Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Business & Technology

How available are charging stations really? Revealing hidden insights from transactional charging data

Thomas Laureys

0568068

Academic year 2023-2024

Promotor: Geoffrey AERTS

Social Sciences & Solvay Business School

DECLARATION OF ORIGINALITY AND PUBLICATION

The signed declaration of originality and publication is an integral component of the written work submitted by the student.

With my signature I confirm that:

- I am the sole author of the written work here enclosed[1];
- I have compiled the written work in my own words;
- I have not committed plagiarism as defined in article 118 of the Teaching and examination regulations of the VUB; of which the most common forms of plagiarism are (non-exhaustive list):
    - form 1: to copy paragraphs/sentences of other authors with a reference, but without the use of appropriate quotation marks when needed;
    - form 2: to copy paragraphs/sentences of other authors, whether literally or not, without any sources;
    - form 3: to refer to primary sources where both text and references, whether literally or not, are copied from a non-referenced secondary source;
    - form 4: to copy paragraphs/sentences of other authors, whether with a reference or not, with only minimal or misleading adaptions;
    - form 5: to copy own previous work without proper citation (self-plagiarism);
    - form 6: to generate or substantially modify text fragments with software tools, without explicit mention of this.
- I have referenced fully in the text and in the reference list any internet sources, published or unpublished works from which I have quoted or drawn reference;
- I have indicated clearly all text from which I have quoted literally;
- I have documented all methods, data and processes truthfully;
- I have not manipulated any data;
- I have mentioned all persons and organisations who were significant facilitators of the work, thus all the work submitted for assessment is my own unaided work except where I have explicitly stated otherwise;
- this work nor part of it has been submitted for a degree to any other programme, university or institution;
- I am aware that the work will be screened for plagiarism;
- I will submit all original research material immediately at the Dean's office when it is requested;
- I am aware that it is my responsibility to check when I am called for a hearing and to be available during the hearing period;
- I am aware of the disciplinary sanctions in case of irregularities as described in article 118 of the Teaching and examination regulations of the VUB;
- (Only for master theses) I agree to the publication of my master thesis. If not, I will attach the 'consent form for publicising master thesis' together with my master thesis (to be found on the VUB website).

Student's surname, first name: Laureys Thomas **Date:** 20/05/2024

Signature: Laureys

[1] For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisor, Geoffrey Aerts, for making the research in this master thesis possible and for the very useful feedback during feedback moments. Next to this, I would also like to express my special thanks to Simon Weekx, who was more involved in the weekly feedback of my master thesis, for giving me feedback on my draft versions of the master thesis and for brainstorming together about determining the scope of the research.

Besides this, I would like to thank my family for the continuous support throughout my studies, which allowed me to fully concentrate on my studies. Furthermore, I extend my thanks towards my fellow students, and especially my girlfriend, with whom I could regularly discuss my master thesis and from whom I also received some useful feedback. Finally, I would also like to express my gratitude towards the professors and the assistants who have taught me the necessary knowledge over the past five years, which now allows me to complete this master thesis.

Thomas Laureys
Etterbeek, 20th of May 2024

Abstract

This master thesis presents a framework that allows to identify availability constraints by analysing transactional charging data, which provides us with insights about the reliability and the functionality of the existing charging infrastructure. The framework consists of 3 steps. In the first step, we analyse the charging behaviour at individual EVSEs. This is then broadened in the second and in the third step, in which we respectively analyse the charging behaviour at adjacent EVSEs at the same CS and at neighbouring CSs. Applying the developed framework to a publicly available dataset for the city of Munich, we find an average percentage of availability constraints of 5,45%, leading to a total downtime of almost 300 000 hours. Next to these insights, this framework also allows to enrich datasets containing transactional charging data. This enables researchers that use these datasets to develop their own models to refine their models, which will lead to more correct conclusions. This master thesis shows thus that, by combining the analyses of the charging behaviour at the CI itself and at the neighbouring CI, it is possible to gain information about periods of potential availability constraints, which allows to quantify these periods and to clean the dataset for these periods, which will positively impact the precisions of models making use of datasets containing transactional charging data. Therefore, this research contributes to the understanding of the reliability of the existing literature, which is an often overlooked topic in the research about increasing charging possibilities.

