time-series temperature data for all weather stations located in Victoria from the official website of the Bureau of Meteorology [11]. The weather station distribution is illustrated in Fig. 2 for reference.

It is worth mentioning that, we find that the Bureau of Meteorology releases the daily maximum and minimum temperatures per day at each weather station, but could not get access to the temperature data with higher resolution, e.g., in hourly intervals. Therefore, due to the data limitation, in this report, we define the extreme hot and cold days based on the daily maximum and minimum temperatures.

---

[2] Note "NMI" does not refer to the real NMI, it is a random identifier. .

[3] For customers with PV, the two channels are labelled as: "E" refers to the energy import channel, and "B" refers to the energy export channel. This label principle has been utilised in all DNSPs.

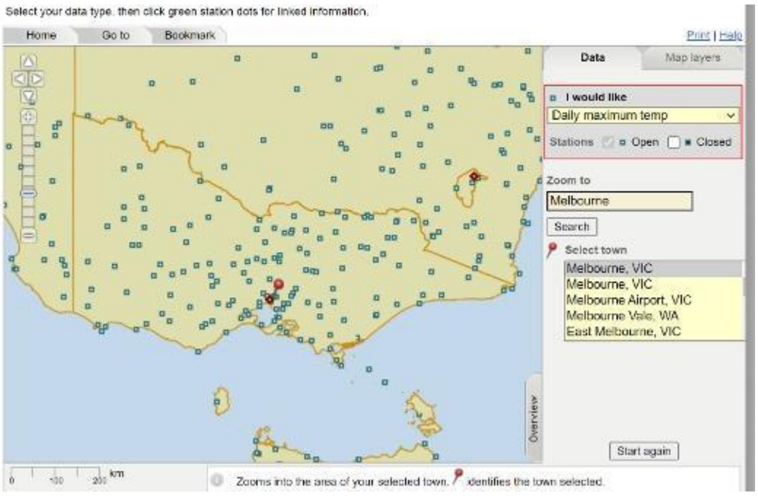

Figure 2 Weather stations across Victoria shown in website of the Bureau of Meteorology

Due to the diverse sources of our data, there are variations in data formats resulting from differences in how each DNSP collects and stores data. Furthermore, the input datasets have other issues that need to be addressed, such as missing values, inconsistencies, and discrepancies in data formats. These challenges are discussed in more detail in the data wrangling section.

3. Data wrangling

3.1. Smart meter and EV rebate datasets

The data provided to us is benefited from the standardisation of NEM12 format. However, the data storage processing may lead to some anomalies, such as data format inconsistency and duplication. To ensure data consistency and accuracy, we performed data wrangling on the raw data (including the EV rebate file and meter data file) to enhance the reliability of the extracted EV data and subsequent analysis. During the data-wrangling process, we identified and corrected some inconsistencies and anomalies in the raw data as below:

- **The storage format of date time inconsistencies:** By examining the data from various DNSPs, we discovered issues where the positions of the day and month in the date were swapped. For example, in some files, the dates were stored in the DD/MM/YY format, while in others, they were stored in the MM/DD/YY format. Additionally, we found that some DNSPs used Australian Eastern Standard Time (AEST) for storing time data, whereas others used Coordinated Universal Time (UTC). Mixing AEST and UTC dates can cause inaccurate analysis, data inconsistency, and complicated processing. Therefore, during our data wrangling process, we standardised the mentioned various date and time format.
- **Column name inconsistencies:** We found inconsistencies in column names across different DNSPs. For instance, the column for identifying the customers, the random unique identifier, appeared under various names in our data[4]. Similarly, this issue was present in columns identifying whether a customer is residential[5], columns identifying meter channels[6], and columns indicating the timestep[7]. Inconsistencies in column names may lead to difficulty in merging or joining data frames, resulting in errors or incomplete datasets. Thus, we standardised the column names and their order across the datasets from different DNSPs.
- **The missing and duplicate data:** By comparing records with the same random unique identifier at the same timestamp, we found instances of data duplication in both the meter data and the EV rebate file. For example, the meter data for the same customer at the same time was recorded multiple times, and the same customer's information appeared multiple times in the EV rebate file. Additionally, we discovered missing meter data for some customers, some meaningful information was recorded as "Not available" or certain customers lacking an entire quarter's worth of meter data within a year, which may cause imprecisely analysis, increased computational costs, misleading metrics, and compromised results of extracted EV charging data. The missing of some key columns can cause significant problems. For example, the NMI_classification used to identify residential and business customers. Its absence would result in our inability to determine whether a customer is residential[8]. In

[4] For example, Jemena uses "Identifier_HASH", for this same term, AusNet uses "NMI", United Energy uses "meter_id" , etc.

[5] For example, Jemena uses "NMI_CLASS", PowerCor uses "cust_type", etc.

[6] For example, Citipower uses " channel_type", AusNet uses "CHANNEL", etc .

[7] For example, PowerCor uses "read_date", United Energy uses "start_timestamptz_ts", ect.

[8] For example, United Energy has a lack of columns "NMI_CLASSIFICATION".

our data wrangling, we deleted the duplicate data, and focused on the customers with complete information.

- **Semantic error:** We identified semantic errors in the records, one is the mislabeling in the classification of customers. Many customers have been labelled as residential despite exhibiting unreasonable energy consumption usage, which are more characteristic of industrial or commercial customers. This mislabeling could result from changes in customer classification over time or errors in the initial data entry. The other one is that if customers install battery storage at home, the variation can be very large, and the factors of uncertainty will increase accordingly. Therefore, we compared the PV installation size (e.g., rated power) with the maximum values recorded in the export channel. This method allowed us to exclude some residents with installed batteries and reduce the uncertainty. To address these semantic errors, we implemented a method to cross-verify the customer classifications. Specifically, for the customer classification mislabeling, we identified outliers by detecting anomalous consumption patterns that do not align with typical residential usage[9].

3.2. Weather dataset

The downloaded weather station dataset includes the longitude and latitude information for each weather station. To match each customer with a specific weather station for the following extreme weather check, we therefore need to match each weather station with a postcode same to or close to each customer with their corresponding postcodes provided in the dataset. To achieve this, we find the geographical boundaries, with longitude and latitude information, of each suburb across Victoria from the governmental DataVic website [12]. After that, we traverse all weather stations and suburbs to determine the station's postcodes by checking whether the station is within a specific suburb's boundary.

[9] In Australia, average daily energy consumption for residential customers is typically ~60 kWh/d [8].

4. Key assumptions

To ensure the robustness and reliability of our analysis, our technical methodology is predicated on several key assumptions. These assumptions provide a foundational framework that allows us to control for various variables and isolate the factors of interest, i.e., EV charging in the meter data. By clearly defining these assumptions, we aim to enhance the interpretability and validity of our findings, acknowledging the inherent limitations and addressing potential sources of variability.

Below are the key assumptions underpinning our study :

- **Consistency in electrical appliance usage and consumption patterns:** Since we do not have ground truth data about EV charging in the meter data, our methodology to be detailed in the next section relies on comparisons between day-to-day load for the same customer in two years. Hence, for each individual customer, we assume that there is no significant change in the energy consumption pattern of other electrical appliances (excluding EV) for the same day of week and time across three continuous years. Typical appliances could include the water heaters (heat pumps), other appliances like kettles, and cooking related devices.
- **Regularity of EV charging events:** We do not consider customers who exhibit sporadic charging events at home. Instead, we focus on those who have regular charging patterns at home, supporting the desired commute travel distance with an adjustable frequency[10]. The main reason is that the extracted signals, which are very sparse and of very short duration, are highly likely to be caused by noises, which mean other appliances' loads (not EV charging). Since we cannot ensure that these noisy signals are EV charging profiles, our methodology inclines to remove them to ensure the reliability of the EV charging profiles extracted.

Some of these assumptions, while strong, are believed to be necessary in order to create a controlled environment for our analysis, with the lack of ground truth charger data. These assumptions help our methodology in reducing the impact of noises and variability in the meter data.

[10] Regarding EV charging, sporadic charging events at home are not the primary focus of our analysis. While we do not exclude them entirely, our emphasis is on regular charging patterns that align with typical usage scenarios. In addition, this assumption is grounded in the report [1], which suggests that typical EV usage involves consistent weekly charging to meet travel needs.

5. Technical methodology

5.1. Proposed toolbox overview

The proposed toolbox integrates with the aforementioned smart meter dataset, EV rebate program dataset, weather dataset, and travel behaviour profile (e.g, travel distance) [1] - [6], [11]. Specifically, the smart meter dataset is divided into two distinct periods: one representing a time when customers do not have EVs (Period 1) and another when they do (Period 2). The key idea is to extract EV charging data by analysing the increased load consumption observed between these two periods[11]. This is because we do not have any ground truth data about EV charging of these customers in the dataset. Through contrasting loads on the basis of day-to-day comparison over two years, in which customers do not have EVs in the first year and have EV charging at home in the later year, we aim to capture the additional power usage in half-hour intervals from the meter data to extract EV charging profiles.

The analysis begins with a rule-based model designed to evaluate energy consumption. This model includes an output layer that applies statistical methods to filter out extreme customers, ensuring the dataset's integrity and focusing on typical energy consumption patterns. The subsequent step involves a detailed analysis of the increase in the energy consumption, aiming to detect potential EV charging curves. This step is critical as it differentiates the additional energy consumption due to EV charging from other possible increases in energy usage. It is worth noting that because of the lack of information about detailed devices and activities of customers at home, various factors, such as the potential impact of COVID-19 and home devices like batteries, may affect the accuracy of the designed method.

In addition, a machine learning model, equipped with a buffer (i.e., knowledge pool covering typical commercial household EV charger level in the relevant area) is developed. This model processes the detected charging curves to cluster charging rates effectively. By identifying clusters within the data, the model can determine typical charging behaviours and rates among different customer segments. Following the clustering, the model refines the detection process to ensure that the identified EV charging patterns are effective and reliable. This refinement is crucial for minimising false positives and ensuring that the charging events identified are indeed related to EV charging.

The final stage of the analysis involves verification and adjustment of the identified EV charging profiles. This step incorporates travel constraints derived from the travel behaviour profiles, which helps to contextualise the charging events within the expected travel patterns of EV customers. By considering these constraints, the toolbox can validate whether the detected charging profiles align with realistic EV usage scenarios. The EV charging data is potentially influenced by weather and batteries. For instance, consumption on the same date in two different years may vary due to differences in weather conditions, and the total consumption recorded by the meter may also differ because of battery impact. We need to acknowledge the limitations of lacking relevant data of residential battery installation. To examine the effects of weather conditions on the

[11] The mentioned two periods of data (one without EVs and one with EVs) are essential inputs for the proposed toolbox. It is recommended that these two periods cover adjacent years to better support and apply our assumptions. Besides, we can appropriately account for seasonal variations in energy consumption, thereby enhancing the accuracy and reliability of our analysis.

disaggregated EV charging for the verification purpose, we used the collected daily minimum and maximum temperatures to identify extreme weather days, which can help further examine the disaggregated EV charging patterns on those days. It is worth mentioning that many of the parameters (e.g., the buffer, travel behaviour profiles, etc) within our proposed rule-based model are adjustable based on actual conditions. These parameters can be fine-tuned to better reflect specific scenarios and requirements, ensuring that the toolbox is adaptable and can provide precise and relevant results depending on the context in which it is applied.

Figure 3 illustrates the complete framework of the proposed toolbox, highlighting the flow of data and the sequential steps involved in the analysis. In the following subsections, we will introduce details of the filters and related mechanisms in the rule-based model step by step.