Table of Contents

1.) Introduction

In 2015, 196 countries signed the Paris Agreement, in which they aim to keep the global warming up well below 2% compared to the pre-industrial temperature level in the late 1800s (United Nations, n.d.). Preferably, they want to limit the global warming up to 1.5%. In order to be able to achieve this goal, greenhouse gas emissions need to be reduced by 45% by 2030 and eventually need to reach zero by 2050 (United Nations, n.d.). The Paris Agreement required countries to submit action plans in which they must outline their strategy in order to reach the objectives set by the Agreement. In the European Union's action plan, they aim to reduce the EU emission by 55% before 2030, compared to the 1990 level (European Council, 2023). Although the EU already reduced its overall greenhouse gas emission by 22.4% in 2016 (compared to the level of 1990), it seems that the transportation sector has some difficulties to catch up with this trend, as the emission of greenhouse gases in the transportation sector increased by 18% in this same time span (Ahn, 2019).

In addition, as the transportation sector is one of the main contributors to the climate change (being responsible for 23% of the total amount of energy-related greenhouse gas emissions in 2017), this increase threatens the achievement of the goal of the reduction in greenhouse gas emission (Lévay et al., 2017). Therefore, as part of their long-term emission reduction strategy, the European Parliament recently approved a new legislation on CO2 emission of new cars sold in the European Union. By 2030, they demand a 55% reduction of CO2 emission for cars and by 2035, they want a reduction of 100% (European Parliament, 2023). This means that car manufacturers will have to look for alternatives for internal combustion engine vehicles.

As research has shown in the past, electric vehicles are a sustainable alternative to substantially reduce and even eliminate the emission of greenhouse gasses (Delucchi et al., 2014; van der Kam et al., 2020). As a result, there has been already a lot of support by European governments to increase the penetration of electric vehicles in their country (Razmjoo et al., 2022). This increase in the number of electric vehicles on the roads in the European Union (Statista, n.d.) also requires an increase in the availability of public charging infrastructure (CI), in order to not slow down this penetration of EVs (Salah & Kama, 2016; Straka et al., 2020). However, in order to convince charging point operators (CPOs) to deploy public CI, they have to be ensured that their installation of this CI will be profitable enough and therefore a sufficient utilization of the CI is required (Friese et al., 2021; Mortimer et al., 2022a; Straka et al. 2020). For policy makers too it is important that the CI they install in public space has a sufficient utilization in order to make this placement justifiable (Friese et al., 2021). On the one hand, as the location of the CI is a very important determinant for its usage and thus also for its profitability (Friese et al., 2021; Straka et al., 2020), this should be carefully considered by the CPOs and the policy makers when deciding upon the installation of new CI (Mortimer et al., 2022b). However, as already a lot of research has been conducted on this quantitative aspect of the CI (i.e. increasing the availability of CI by expanding the CI network) (Karanam & Tal, 2023), this master thesis will not focus on this topic. But on the other hand, the availability of the existing CI also determines the utilization rate of the CI and thus again its profitability (Karanam & Tal, 2023). As stated by Karanam & Tal (2023), this is an aspect of CI that has been less researched upon in the past years. Therefore, this master thesis will be about the availability of the existing CI. More specifically, in this master thesis, we will answer the following research question, which is further divided into 3 sub questions (each corresponding to one step of the framework).

- What can the analysis of charging behaviour at Electric Vehicle Supply Equipment (EVSE) tell us about their availability?

  Availability means that the EVSE is charging an EV or is available for charging an EV.

  - How can we identify suspicious periods in availability by only looking at the charging behaviour of the EVSE itself?
  - What can the analysis of the charging behaviour at adjacent EVSEs at the same Charging Station (CS) add to our understanding of what happens during these periods?
  - What can the analysis of the charging behaviour at nearby CSs add to our understanding of what happens during these suspicious periods?

In the remainder of this master thesis, we will start by providing an overview of the existing literature about increasing the availability of CI (i.e. why this increase is needed and what has already been investigated in the existing literature), also emphasizing the topics that are less researched upon so far (i.e. the research gap). Then, we will explain the methodology that will be used to develop the framework. However, before developing the framework, we will first go further into detail on the data that is used to develop the framework, after which we will then discuss the framework itself in detail. Once discussed this, we will (in section 7) discuss the results of our framework applied on the publicly available dataset about CEs in Munich between 06/05/2021 and 06/10/2021, which we will then (in section 8) reflect on the existing literature. Finally, we will make a critical reflection about our framework, in which we will discuss its strengths and its limitations as well as potential future research. We close with a conclusion in which we summarize our main findings.

A list of abbreviations can be found at the end of this master thesis after the conclusion. The code developed in this framework is available upon request.

2.) State of the Art

The following sections will discuss the relevant literature concerning the utilization of CI. The first part, '*Penetration of Electric Vehicles*', describes the most important factors influencing the demand of EVs. As the presence of sufficient available CI is an important determinant for the demand of EVs and thus for its level of penetration (Egbue & Long, 2012), the following part, *'Availability of Charging Infrastructure'*, then dives into more detail about this, discussing the 2 aspects of availability of CI, namely the *'Quantitative aspect'* and the '*Qualitative* aspect'. The last section provides a brief overview of the research gap that exists in the current literature and which will thus be the focus of this research.

Note that a part of this literature review is based on the research paper written in the academic year 2022-2023 for the course *Research Paper*.

2.1) Penetration of Electric Vehicles

In recent years, European governments tried to encourage people to switch from ICE vehicles to electric vehicles (EVs) (Razmjoo et al., 2022). However, in order to be able to increase the public demand for electric vehicles, there should be first of all a good understanding of which factors influence this demand. A first factor influencing the demand for EVs is the related cost of owning an EV (compared to owning an ICE vehicle). The high purchase price of an EV, compared to the purchase price of an ICE vehicle, forms a barrier to the penetration of EVs. In order to reduce this additional cost, governments introduce fiscal incentives, such as tax reductions or direct subsidies, in order to make EVs more attractive to people (Dost et al., 2016; Lévay et al., 2017). A second factor influencing the demand for EVs is the presence of sufficient reliable charging infrastructure. So, in other words, the ease of finding an available, good working charging point for charging your EV (Egbue & Long, 2012). As EVs have still a limited range compared to ICE vehicles, EV users can get worried about a power outage when driving if there are insufficient charging stations (CSs). This range anxiety limits the penetration of EVs (Salah & Kama, 2016; Pevec et al., 2019). According to Salah &

Kama (2016), there are several solutions to lower this range anxiety and thus to accelerate the penetration of EVs. On the one hand, there are some solutions at the level of the EV users and the EVs itself. For example, one of the solutions is to make the estimations of the remaining state of charge (SoC) more accurate, such that EV users can rely on it more. The other solutions at the level of the EV driver and the EVs itself proposed by Salah & Kama (2016) are "efficient route selection" (i.e. selecting the route with the lowest energy consumption) and improving the management of the energy consumption of the EV in order to extend the range. On the other hand, they also propose another solution, which is not on the layer of the EV driver or of the EVs itself, but on the layer of CI. As the lack of available CSs is one of the major causes of range anxiety, another possible solution for reducing this range anxiety is an increase of the availability of charging infrastructure. As the number of electric vehicles sold in the European Union increases exponentially each year (Statista, n.d.), there is thus an urgent need for sufficient available charging infrastructure, in order to not slow down this increase in EVs (Metais et al., 2022). If we take a look at the penetration of EVs in Belgium, the number of electric vehicles increased from 10 EVs in 2008 (Statista, 2023) to 94,390 EVs in 2023 (Statista, n.d.-a). This exponential increase can be explained by the continuous technological innovations in electric vehicles, the increased availability of cost-efficient EV's and also by favourable governmental policies, such as the EU's new legislation on the CO_2 emission of new cars sold (Balasubramaniam et al., 2022; Li et al., 2023). As the number of EVs in Belgium is expected to increase further (an estimated 188,000 EVs by 2027) (Statista, n.d.-a), the availability of CI in Belgium will also have to keep growing.