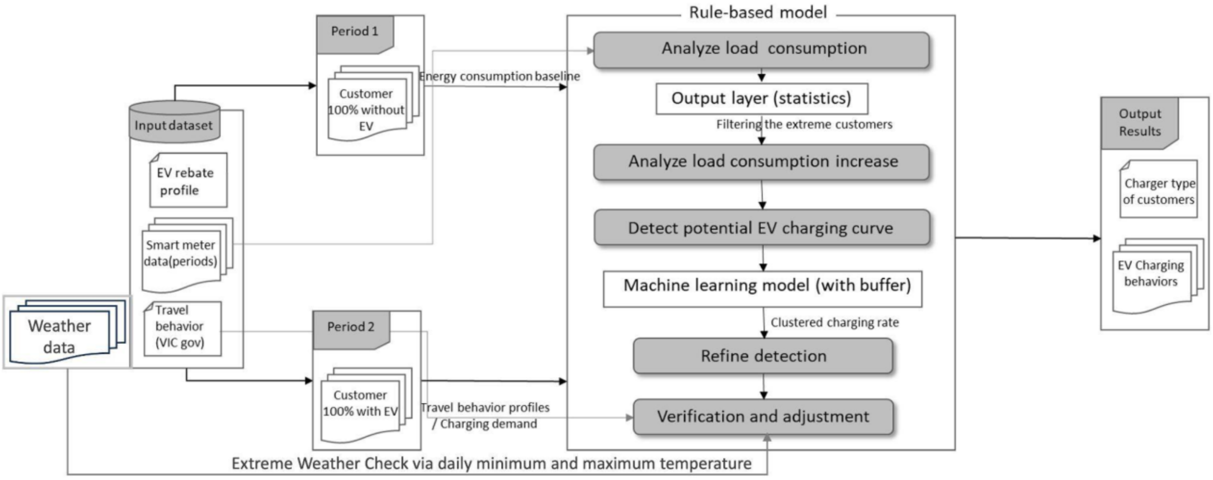

Figure 3 Framework of the proposed toolbox

5.2 Rule-based model

5.2.1. Smart meter data initialisation

During this step, we initialise the smart meter data by addressing various issues identified during the data wrangling process. This includes standardising date-time formats, unifying column names, and handling missing or duplicate data. Additionally, we correct semantic errors such as mislabeling in customer classifications. These steps ensure that the data is clean, consistent, and reliable, providing a solid foundation for subsequent analysis.

5.2.2. EV customer identification

In this step, we identify EV customers based on the EV rebate program dataset and the smart meter dataset. This identification process involves a crucial data integration: We merge the EV rebate program data with the smart meter data, aligning the two datasets based on customer identifiers. This integration ensures that we can accurately correlate rebate participation with actual energy consumption patterns. By carefully matching these identifiers, we create a dataset that combines customer participation in the EV rebate program with their detailed energy usage data. The outcome of this identification process is a list of customers who are EV customers, along with their associated charging patterns.

5.2.3. Energy consumption differentiation

In this step, we perform a detailed differentiation analysis to quantify the additional energy consumption specifically attributed to EV charging. This analysis compares energy usage between the baseline period[12], when customers did not have EVs, and the periods[13] with EV adoption. The goal is to isolate the impact of EV charging by considering typical consumption patterns.

First, we incorporate data from two distinct periods, ensuring that only customers with complete data in both periods are considered for further analysis.

Next, based on the given smart meter data, we calculate the load balance for customers to understand the overall energy dynamics within their households.

By analysing the load balance, we can identify patterns of energy usage and distribution, which helps in isolating the specific impact of EV charging on the overall energy consumption.

In detail, when customers do not own EVs, for non-PV customers, the load balance can be calculated by:

$$E_{i,y0}^{Meter} = E_{i,y0}^{App} , \tag{1}$$

when customers own EVs, for non-PV customers, the load balance can be calculated by:

$$E_{i,y2}^{Meter} = E_{i,y2}^{App} + E_{i,y2}^{EV} , \tag{2}$$

where

- Subscript i represents non-PV customer index, and $y0, y2$ represent period index .
- Superscript is applied to separate different types of energy, for example, $Meter$ refers to the energy consumption recorded meter data.
- $E_{i,y0}^{Meter}$, $E_{i,y2}^{Meter}$represent the vector of energy consumption of non-PV customer i recorded in $y0$, $y2$.
- $E_{i,y2}^{EV}$ denotes vector of energy related EV charging of non-PV customer i recorded in $y2$.
- $E_{i,y0}^{App}$, $E_{i,y2}^{App}$ denotes vector of energy consumed by other electrical appliances excluding EV of non-PV customer i at $y0$, $y2$.

Then, we determine the overall increase in energy consumption by comparing the total energy usage between y2 and y0. This step identifies the gross change in consumption, ΔE^{Meter} , which includes all factors influencing energy use during these periods.

For non-PV customers, it can be calculated by (3):

$$\Delta E_{i}^{Meter} = E_{i,y2}^{Meter} - E_{i,y0}^{Meter} = E_{i,y2}^{EV} + (E_{i,y2}^{App} - E_{i,y0}^{App}). \tag{3}$$

According to the first two key assumptions, for all customers, we have $E_{i,y2}^{App} \approx E_{i,y0}^{App}$.

Therefore, the EV charging profiles can be approximated as (4) for non-PV customers:

$$E_{i,y2}^{EV} = \Delta E_{i}^{Meter} . \tag{4}$$

By performing these calculations, we effectively isolate the energy consumption specifically attributed to EV charging. This differentiation helps in understanding the true impact of EV adoption on the energy grid, free from the confounding effects of other electrical appliances.

---

[12] We define y0 as the year of the baseline period.

[13] We define y2 as the year with EV adoption.

After disaggregating EV charging profiles from the load difference between 2020 and 2022, we also need to design an additional filter to make sure that the EV charging behaviours align with the common customers' practice, i.e., the EV charging will last for a relatively long time. This filter operation will help us filter out load increases which are unlikely to be EV charging, e.g., kettle load. To achieve this, we set a parameter for such an operation, with an default value as one hour, which means that the charging duration for our extracted EV charging profiles is at least one hour, such that short-duration load will be extracted as EV charging. This parameter is also tunable to control the outputs and mitigate the impact of other devices.

Regarding other power increases which might be due to EV charging such as hot water load, given that our disaggregation approach applies contrasting load in 2020 and load in 2022. The assumption is that the hot water consumption patterns do not have a significant change and hot water consumption has existed in the 2020 load. The biggest change in energy consumption and in particular increases in power from the meter readings in 2022, compared to that in 2020, is because EV charging, as those customers selected in this study have applied for EV rebate in 2021 between 2020 and 2022.

5.2.4. Meter load data alignment between 2020 and 2022

Additionally, instead of performing the above differentiation using the same calendar date in 2020 and 2022 (e.g., 01/01/2020 and 01/01/2022), we conduct the day-of-week alignment using the same day of weeks between two years in order to mitigate the impact of weekday/weekend differences in load patterns. This alignment could further enhance the accuracy of our EV charging disaggregation model, as the load profiles may have more consistency for the same day of the week at a similar time of the year and season. Specifically, our disaggregation starts from Monday. The first Mondays of 2020 and 2022 are 03/01/2020 and 06/01/2022. We use the meter load data from these two days to disaggregate EV charging profiles for this specific Monday and so on and so forth for the rest of the year. The last day of 2022, i.e., 31/12/2022, corresponds to 28/12/2020 in our disaggregation scenario. To complete the whole disaggregation process, the meter loads between 29,30,31/12/2020 and 01,02,03/01/2022 are aligned for disaggregation, so as well for meter loads between 01,02/01/2020 and 04,05/01/2022.

5.2.5. Ensuring consistent consumption via outlier detection

Despite our assumptions, real-world scenarios may introduce changes in energy consumption behaviour, such as a high diverse energy consumption increase, the purchase of new high-energy appliances by customers, etc., which could affect energy usage patterns. To support the first key assumption, it is crucial to identify and filter out customers whose consumption patterns have changed at the same date and time across different years. This ensures that observed increases in energy consumption are primarily due to EV charging rather than other variables. To achieve this, we propose a statistical-based approach to filter out these unstable customers (outliers). We would like to highlight that the omission of outliers is reasonably based on the two key assumptions detailed in Section 4, aimed at improving the integrity of the disaggregation results. Additionally, the load patterns and EV usage of outliers may not accurately reflect the real-world scenarios. Therefore, our main focus is on investigating and analysing customers who align with our assumptions.

First, following the above energy consumption differentiation, we need to identify customers who did not have an EV in y0, and adjacent year[14], and then find the intersection of these customers using the same merge principal as mentioned in the last subsection. Then, we use the meter data from y1 and subtract the meter data from y0 to obtain the increase in energy consumption.

$$E_i^{Meter\,*} = E_{i,y1}^{Meter} - E_{i,y0}^{Meter} \quad (5)$$

According to our assumption, if a customer has not purchased new high-energy appliances and their electricity usage habits are stable, the increase in load at the same date and time across the two years should be minimal. Customers with significant increases in energy consumption are identified as outliers and need to be filtered out. To quantitatively determine the extent of load increase that qualifies as unstable, we use the three-sigma method to establish the threshold.

In statistics, the three-sigma rule states that approximately 99.7% of the data points lie within three standard deviations (σ) of the mean (μ). Data points outside this range are considered outliers. We apply this rule to filter outliers in our data.

$$Pr(|X - u| \geq 3\sigma) < 0.003 \quad (6)$$

Further, in our method, we apply L2 normalisation to the differentiation results of energy consumption between years 0 and 1. Apply L2 normalisation to the differentiation results. The result of this differencing could be positive or negative. Regardless of whether the result is positive or negative, the magnitude of the absolute value indicates stability or instability. Therefore, the L2 norm meets our requirements, ensuring that the differencing results are always greater than or equal to zero.

Let $x_i = ||E_i^{Meter\,*}||_2$. The whole process can be described as follows:

- Calculate the mean (μ_x) and standard deviation (σ_x) of the power consumption increases for N customers with index i:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad (7)$$

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu_x)^2}. \quad (8)$$

- Identify outliers as customers whose increase in power consumption falls outside the range.

The flowchart of this step is illustrated as Fig. 3. By filtering out these outliers, we minimise the impacts of new high-power electrical appliances on our analysis. This ensures that the calculated power consumption increase is more likely attributed to EV charging rather than other variables. This approach supports the first key assumption and enhances the accuracy of our methodology.

Finally, depending on the availability and quality of supporting data, we may choose to remove these outlier customers from our dataset. This optional step helps maintain the integrity of our analysis by ensuring that only stable consumption patterns are considered, leading to more effective and reliable results in understanding EV charging impacts.

[14] This processing step is an optional step based on the dataset you have, you can neglect this step, if you do not have this kind of dataset. We define another adjacent year in which customers do not have an EV as y1.

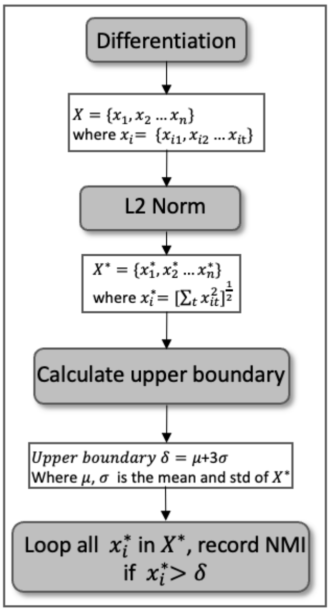

Figure 3 Diagram to identify the customers with unstable energy usage patterns

5.2.6. EV charger type infer

It is necessary to have the information of EV chargers utilisation when doing EV charging extraction, but EV chargers come in various types and power levels in Australia. According to the Australia Electric Vehicle Council [3], chargers range from simple household outlets to high-powered DC fast chargers. However, our dataset lacks explicit information about the specific charger types and levels used by customers. To overcome this data limitation, we employed a machine learning approach to cluster the relevant EV charging data extracted in the previous steps.

We utilised the K-means algorithm, a well-known unsupervised learning method, to infer the possible charger types within DNSP. The K-means algorithm divides data points into clusters based on their similarities, where each point belongs to the cluster with the nearest centroid. This method allows us to group potential charging events that exhibit similar characteristics, potentially indicating the use of similar charger types.

To support this clustering, we referred to a knowledge pool (buffer) encompassing the power levels of commonly used EV chargers in the Australian market. This buffer includes data on typical charger power ratings as shown in Table 1. This information provides a reference framework that guides the clustering process and helps interpret the results.

Specifically, the general K-means algorithm process is described below:

1. Set value for hyperparameters K, which represents the number of classes in the dataset.
2. Randomly select the cluster centroid for each class.
3. Loop each point, and calculate the distance between the point and each cluster centroid.
4. Assign each data point to the nearest centroid, forming k clusters.
5. Recalculate the centroid for each cluster.
6. Repeat 3-5 until converged.