2.2) Availability of Charging Infrastructure

There are 2 main ways in which the availability of CI can be increased. Option 1 is to install new CI. This is the more quantitative aspect of the solution for improving the availability of CI. Option 2 is to make the existing CI more usable, which means that the currently existing CI will be more available. This is the more qualitative aspect of the solution, which is often overlooked in the existing literature. This second option can be further divided into two subcategories, namely CI that is not usable because of technical causes and CI that is not usable because of behavioural causes. However, before continuing on these 2 options, we will first explain the different layers of CI in a very concise way.

2.2.1) Charging Infrastructure Typology

Before analysing the charging behaviour of EV drivers, first of all a clear distinction between the different layers of CI is needed. The figure below visually presents these different layers.

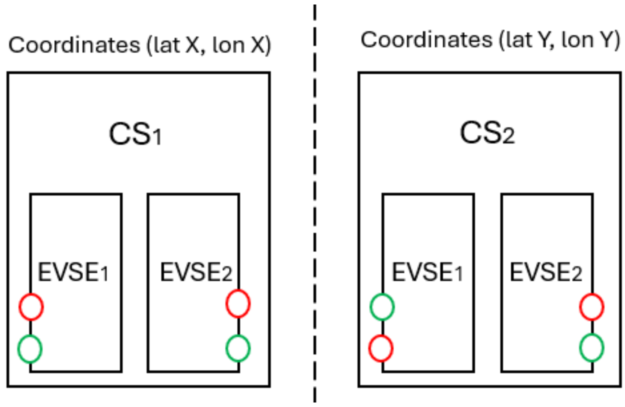


Figure 1. Different layers CI (Own creation)

The green and the red circles represent the CPs. When one CP at the same EVSE is occupied (i.e. green), then the other CP cannot be occupied anymore (i.e. red).

Starting from the level at which the analysis will be conducted in this master thesis, the Electric Vehicle Supply Equipment (EVSE) is that part of the CI through which the energy is delivered to the EV (Friese et al., 2021). One EVSE may contain multiple plugs (i.e. Charging Points (CPs)). However, only one CP can be used at the same time. The different plugs are there in order to ensure compatibility with different EV types (Friese et al., 2021). Moving to a higher layer of CI, we have a Charging Station (CS). A CS is the overarching term for all EVSEs at the same location (i.e. with the same coordinates). Unlike for the different CPs at the same EVSE, the different EVSEs at the same CS can be used simultaneously. Each CS has a unique location.

2.2.2) Quantitative Aspect

Simply increasing the amount of CI is one way to increase the availability of charging possibilities for EV drivers (van der Kam et al., 2020). This is a topic on which already a lot of research has been conducted in the existing literature. In the research of van der Kam et al. (2020), their main finding was that charging behaviour differs substantially between different locations and as a consequence also the optimal strategy for the roll-out of new CI differs between these different locations. In the research of Salah & Kama (2016) and Sankaran et al. (2020), this quantitative aspect is also being discussed, but then in the context of reducing range anxiety. When there is an insufficient amount of CI, this will lead to range anxiety for EV drivers, which will limit the penetration of EVs. Therefore, in order to limit this range anxiety, they say that the availability of CI needs to be increased, which can be achieved by deploying new CI.

When deciding about the roll-out of new CI, the location of the CI is a key determinant for its usage, which means that this needs to be carefully considered by the providers before installing the new CI (Mortimer et al., 2022b). In the literature, already a lot of research has been done in which models are developed that optimize the roll-out of new CI, in order to maximize utilization and thus also profitability of the CI. Friese et al. (2021) develop a predictive model for the usage of CI, using a two-step machine learning approach. In the first step, they use an agglomerative clustering approach to identify usage patterns of public charging infrastructure in Munich (Germany). Then, in the second step, they utilize Random Forest Classification in order to predict usage behaviour at locations for potential new CI, based on socio-demographic data in Munich. This model then can be used by providers to locate places where the new CI will be the most profitable. However, an important consideration that must be made is that these findings are contradictory to the findings of Philipsen et al. (2018), in which it is stated that there is a weak correlation between the charging behaviour of EV users and the socio-demographic characteristics of these users. Mortimer et al. (2022b) too develop a model in order to estimate the usage potential of CI, using data about places of common interest as a possible predictor of CS utilization. In the research of Pevec et al. (2018), geographical data is being used for developing a model that predicts the utilization of CSs and thus helps providers to decide on the deployment of new CI. This geographical data consists of places of common interest (such as in the research of Mortimer et al. (2022b)), but also of distances between available CSs. In the research of Straka et al. (2020), they develop 3 different predictive models that enables CPOs to estimate the popularity of places for new CI, based on social, demographic, urban and transport characteristics of that place. When evaluating their 3 predictive models, they found for all these 3 models an accuracy score of over 0.80 and an F-score of about 0.68, which means that the models are better than random models (so they have predictive value), but they are not strong enough to base decisions only on them. All these optimization models allow providers to estimate the future usage of several potential locations for new CI, and thus also to estimate the profitability of the new CI.

Although increasing the amount of CI will indeed also increase the availability of the CI, there are also some disadvantages about installing new CI. First of all, when increasing the amount of CI, this will entail high installation costs. Furthermore, when considering cities, there is often not enough space to install a lot of new CI (van der Kam et al., 2020; Wolbertus & Van Den Hoed, 2017). Therefore, research should not only focus on increasing the total amount of CI, but also on making the existing CI more reliable, as this is another possible way to increase the availability of CI without requiring additional installation costs or additional new space.

2.2.3) Qualitative Aspect

The second option to increase the availability of CI is to make the existing CI more available. This qualitative aspect can be further subdivided into two subcategories, based on the cause of the CP being unavailable for charging.

2.2.3.1) Technical Causes

This first subcategory within the qualitative aspect states that in order to increase the availability of CI, the existing CI itself should be made more reliable (Karanam & Tal, 2023). Rempel et al. (2022) define the *reliability* of an EVSE as an EVSE "*that charges the EV, for the expected duration, after using an appropriate payment method, at the expected rate (i.e. kW)"* (Rempel et al., 2022, p.3). Unlike for the deployment of new CI, research about making existing CI more reliable is less represented. Instead, most researchers, when they conduct research about increasing the availability of CI, they only focus on the quantitative aspect, such as for example in the research of Salah & Kama (2016), Straka et al. (2020) and Pevec et al. (2018).

In the scarce research that has already been conducted about the reliability of CI, the focus was mainly on giving a descriptive overview of the reliability of the existing CI. This is for example the case in the research of Stosur et al. (2020), in which they investigate the different potential failure states that can occur in CI when there is an EV charging. In the research of Rempel et al. (2022), they also conduct a very descriptive analysis about the functioning and the non-functioning of CI. More specifically, they evaluate the reliability of 657 fast charging EVSEs, located in California (USA), by visiting them and testing whether they can charge their EV for 2 minutes or not. They found that only in 72.5% of the cases the EVSE was functional, meaning that the EVSE had "*charged an EV for 2 minutes or was charging an EV at the time the station was evaluated*" (Rempel et al., 2022, p.1). In 4.9% of the cases, the cable of the EVSE was too short in order to reach the EV that needed to be charged (which is a design failure). In 22.7% of the cases, the EVSE was non-functional, which was the result of many different causes (e.g. screens that were not responding, a broken connector, payment system failures etc.). Besides this, their research also showed that these failures are not quickly repaired, since after revisiting some of the initially investigated EVSEs (8 days later), they found that of the 80 EVSEs revisited, 48 were still functional, 8 changed from non-functional to functional, 10 from functional to non-functional and 14 remained non-functional, 13 of which for the same reason as they were originally non-functional.