Prior to conducting the clustering analysis, we aimed to minimise the influence caused by the sparsity of potential charging events during the K-means [9] clustering. To achieve this, we counted the number of

potential charging events in several periods within a day and selected the periods when charging events were most likely to occur. This approach allowed us to retrieve more accurate candidate charger levels.
After identifying the candidate charging rates, for each customer, we calculate the mode of their potential EV charging rates, which represents the most frequently observed charging rate. We then compare this mode with the maximum power $max(E_{i,y0}^{Meter} / \Delta t)$ when the customer has no EV to check if the load increase[15] is caused by the EV charging.

Table 1 Guide to Electric Vehicle Charging [4]

EV charger types	Level 1 AC[16]		Level 2 AC/DC[17]			Level 3 DC
Example	An electrical outlet		A wall charging unit			A high-powered DC fast-charger
Typical application	Home		Home, work, shopping centre and car park			Regional near highways, motorways and key routes
Power	**2.3 kW**	**3.7 kW**	**7.4 kW**	11 kW	22 kW	50-350kW
Costs	Starting from $500 per car space		Starting from $2,000 per car space			Starting from $10,000 per car space (shared)

Considering that the extracted EV data may sometimes be affected by noise (including the uncertainty mentioned before), we added a noise proximal term to cover more charging events. The proximal term was established based on the maximum consumption observed during non-EV and non-PV periods for each customer. The proximal term can be formulated as follows:

$$\rho = \tau * max(E_{i,y0}^{Meter}), \tag{9}$$

where

- τ is coefficient for the proximal term, which indicates the noise level caused by other appliances. The reason we include this term is that the extracted EV charging profiles may be influenced by some noises in the inherent meter data during data collection. To avoid missing some charging events and achieve more accurate EV charging disaggregation, we first calculate the maximum meter load and use this coefficient to control the noise level. This coefficient is usually a small constant near zero, e.g., 0.05, to avoid over-amplifying the noise for our EV charging disaggregation.

[15] For PV customers, it is supposed to neglect their export channel but instead focus on the import channel.
[16] The charging level of type 1 uses a standard 230-volt AC power point, providing approximately 10 to 12 kilometres of range per hour.
[17] A Level 2 charger is at least twice as powerful as a Level 1 charger and can provide about 35 to 110 kilometers of range per hour. The Level 2 charger supports both AC and DC supply, with the AC supply at high power (e.g., 11 kW and above) being three-phase.

- $max(E_{i,y0}^{Meter})$ represent the maximum consumption observed for the customer during the non-EV and non-PV periods.

When identifying the coverage range for each customer, we can establish a profile that contains the charging level and the coverage range for all of the customers, which will contribute to further refinements.

Furthermore, it could happen that customers may upgrade their chargers to have faster EV charging, e.g., from 2.3kW charger to 3.7kW or 7.4kW chargers. However, such upgrading information is lacking in the current dataset. Thus, it can be challenging to identify those customers with more than one charger at home. We therefore make the reasonable assumption that each customer only has one kind of charger at home. Given the potential for customers using 2.3kW chargers to upgrade their chargers in the future, it would be an interesting point to analyse the charging behaviours for this type of customer. Detailed data regarding how many chargers and corresponding charger types for each customer may help our rule-based model better extract EV charging load under such a multi-charger circumstance.

5.2.7. Refine EV charging based on travel behaviour filter

In this step, we refine our analysis by applying a travel behaviour filter involving average annual travel distance, driving frequency[18], unit distance consumed energy, ect. This ensures that the identified EV charging patterns align with actual travel behaviour, allowing for a more accurate estimation of EV charging needs and optimised charging patterns. By integrating travel behaviour data, we can better understand and predict the energy requirements for EV customers.

First, we calculate the yearly total EV consumption demand via, this is based on the annual travel distance and the energy consumption per kilometer, which gives us an understanding of the yearly energy needs for an EV based on actual driving distances:

$$E^{Year} = D^{Year} \times \beta \ . \tag{14}$$

Next, to understand the regular energy needs within a smaller period (e.g., season, month, week), we distribute the yearly energy demand by dividing it by L, the number of such periods in a year. This step is crucial because energy consumption patterns can fluctuate but typically follow a consistent pattern over the year:

$$E^{L} = \frac{E^{Year}}{L}. \tag{10}$$

Once we have the energy demand for a specific period, we need to determine the required charging duration charge to meet this demand. This is calculated by dividing the energy demand for the period by the charging rate:

$$t_k = \frac{E^{L}}{P_k}. \tag{11}$$

Finally, we rearrange the yearly total EV consumption demand to link the charging rate, duration, and the number of charging sessions per year. This provides a view of how different factors interact to meet the annual energy requirements of an EV:

$$E^{Year} = P_k \times t_k \times n_k \ , \tag{12}$$

[18] There are about 70% travellers who choose private vehicles in VIC involving both self-driving and being passengers [4].

where:

- k denotes the charging rate index.
- E^{Year} denotes the yearly total EV consumption demand.
- D^{Year} is the distance travelled per year.
- β refers to the energy consumption per kilometre.
- E^{L} is the energy consumption demand for a smaller time frame.
- L is the number of specific periods in a year.
- P_k is the charging rate corresponding to the index k.
- t_k is the charging duration corresponding to the index k.
- n_k is the number of charging sessions per year corresponding to the index k.

It is important to note that the charging duration and the number of charging sessions can be adjusted based on different charging rates. For instance, higher charging rates will typically reduce the required charging duration but might increase the number of charging sessions needed to distribute the EV energy load effectively [10].

By understanding the total yearly and period-specific energy demands, we can adjust the charging patterns to reflect actual usage better. For example, if a customer's annual travel distance is known, we can calculate the total energy required for the year. Breaking this down into smaller periods allows us to determine how long the EV needs to be charged in each period to meet these requirements.

Additionally, by rearranging the formula, we can understand how the charging rate and frequency affect the overall energy consumption. This allows for adjustments based on different charging infrastructures and customer behaviours.

Implementing this travel behaviour filter ensures that our analysis is rooted in real-world data, leading to more reliable and actionable insights. It enables us to tailor the charging strategies to match actual driving patterns, improving the efficiency and effectiveness of EV charging profile extraction.

Besides the travel filter, as customers with EV are expected to charge their EVs regularly and relatively frequently, we also introduce a parameter as a threshold to filter out those customers who charge less than commonly expected throughout a year. This parameter is tunable with a smaller number leading to more customers who may charge less frequently to be selected.

5.2.8. Extreme Weather Verification

As each weather station collects daily minimum and maximum temperatures, we first identify extreme hot and cold days in 2022 using the two-sigma principle, which considers approximately 95.45% days as normal weather days and the rest of about 4.55% days as extreme weather days.. We denote the daily minimum and maximum temperatures as t_{min} and t_{max}, respectively.

For extreme hot days at each weather station, we calculate the average maximum temperature and the standard deviation, denoted as u_{max} and σ_{max}, respectively. For those days with maximum temperatures higher than $u_{max} + 2\sigma_{max}$, we consider them as extreme hot days. Similarly, for extreme cold days, we compute the average and standard deviation of minimum temperature, denoted by μ_{min} and σ_{min}, respectively. For those days with minimum temperatures lower than $u_{min} - 2\sigma_{min}$, we consider them as extreme cold days. Our adopted two-sigma principle is adjustable to one-sigma principle (i.e., $u_{max} + \sigma_{max}$ and $u_{min} - \sigma_{min}$) or three-sigma principle (i.e., $u_{max} + 3\sigma_{max}$ and $u_{min} - 3\sigma_{min}$) to identify more or fewer extreme weather days, respectively. The overall flowchart is depicted in Fig. 4.

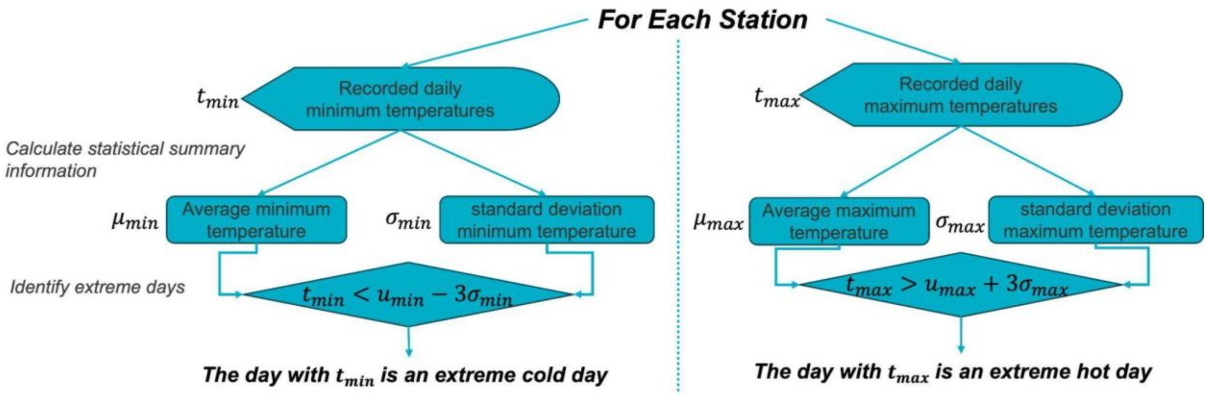

Figure 4 Flowchart of identifying extreme weather days

After identifying the extreme hot and cold days at each weather station, we then match each weather station with each customer using their postcodes. In the way, we can further categorise each customer's disaggregated EV charging profiles through normal weather and extreme weather days. To carefully examine derived EV charging at extreme weather days (e.g.., considering potential impacts of heating or cooling loads), for each customer, we perform the following verification:

1. Customer normal weather days' EV charging to create an envelope representing normal EV charging patterns. Specifically, for each 30-min interval, we collect this interval's EV charging throughout 2022 and calculate its average and standard deviation which are denoted as μ and σ, respectively. Then, three-sigma principle is applied to build the lower and upper bounds for the envelope, which are equal to $\mu - 3\sigma$ and $\mu + 3\sigma$, respectively.
2. We traverse the daily EV charging profiles for all extreme weather days to see whether these EV charging profiles are covered by our constructed envelope or not. If EV charging is within the envelope, the disaggregated EV charging may not be influenced by the extreme weather. Otherwise, we will conduct the following step to further investigate whether the EV charging is affected by the extreme weather conditions.
3. For EV charging profiles that are not fully covered by the envelope, we further identify the part of EV charging that is out of the envelope, including the charging duration and charging periods (e.g., morning, noon, or night times). Based on the average charging duration calculated at normal weather days and corresponding common charging periods, we can then check if the EV charging might be affected by heating or cooling load. If the EV charging does not align with regular patterns, it is then corrected to zero, as the EV charging disaggregation on this specific extreme weather day may be affected by the extreme weather conditions.

It is worth noting that, due to the data limitations, we were only able to classify extreme hot and cold days based on the daily minimum and maximum temperatures. More fine-grained weather dataset, if accessible, could further improve the extreme weather check component to capture extreme weather periods.

6. EV charging disaggregation results

This section presents the EV charging disaggregation results produced by our rule-based model across all DNSPs, including a general overview of the disaggregation from the perspective of different kinds of chargers, analysis on the energy uplift between 2022 and 2020 with respect to EV charging, customers' weekly charging pattern investigation, discussion on the weekday and weekend EV charging behaviours, and verification on extracted EV charging profile on extreme hot and cold days.

6.1. Disaggregation overview

The following analysis provides insights into the dedicated datasets after implementing various data processing steps. After meticulously applying our data filters, we identified the valid residential EV customers without solar PV installation. The total number of identified customers is 265. The identified charger levels are 2.3 kW, 3.7 kW, and 7.4 kW. After data wrangling including the removal of incomplete, inconsistent, and erroneous data, the overall customer number for those without residential PV systems is 493. After the filter described in Section 5.2.5, the remaining number of customers without installing solar PV is 402. This filter relies on the three-sigma principle to remove customers ith inconsistent energy consumption profiles. This three-sigma principle can be relaxed to two-sigma or even one-sigma principles, depending on the context. The three-sigma principle used in our approach aims to ensure that we only exclude customers with significantly inconsistent load patterns. After applying the travel filter, as defined in 5.2.7, the remaining number of customers is 394. We set 5000 kilometres travel distance per year as the threshold to filter out invalid customers. This parameter setting is based on the Ausgrid report [15], which states that 96% of EV drivers tend to drive more than 5000 kilometres per year. This parameter can be adjusted to suit various contexts accordingly. Finally, after applying the charging frequency filter (specifically, we set the threshold for the number of active charging – the number of charging hours as 200), the number of customers drops to 265. Again, this charging frequency parameter is also tunable based on the needs. These levels provide a detailed understanding of the charging rates, showing how different power levels are utilised. This information is crucial for extracting EV charging profiles and for managing and balancing local energy consumption.

Customers' charging behaviours can be influenced by the charger's rate. This variance in charging rates, aligned with the number of customers for each type, distinctly contributes to the overall load on the power grid. For example, Fig. 11 illustrates the distribution of customers by their type of charger. The pie chart shows that 87.5% of the customers use 2.3 kW chargers, 5.3% use 3.7 kW chargers, and 7.2% use 7.4 kW chargers, with the number of customers being 232, 14, and 19, respectively. The average EV charging consumption per day is 10.45 kWh across all non-PV customers.

More importantly, it is worth noticing that, after classifying each customer to a specific type of charger, the majority of customers use 2.3kW charger to charge their EVs at home, while a small number of customers use 3.7kW or 7.4kW chargers. Therefore, we acknowledge that the disaggregated EV charging profiles for customers with these two kinds of chargers may not represent the general EV charging profiles for 3.7kW or 7.4kW charger due to this small number of customers. More customer data associated with 3.7kW and 7.4kW chargers will make the results more robust.

Furthermore, Fig. 12 illustrates the distribution of load differences between 2020 and 2022 based on the original smart meter dataset (i.e., without any operations of our disaggregation approach). The total number of non-PV customers before disaggregation filters, as mentioned above, is 493. The results indicate that most load differences fall into the range between 2.3kW and 3.7kW (not included), which align with the charger mix derived through our disaggregation method that the majority of customers use 2.3kW chargers. Again, we would like to highlight that the charger mix may not reflect the real-world charger mix but is only derived from the provided dataset (with possible inherent bias).

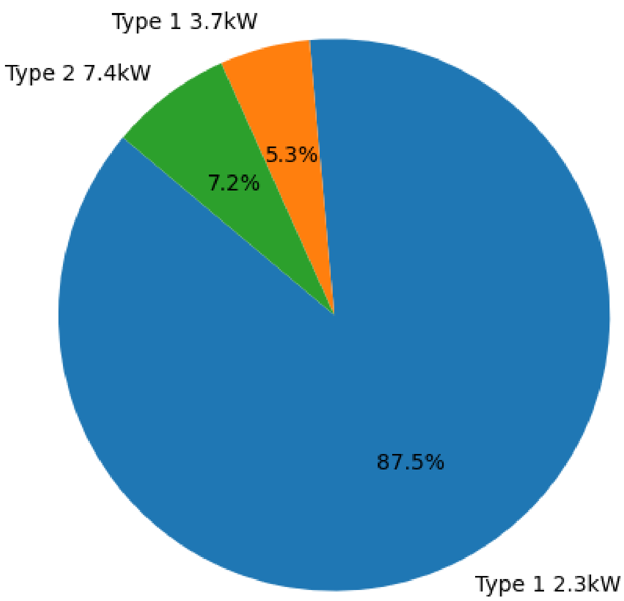

Figure 11 Distribution of EV customers by their using types of chargers in the process dataset (Note that this mix does not necessarily reflect the actual mix in the real world as the processed data can be biased due to limited sample size.)

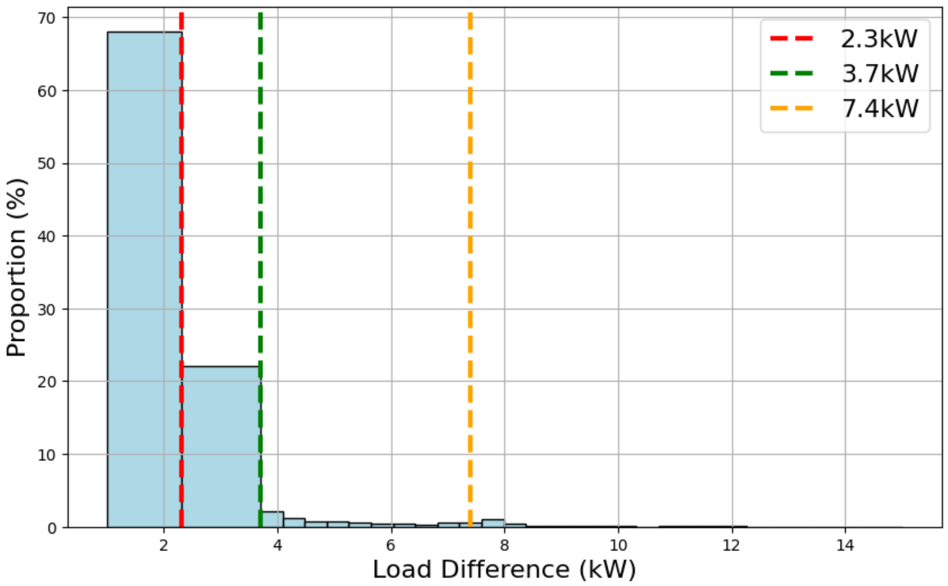

Figure 12 Load difference distribution of the original dataset (the total number of valid non-PV customers in the original dataset is 493)

6.2. Weekly meter load and EV charging visualisations

The following samples illustrate disaggregated EV charging profiles for customers using different types of chargers (with partially masked NMI). These profiles reflect different charging events on a weekly time scale and include customers using different levels of chargers. All the following visualisations start from Monday. Also, the blue and orange lines represent the disaggregated EV charging load and the smart meter load, respectively.

- 2.3kW charger:

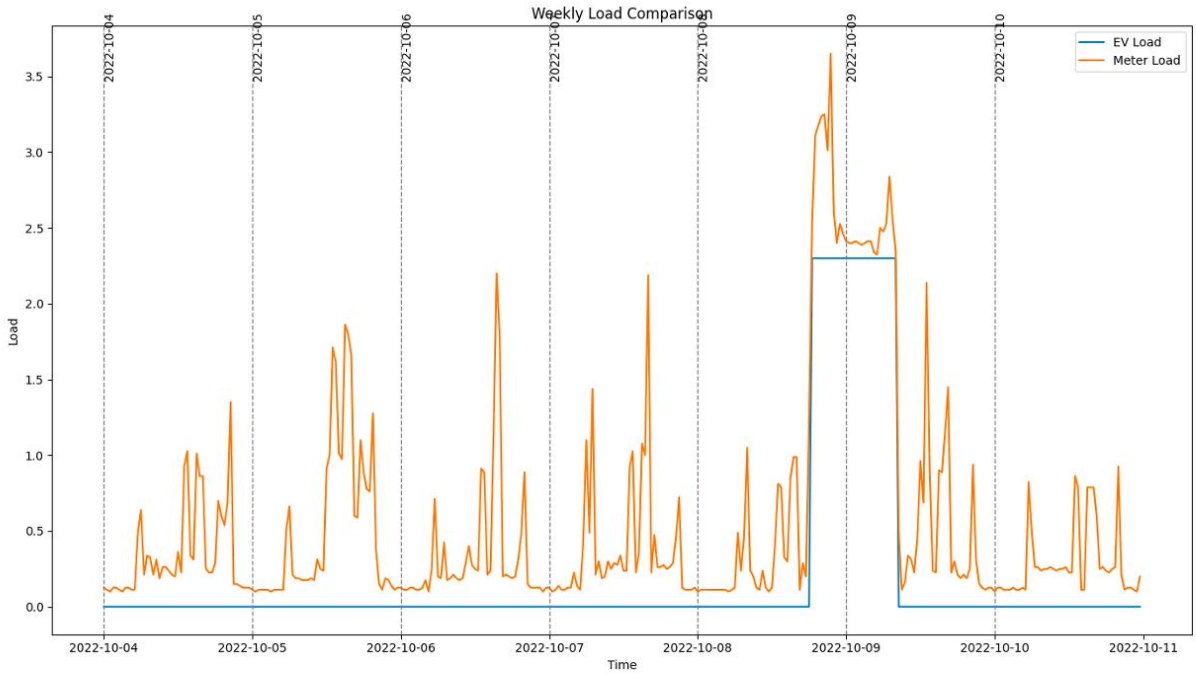

Figure 13 Weekly meter load and EV charging for customer using 2.3kW charger 1exxxxad9 from 2022-10-04 to 2022-10-11

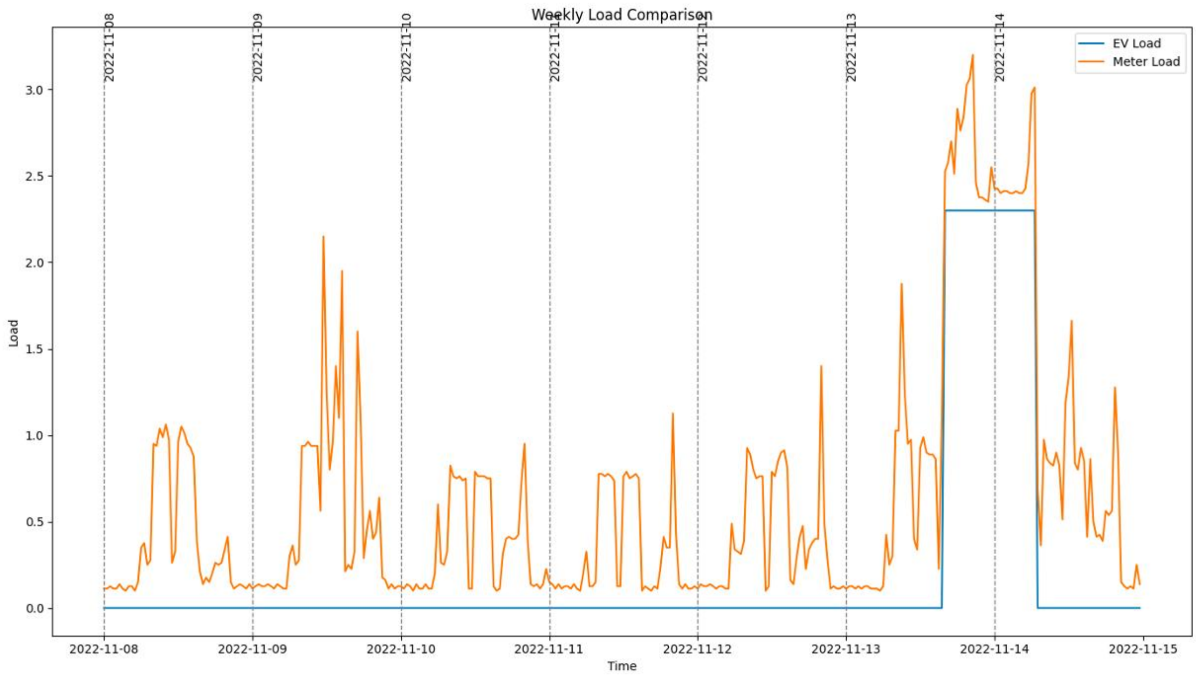

Figure 14 Weekly meter load and EV charging for customer using 2.3kW charger 1exxxad9 from 2022-11-08 to 2022-11-15