Although most of the existing literature focuses more on giving a description of the reliability of CI, there has already been done some research about developing a model to predict the reliability and hereby also the availability of a CS, such as in the research of Zhang & Wozniak (2017). They start their analysis by looking at the CS as a system. In that system, several components need to cooperate in order for the system to be able to work reliably. When a certain component failures, this can lead to a (partial) loss of function of that component, which can in return result in a (partial) loss of the whole system (i.e. the CS). In other words, such component failures can have a major impact on the reliability and as a consequence also on the availability of the CS. In their research, they develop a model that allows to calculate this impact. For this calculation, they take into account the number of EVs that are charging at the same CS (at different ports or connectors of that CS) and also the power output at each of these connectors. Based on this developed model, it is possible to calculate the probability that a failure will occur at a specific CS, given the number of EVs that are charging at that CS and the power output of the connectors of that CS, at a given moment in time. They find for both variables an inverse relationship with regard to the reliability and thus also the availability of the CS. The more EVs are charging at the same CS, the higher the probability of failures at that CS and thus the lower the reliability and the availability of that CS. The same applies to the power output.

2.2.3.2) Behavioural Causes

This second subcategory within the qualitative aspect is about all possible (behavioural) causes, that can prevent a specific CP to be available for an EV driver for charging, in scenarios where neither the CP is charging another EV nor having any technical failures (see *2.2.3.1) Technical causes*).

One example of such a behavioural cause, on which already some research has been conducted, is the phenomenon called “charging station hogging”. Charging Station Hogging (CSH) occurs when EV drivers leave their vehicle plugged at the CS much longer than necessary, because they use the CS not only to recharge their EV, but also as a parking spot (Wolbertus & Van Den Hoed, 2017). Due to this CSH the existing charging infrastructure is not optimally used, which may cause problems in regions lacking sufficient CI. This CSH is mainly observed in cities (where parking spots are sometimes hard to find) and less at CI placed along highways (Wolbertus & Van Den Hoed, 2017). For example, the research of Van Den Hoed et al. (2013), conducted in the city of Amsterdam (in the Netherlands), found that an EV occupying a CP was actually charging only 12% to 18% of the time and thus using the CP as a parking spot the rest of the time. In the research of Wolbertus & Van Den Hoed (2017), they state that by reducing this CSH, the existing CI will be more available for EV drivers, which will increase the utilization of the existing CI.

Although this CSH is something that has more to do with the behaviour of other EV drivers instead of with the CI itself, this does not mean that CI providers cannot take measures against it. In the research of Wolbertus & Van Den Hoed (2017), they did a data-driven analysis about CSH on a real-life dataset in the Netherlands. After their analysis, they conclude that CSH causes a big threat for an effective usage of the existing CI. In order to reduce this CSH substantially, they propose several solutions which CI providers can apply, including "*an idle fee, a monetary reward for positive behaviour, connecting users through the internet, valet charging and allowing unplugging by other users*" (Wolbertus & Van Den Hoed, 2017, p.10). They conclude that the most effective solutions are the ones in which the users can get a monetary reward or a monetary penalty.

However, besides this CSH, there exist also a lot of other causes which can prevent a specific CP to be available for an EV driver for charging, in scenarios where neither the CP is charging another EV nor having any technical failures. These other causes include for example road constructions which can make the CP inaccessible or a vehicle that is parked before the CP, which makes that CP inaccessible too. However, in the existing literature, there is not yet a lot of research conducted about these types of causes. Nevertheless, we can observe in the literature that these causes impact the availability of the CI. This is the fact for the research of Rempel et al. (2022). In their article, they state that from the 678 EVSEs they initially identified in the Great Barrier Area, they had to exclude 21 of them from their evaluation for reliability, because the "*adjacent parking space was occupied by a non-charging vehicle*" (Rempel et al., 2022, p.5), which is a reasonable number. These non-charging vehicles were both non-EVs (in 7 cases) and EVs not charging (in 14 cases). However, although they identify in their research that this has an impact on the availability of the CI, they do not further take it into account for their research, but instead exclude them.

By countering these behavioural causes of CI inaccessibility, providers of this CI can increase the availability of their CI further.