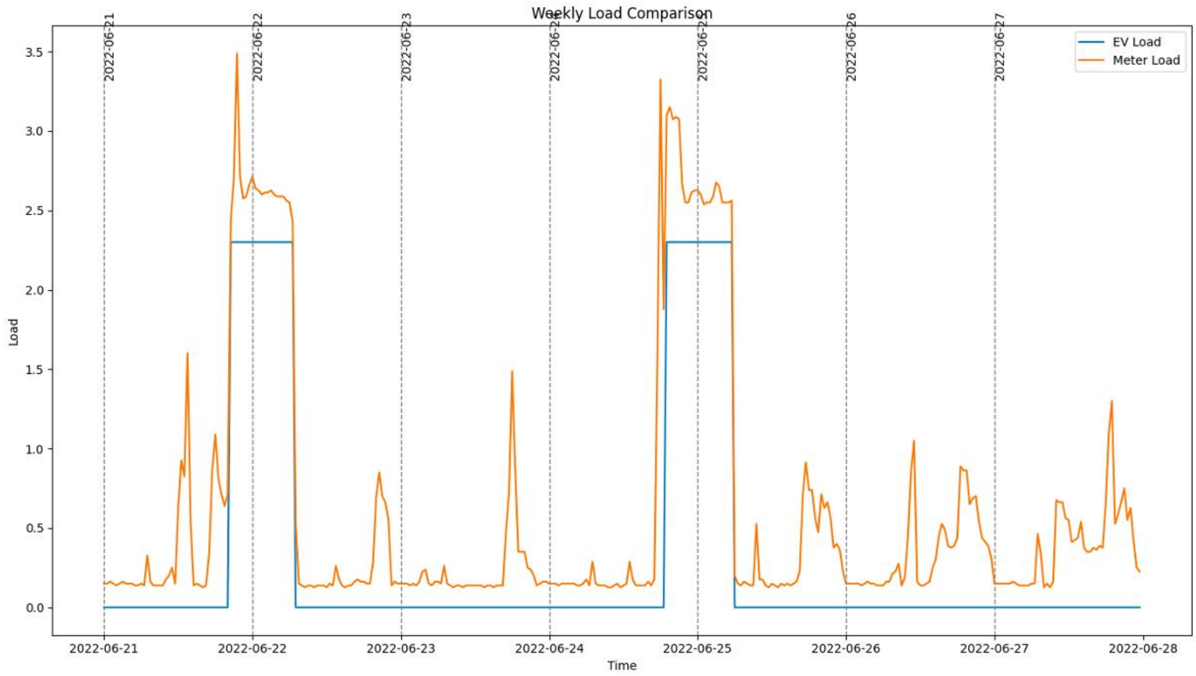

Figure 15 Weekly meter load and EV charging for customer using 2.3kW charger 7f2xxxxe8 from 2022-06-21 to 2022-06-28

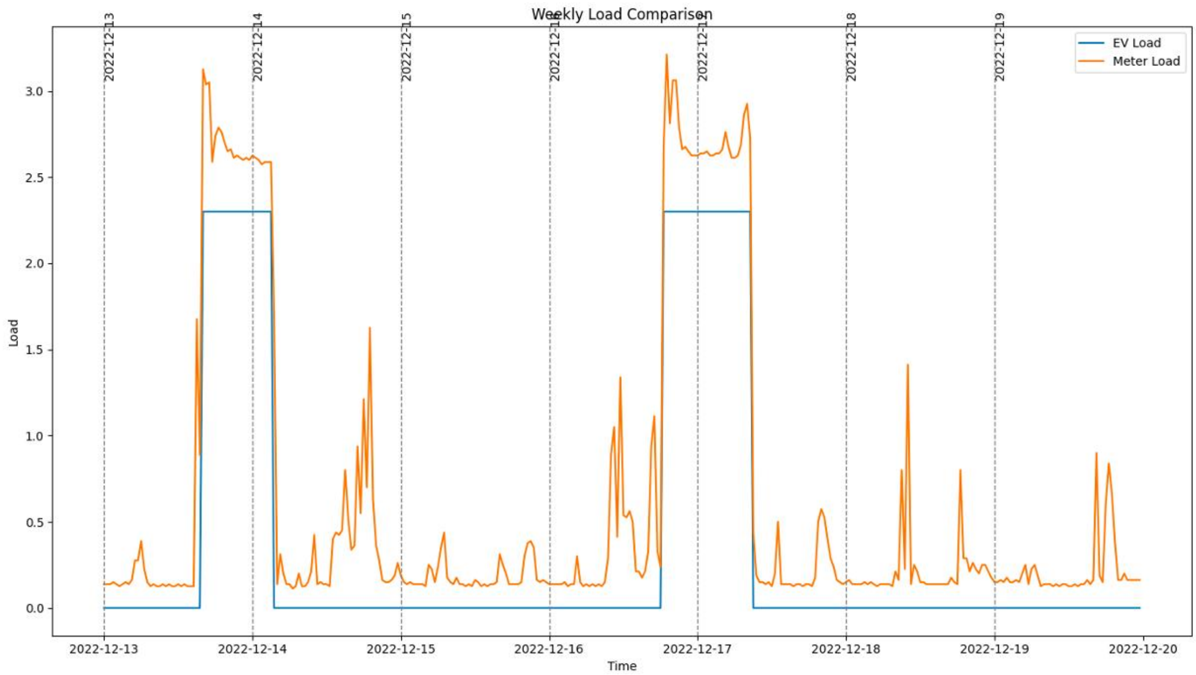

Figure 16 Weekly meter load and EV charging for customer using 2.3kW charger 7fxbe8 from 2022-12-13 to 2022-12-20

- 3.7kW charger

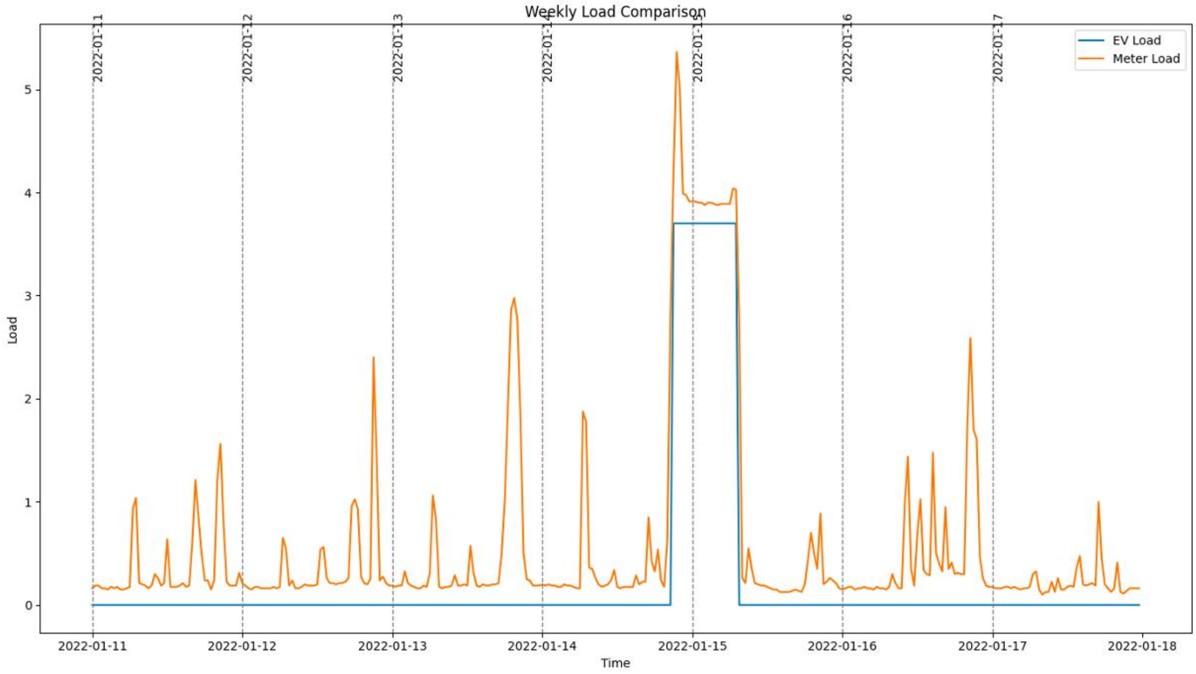

Figure 17 Weekly meter load and EV charging for customer using 3.7kW charger 51xxx9cd52 from 2022-01-11 to 2022-01-18

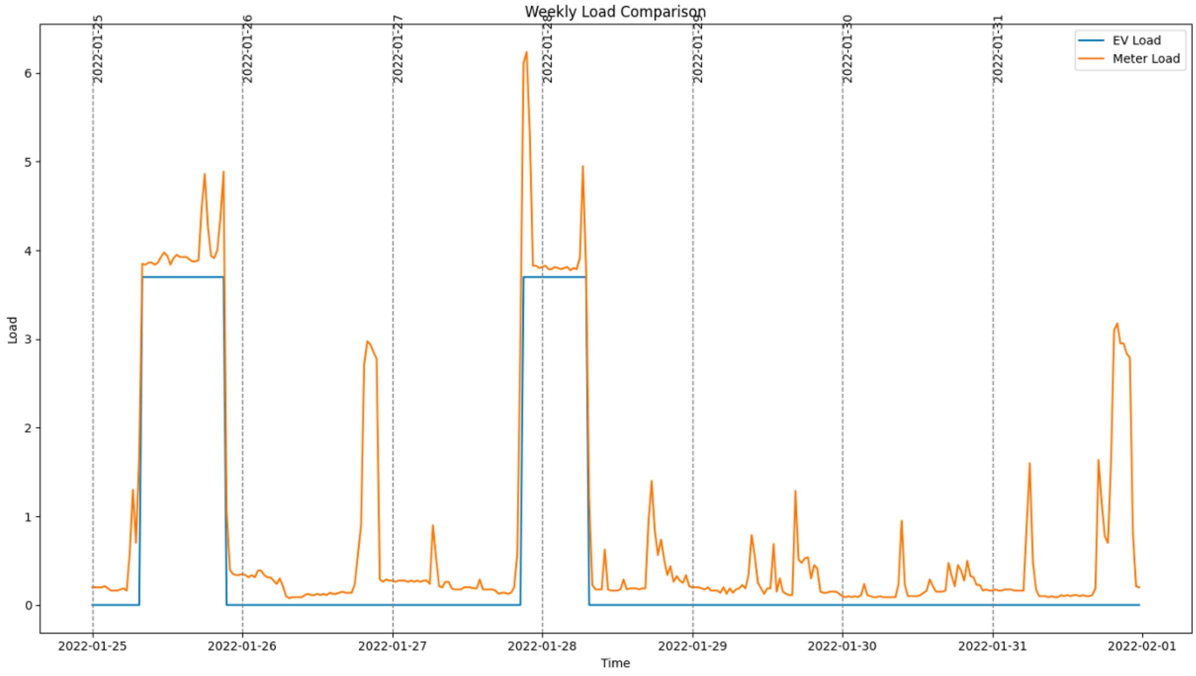

Figure 18 Weekly meter load and EV charging for customer using 3.7kW charger 51x5009cd52 from 2022-01-25 to 2022-02-01

- 7.4kW charger:

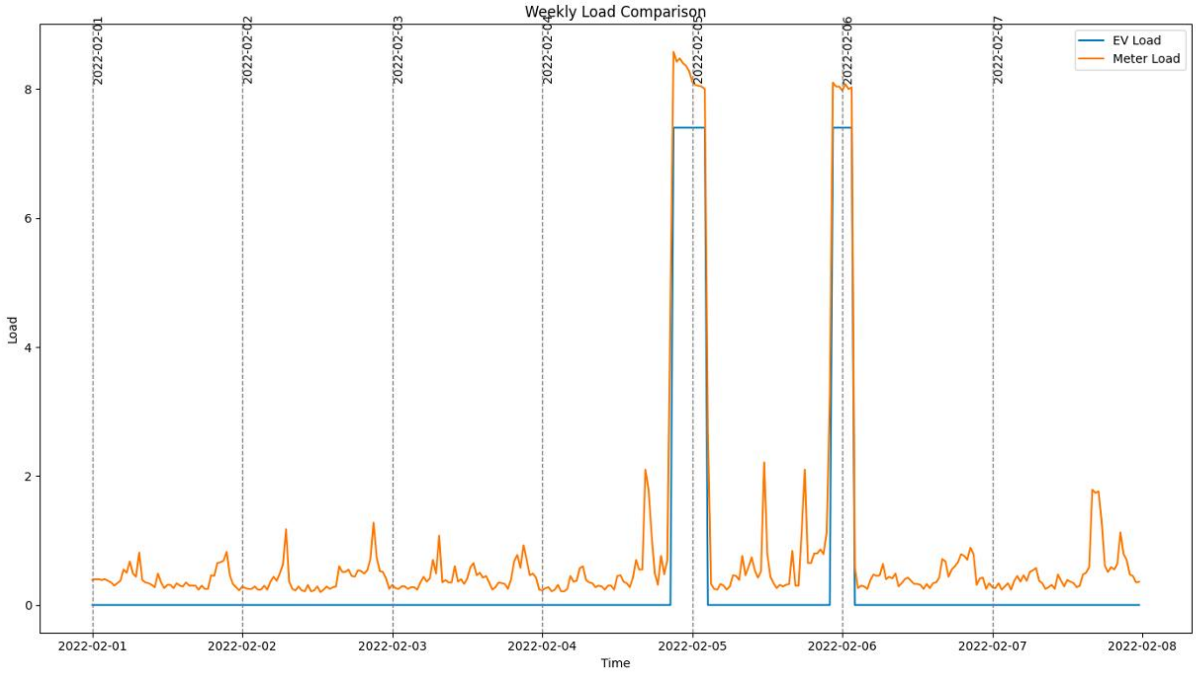

Figure 19 Weekly meter load and EV charging for customer using 7.4kW charger 6axxx8760 from 2022-02-01 to 2022-02-08

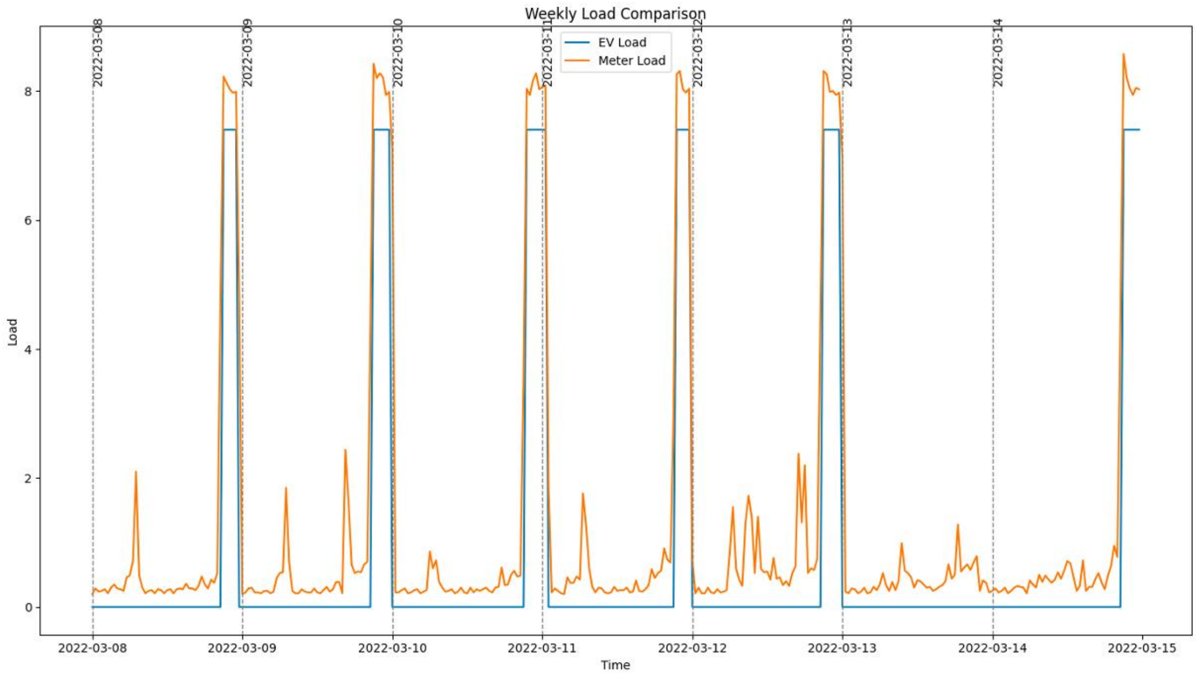

Figure 20 Weekly meter load and EV charging for customer using 7.4kW charger 51xxxx09cd52 from 2022-03-08 to 2022-03-15

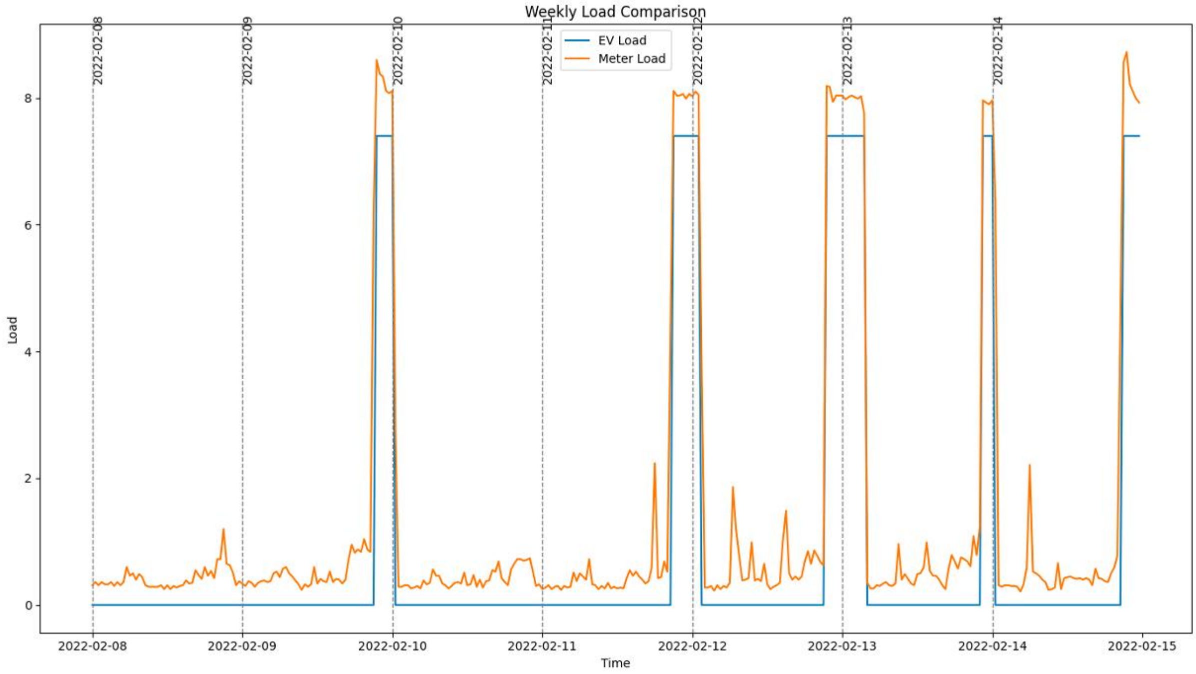

Figure 21 Weekly meter load and EV charging for customer using 7.4kW charger 51xxxxd52 from 2022-02-08 to 2022-02-15

The above illustrations across 2.3kW, 3.7kW, and 7.4kW chargers tell one significant difference in 7.4kW charger. While the former two chargers (i.e., 2.3 and 3.7kW) tend to charge EVs often overnight at most three times per week, customers with 7.4kW chargers tend to perform frequent yet short periods of EV charging. In addition, as the charger size increases from 2.3 kW to 7.4 kW, the EV load contributes more significantly to the overall meter load. For instance, 2.3kW charger has a relatively small impact on the total load, while the 7.4kW charger plays a much larger role, and its charging sessions coincide with the highest peaks in meter load. Moreover, Higher capacity chargers result in more pronounced and consistent peaks in the total load. The 7.4 kW charger creates substantial spikes in the overall load. Furthermore, The visualisations suggest that with larger chargers, the EV load has more regular and higher peaks and the meter load is more synchronised with the EV load for the larger chargers, which also aligns with the results that for high-powered chargers, the proportion of EV charging takes over 90% of the overall energy uplift between 2022 and 2022.

Again, we would like to emphasise that, for disaggregated EV charging profiles using 3.7kW and 7.4kW chargers, due to the small number of corresponding customers, the above weekly meter load and EV charging profiles may be influenced by specific customers' daily electricity usage patterns and may not be able to represent the common patterns for general customers with 3.7kW and 7.4kW chargers.

6.3. Weekday and weekend EV charging analysis

Fig. 22 below shows the daily average EV charging across all customers on weekdays and weekends.

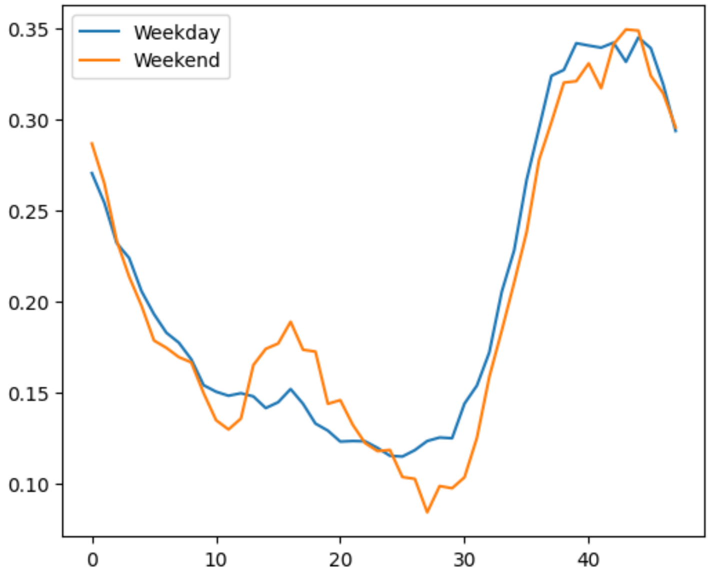

Figure 22 Average EV charging on weekday and weekend

Note that the results for both weekday and weekend align with the study by the Central University of Queensland [14] – most of the EV charging events happen during either morning or night times. Moreover, Fig. 22 reveals that the EV charging at weekends from 8am to 12am are significantly higher than that at weekdays. This is probably because customers can have more opportunities to charge their EVs at home during the beginning of daytime at weekends.

Besides the weekday and weekend comparison, we also calculate the average EV charging energy consumption for each day of week. The results are shown in Table 2, which align with the research findings by the University of Queensland [13]. Specifically, their study shows that EV drivers tend to drive long and short (mainly for commuting) distances at weekends and weekdays, respectively. Our EV charging consumption results indicate that customers tend to charge more at Fridays and may get prepared for long-distance driving at weekends. Customers also tend to charge more at early weekdays, e.g., Tuesday. This might be because their EVs may run out of battery capacity during weekend long-distance travel and need to get charged for daily commuting.

Table 2 Average EV charging energy consumption for each day of week

Day of Week	Average EV Energy Consumption
Monday	10.25 kWh
Tuesday	10.53kWh
Wednesday	10.56kWh
Thursday	10.33kWh
Friday	10.67kWh
Saturday	9.73kWh
Sunday	8.88kWh

7. Validation on disaggregated EV charging profiles

As we do not have the ground truth knowledge on customers' EV charging behaviours, to assess and validate our disaggregated EV charging results, in this section, we examine our results using two validation approaches. The first one is the EV charging proportion that can explain the energy uplift from 2020 to 2022. Based on our assumption, for a single customer, we assume that there is no significant change in the energy consumption of other electrical appliances (excluding EV) for the same date and time across three continuous years. Note that this is the best-effort assumption we could make due to the lack of necessary information to relax it. Therefore, ideally, the energy uplift should mostly be related to EV charging. Therefore, the first validation is to examine whether our disaggregated EV charging profiles can explain the major energy uplift between 2022 and 2020. For the second validation, we rely on comparisons between time-series load and extracted EV charging profiles over a week from Monday to Sunday to check if the results are reasonable. In addition, we delve into the extreme cold and extreme hot days to check that our disaggregation model is not influenced by these extreme weather conditions.

7.1. Validation 1: Energy uplift between 2022 and 2020 w.r.t. EV charging

In Fig. 23, the blue line illustrates the meter load difference between 2022 and 2020 (i.e., using 2022's meter load subtracting 2020's meter load), and the orange line depicts our disaggregated EV charging load (in kW).

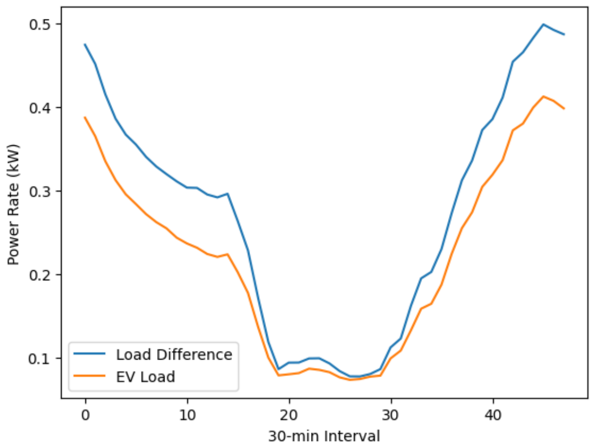

Figure 23 Illustration of meter load difference and disaggregated EV charging load

Each line is averaged across all customers (i.e., not specifying charger types) throughout 2022. From Fig. 12, we can observe that the energy uplift between 2022 and 2020 is mainly due to the EV charging, which is reasonable as we have assumed the stability in electrical appliance energy consumption in Section 4 – Key assumptions. The detailed energy uplift proportions caused by EV charging for three types of chargers are presented in Table 3 below.

Table 3 Energy uplift proportion caused by EV charging for three types of charger

Charger Type	Customer Number	EV Charging Proportion in Energy Uplift
2.3kW	232	74.51%

3.7kW	14	80.10%
7.4kW	19	90.96%

.7.2. Validation 2: EV charging under extreme weather days

To examine the effectiveness of our disaggregation under extreme weather conditions, we here provide samples with meter loads and EV charging through one week including extreme weather days. For instance, in Fig. 24, 2022-12-23 is an extremely hot day as illustrated in the red line (i.e., the daily maximum temperature). Our disaggregation does not consider the dual spikes during daytime as EV charging for the following reasons: 1) these spikes at extreme hot days do not align with this customer's regular EV charging patterns – usually at nights; and 2) the short charging durations for 2.3kW charger does not look like EV charging. Given the hot temperatures, the spikes shown in the meter load is more likely to be caused by the cooling load.

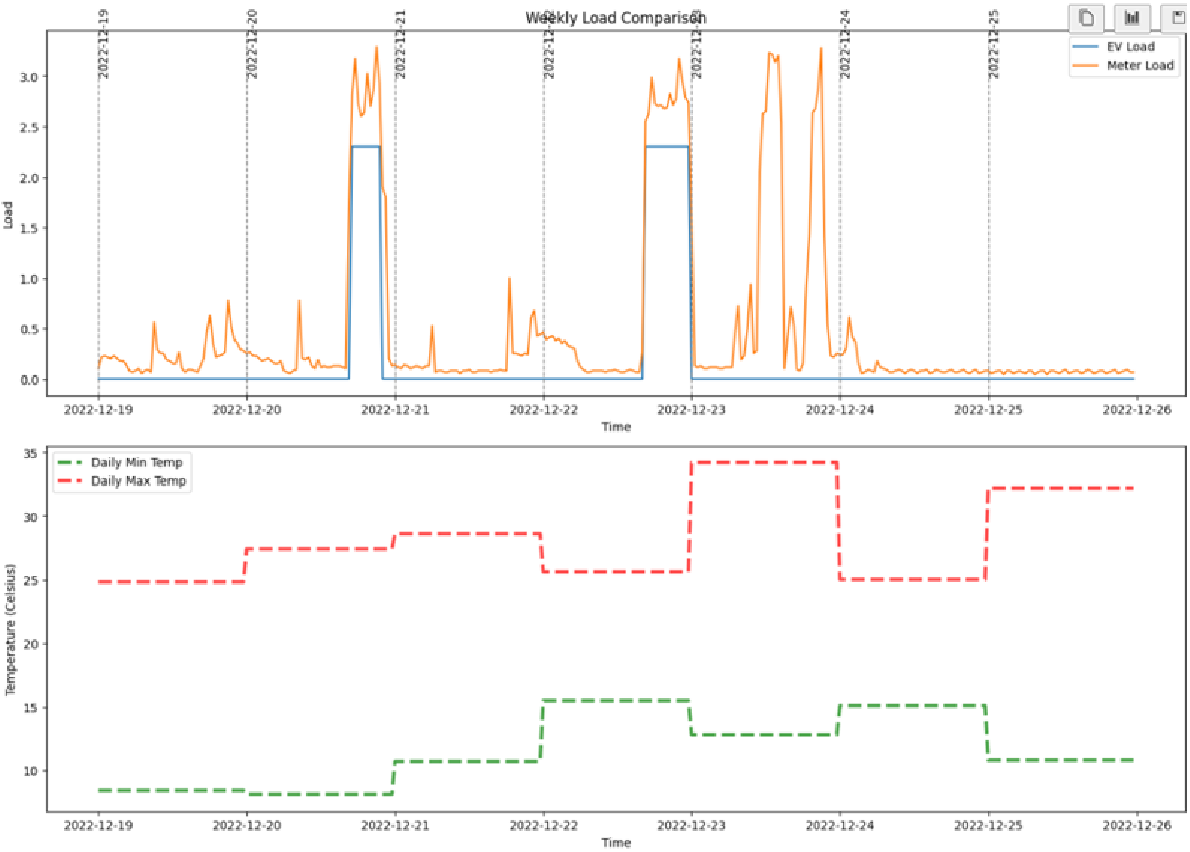

Figure 24 Meter loads and EV charging from 2022-12-19 to 2022-12-26 (2022-12-23 is extreme hot day)

Similarly, in Fig. 25, 2022-01-24 is also an extremely hot day. Besides the identified EV charging, our disaggregation model does not consider the additional load as EV charging. Given the hot temperature on that day, this additional load during daytime may also be due to the cooling load.

The above two examples demonstrate that 1) our disaggregation model will not consider the additional load as EV charging, in particular at extreme weather days and 2) our disaggregation carefully checks the meter load at extreme weather days to avoid mistakenly extracting EV charging load which may be caused by heating or cooling load.

It is noteworthy that identifying the impacts of extreme weathers on EV charging load is challenging, since each customer may have different EV charging behaviours, daily load profiles, and various power rates of heating and cooling appliances.

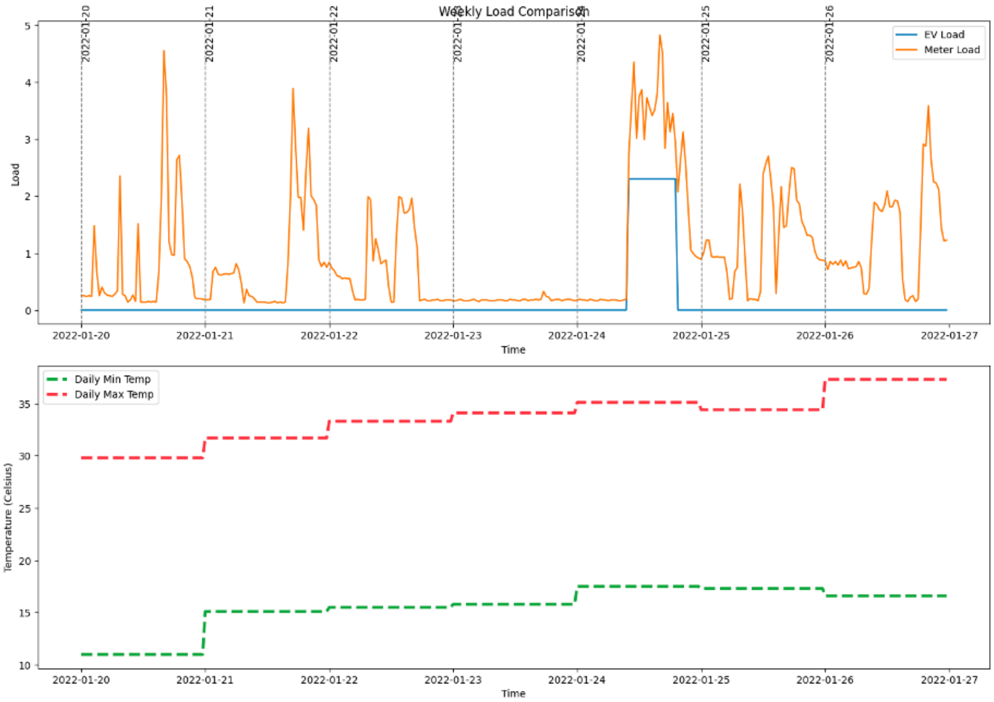

Figure 25 Meter loads and EV charging from 2022-12-19 to 2022-12-26 (2022-12-23 is extreme hot day)

7.3. Validation 3: Examining the impacts of other loads (possible kettle load and heat pump consumption) on our EV charging extraction method and results

In addition to checking the disaggregated EV charging profiles under extreme weather conditions, we further examine if other loads (possible kettle load and heat pump consumption) that exhibit power increases/jumps would affect our EV charging extraction method and results.

As mentioned in Section 5.2.3, in our disaggregation toolbox, we have filtered out EV charging events which do not last for at least one hour. Therefore, our extracted EV charging profile will not capture short-duration power increases, for example those caused by the kettle use usually lasting for 5 minutes. Regarding the heat pump energy consumption, as our disaggregation approach contrasts day to day load in 2020 and 2022, we assume that hot water consumption patterns remain relatively stable from 2020 to 2022. The most significant change in power usage, particularly the increase in 2022 compared to 2020, is attributed to EV charging. From the Vic Government Rebate dataset, the customers included in our study all applied for EV rebates in 2021, between the two periods in which their meter data are used in our method. In other words, In Fig. 26, we take a 3.7kW charger customer as an example. The blue dashed rectangle shows a short-duration load increase/jump (with a power consumption of ~2kW) lasting only in one 30-min interval, and our

EV charging profile is not affected by such a sudden power increase. In addition to the above 2kW load increase, there are many other small load increases of over 1kW in the week, but none of them are taken as possible EV charging profiles by our method.

In Fig. 27, we take a 2.3kW charger customer as an example. In the meter readings, we observe that there exist cyclic and periodic load patterns, which could contain heat pump consumption if the customers have one. Our method does not take any of those cyclic and periodic load patterns as possible EV charging, even not at the low charging power, i.e., 2.3kW. This example, to an extent, shows that our disaggregated results do not count in cyclic and periodic loads, including possible heat pump loads, as EV charging.

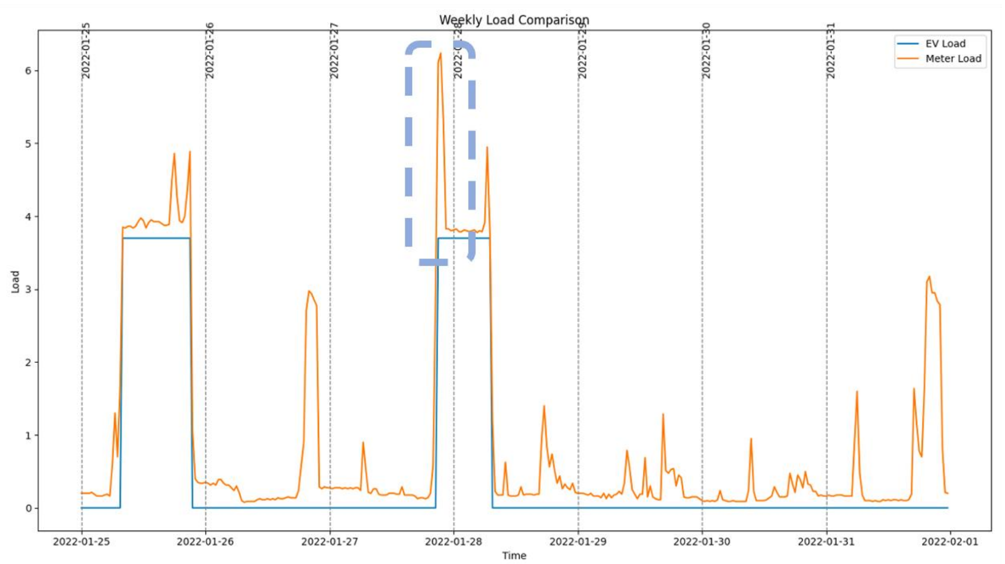

Figure 26 Other short-duration load increases/jumps not captured as EV charging by our disaggregation model - 3.7kW charger customer as an example.

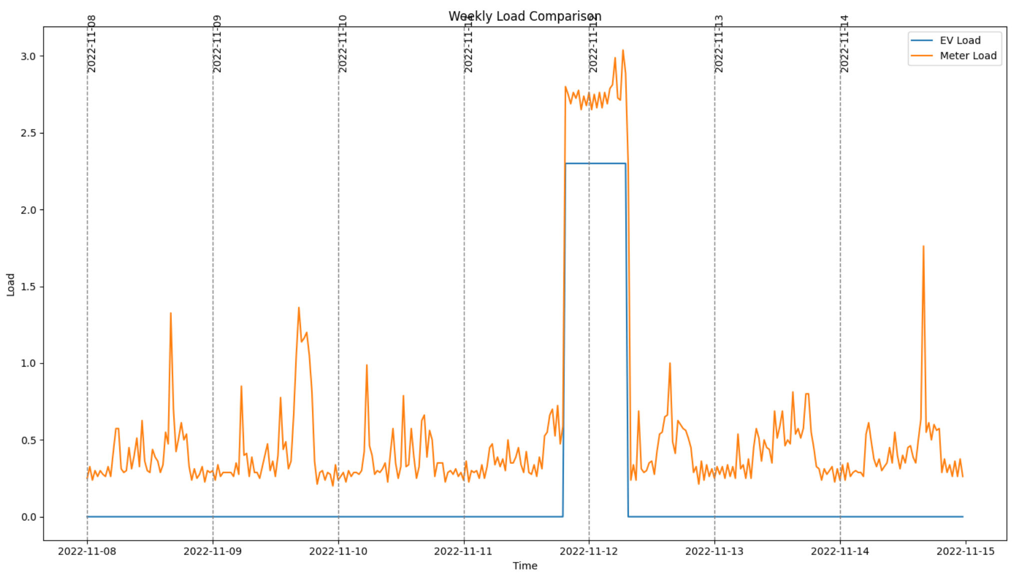

Figure 27 Other cyclic and periodic loads not captured as EV charging by our disaggregation model - 2.3kW charger customer as an example.

7.4. Validation 4: Investigation of charger mix in the provided dataset

To validate that the charger mix of non-PV customers after our disaggregation method reflects the original mix of the provided dataset, for the raw meter load data without any data processing proposed in our method, we calculate the number of hours (per week per customer) when the load difference is between 3.7kW and 7.4kW (where 7.4kW is not included).

The result is depicted in the Fig. 28. The X-axis and Y-axis represent the number of hours and the number of customers, respectively. The values on each bar represent the number of customers of certain hours. The green dashed line indicates how many hours is required to reach the ideal charging consumption per week if the customer is equipped with a 3.7kW charger. Note that the real-world EV charging consumption is generally around 10kWh, and the exact number in our analysis is 10.45kWh per day. From the figure, we can obviously find that only 12 customers satisfied the weekly EV charging consumption threshold. After relaxing the threshold by 40%, about 15 customers possibly have installed 3.7kW chargers, aligning with our disaggregation results that the number of non-PV customers with 3.7kW chargers is 14.

Similarly, we calculate the number of hours (per week per customer) when load difference is bigger than or equal to 7.4kW, which is still based on the raw input without any preprocessing. The results in Fig. 29 show that there are about 20 customers who may install 7.4kW chargers when we relax the threshold by 50%, aligning with our disaggregation results that the number of non-PV customers with 7.4kW chargers is 19.

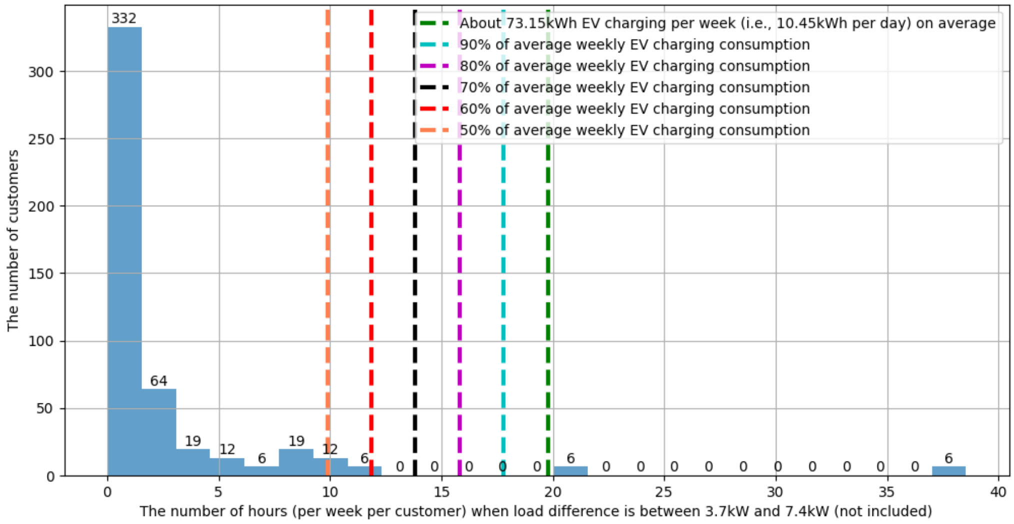

Figure 28 The number of hours (per week per customer) when load difference is between 3.7kW and 7.4kW

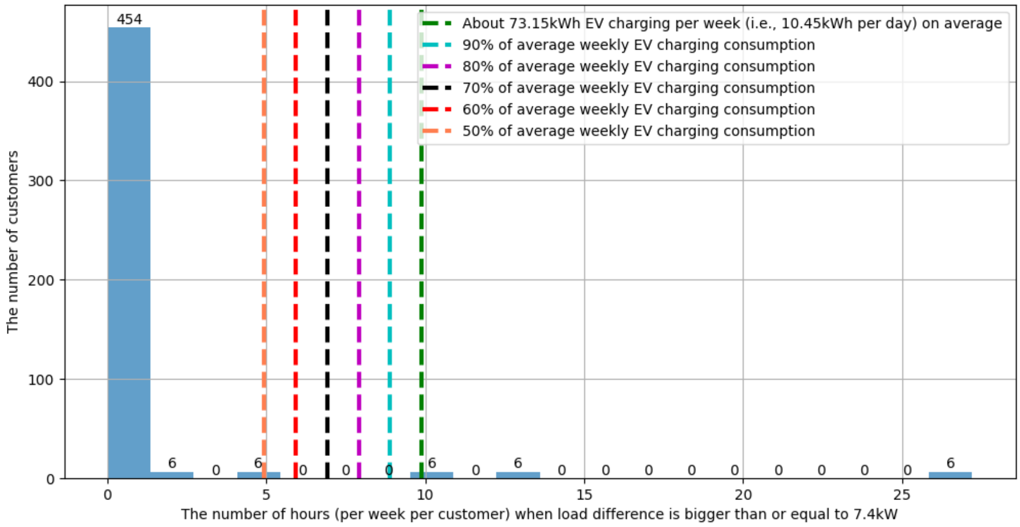

Figure 29 The number of hours (per week per customer) when load difference is bigger than or equal to 7.4kW

Conclusion

This milestone report outlines a framework for the technical modelling of electrification of transport profiles, particularly focusing on the extraction of EV charging profiles from smart meter data for customers without installing solar PV, as these customers are not affected by the solar PV generation that may lead to degraded performance of the method. Thanks to the meter data that is made available for this project, we are able to design the toolbox based on a contrastive approach based on the changes in the meter reading between two years. Given the available datasets, our approach introduces a set of filtering mechanisms to ensure reliable and accurate analysis. In addition, we must acknowledge that disaggregating EV charging profiles without ground truth data is a challenging task. The accuracy of the extracted profiles can be influenced by several factors, including but are not limited to COVID-19 and individual customers' specific devices, such as battery storage, and behaviours that are not easily visible from the meter data we used.

It is worth noting that, since we do not have ground truth, we cannot assess the accuracy of the EV data extracted through the toolbox. To address the challenge, we designed the toolbox to filter out signals that are not likely to be EV charging as much as possible, such as kettle and heat pump loads. If additional data could be made available in the future, users can potentially further improve the disaggregation results. Additionally, our toolbox does not focus on discrete irregular charging, as these signals have short durations, and we cannot ensure they are EV charging signals.

The core concept of our approach is the differentiation of energy consumption to identify EV charging loads. By comparing smart meter data from periods without EV usage to periods with EV usage, we distinguish the additional energy consumption attributable to EV charging. This process involved the careful initialization of smart meter data, cross-referencing with EV rebate programs, the application of travel behaviour filters to align charging patterns with realistic driving habits, and the check of EV charging profiles at extreme weather days.

Our data processing results underscore a critical insight: while our toolbox can extract potential EV charging profiles for all EV customers, there is a trade-off between extraction accuracy and customer group size. Customers of the toolbox must balance the need for accuracy with the desire to include a larger customer base. The flexibility of our toolbox allows customers to manually control which filtering steps to apply, enabling them to tailor the analysis to their specific requirements.
In addition to refining the dataset, we have identified the EV charging rates for customers following common charger types. The output includes the charger levels used by customers to charge their EVs at home. These charger levels—2.3 kW, 3.7 kW, and 7.4 kW—provide a detailed understanding of the charging rates, illustrating how different power levels are utilised. This information is crucial for extracting EV charging profiles and for managing and balancing local energy consumption.

Furthermore, we obtained preliminary distributions of customers across different charger types and conducted a detailed analysis of charging behaviour throughout the week. This analysis provided insights into the daily and weekly usage patterns. Additionally, we provided sample EV charging curves, which serve as practical examples of how our methodology can be applied to real-world data. The further analysis on weekdays and weekends' EV charging profiles provide insights into customers' EV charging behaviours.

Through our approach, we have established a framework that not only refines EV charging data but also provides insights into charging behaviours. The identified charging rates and patterns are essential for future EV modelling efforts, load balancing strategies, and energy management. These insights can potentially support the development of a more sustainable and efficient energy system, for energy management, infrastructure planning and operation in the evolving landscape of EV adoption.

MONASH University

References

[1] Department of Energy, Environment, and Climate Action, “Zero emissions vehicle roadmap”, 2021.
[2] Department of Transport and Planning, “Victorian Integrated Survey of Travel and Activity”, 2022.
[3] Australia Electric Vehicle Council, “State of Electric Vehicles Report 2023”, 2023.
[4] Department of Energy, Environment, and Climate Action, “EV-ready buildings for owners corporations”, 2023.
[5] Evenergi, “EV Owner Demographics and Behaviours - Results of EVenergi’s relevant surveys”, 2023.
[6] P. Lavieri and C.B. Domenech, “Large-Scale Network and System Integration of Electric Vehicles: A Techno-Economic Perspective”, 2021.
[7] Department of Energy, Environment, and Climate Action, “Find your energy distributor," [Online]. Available:https://www.energy.vic.gov.au/households/find-your-energy-distributor.
[8] Frontier Economics, “Residential energy consumption benchmarks-Final report for the Australian Energy Regulator”. [Online] Available:https://www.aer.gov.au/system/files/Residential%20energy%20consumption%20benchmarks%20-%209%20December%202020_0.pdf.
[9] K. P. Sinaga and M. -S. Yang, "Unsupervised K-Means Clustering Algorithm," *IEEE Access*, vol. 8, pp. 80716-80727, 2020.
[10] J.Dixon & K. Bell, “Electric vehicles: Battery capacity, charger power, access to charging and the impacts on distribution networks”. *ETransportation*, 4, 100059, 2020.
[11] Bureau of Meteorology. Available: http://www.bom.gov.au/.
[12] DATA VIC, “Vicmap Admin - Postcode Polygon,” [Online]. Available: https://discover.data.vic.gov.au/dataset/vicmap-admin-postcode-polygon.
[13] Philip, Thara, Lim, Kai Li, and Whitehead, Jake (2022). Driving and charging an EV in Australia: A real-world analysis. Australasian Transport Research Forum (ATRF), Adelaide, SA, Australia, 28-30 September 2022.
[14] Mumtahina, Umme, Sanath Alahakoon, Peter Wolfs, and Jiannan Liu. 2024. "Constructing Australian Residential Electricity Load Profile for Supporting Future Network Studies" Energies 17, no. 12: 2908. https://doi.org/10.3390/en17122908
[15] Ausgrid - Revised proposal - EVenergi - Att. 5.7.5 - EV owner demographics and behaviours - 30 Nov 2023. https://www.aer.gov.au/documents/ausgrid-revised-proposal-evenergi-att-575-ev-owner-demographics-and-behaviours-30-nov-2023.

Faculty of XXXXX
Monash University
Wellington Road
Clayton, Victoria 3800
Australia

T: +61 3 XXXX XXXX