2.3) Research Gap

A thorough study of the existing literature about increasing the availability of CI pointed out that until now, most of the research has been focused on the quantitative aspect. However, as has become clear in this literature review, there exists also another possibility for achieving more available CI, namely the qualitative aspect. The fact that the qualitative aspect also has a significant impact on the availability of the CI, is clearly demonstrated in the research of Rempel et al. (2022). Not only did they find that a significant portion of the EVSEs were non-functioning (i.e. 27,5%, only considering failures with a technical cause), but also that these non-functioning EVSEs are not quickly repaired. Also in the research of Liu et al. (2023), in which a more holistic study about increasing the availability of CI is being conducted, the importance of the qualitative aspect is highlighted. By applying machine learning techniques on 20 880 consumer reviews about CI over a time horizon of 10 years, they were able to make a classification of these reviews. The 3 main concerns they identified by doing this classification were about (1) "*location features or amenities related to the station*", (2) "*functionality*" of the station and (3) "*availability*" (i.e. whether the station is occupied or not and available for charging) (Liu et al., 2023, p.4). In this classification, we can clearly observe the categorization of this master thesis, wherein the first concern is related to the *quantitative aspect*, the second concern related to the *technical aspects* of the *qualitative aspect* and the third concern partially related to the *behavioural aspect* of the *qualitative aspect* (depending on whether the EV occupying the CS is still charging or not, which is then CSH if not, or whether the CS is not available because of other causes, such as for example road constructions). Concern (2) and (3) together accounted for 43,7% of all concerns, which is a substantial portion (Liu et al., 2023).

Although its importance, still less research has been conducted on this aspect. Nevertheless, some researchers already focus in their research on this qualitative aspect. However, these researchers only consider one part of the qualitative aspects, such as for example Rempel et al. (2022) focusing on failures with a technical cause or Van Den Hoed et al. (2013), discussing only CSH (which is a part of the behavioural aspect). Therefore, this master thesis will close this research gap by conducting a holistic analysis of the qualitative aspect, in which all potential causes why a certain EVSE is unavailable for charging will be taken into account in the analysis. Concrete, we will develop a framework that will allow us, for a given unlabelled dataset containing transactional charging data, to identify availability constraints for the EVSEs present in the dataset. For building this framework, we will not only analyse the charging behaviour at the EVSE itself, but also at neighbouring EVSEs. This framework will add to the understanding of how efficient the current CI is, which will not only be useful for the providers of CI and policy makers, but also for building predictive models about usage potential of new CI (e.g. optimization models that are being used for the quantitative aspect of increasing the availability of CI). These models are namely trained on datasets about transactional charging data. However, these datasets often do not contain information about potential defects of the EVSEs, which can lead to biased results for the models, since they will treat EVSEs which are unavailable for charging for a long period of time as EVSEs that are not popular (due to for example their location).

3.) Overview Table Availability CI

The table below provides a concise overview of the different ways that exist for increasing the availability of CI. For each of the 2 aspects, it gives a brief description of what it exactly means, followed by an overview of what already has been studied in the existing research about this aspect, together with some references to articles in which this research has been conducted. For the qualitative aspect, there is also a subdivision into 2 sub aspects. The green, circled part about increasing the availability of charging possibilities by making the existing CI more available is what will be researched upon in this master thesis. (For a more detailed explanation, see point *2.3) Research gap*.)

Aspect	Meaning aspect	Subaspect	Meaning subaspect	Existing research
Quantitative aspect	Increasing amount of CI	Not applicable	Not applicable	• Different locations have different utilization rates (van der Kam et al., 2020) • Optimization models to determine deployment place with highest potential utilization (Friese et al., 2021)
Qualitative aspect	Increasing availability of existing CI	Technical aspect	Making existing CI itself more reliable (i.e. less technical defects)	• Focus on descriptive analysis (Rempel et al., 2022) • Predictive model: bottom-up approach (Zhang & Wozniak, 2017)
		Behavioural aspect	Making existing CI more accessible for charging	• CSH limits utilization of existing CI (Wolbertus & Van Den Hoed, 2017) • Other causes of inaccessibilities related to behaviour (research shortage)

Table 1. Overview increasing availability CI (Own creation)

In the overview above, we can see that there exist 2 different causes resulting in availability constraints of CI (namely technical causes and behavioural causes). However, these 2 different causes of availability constraints do not explain how the availability constraints can be observed in the data. For this, we have to think about the consequences of these causes. On the one hand, we can observe an extremely long connection time for a certain CE, which is called CSH (a behavioural cause). On the other hand, we can observe an extremely long idle time between 2 consecutive CEs. These long idle times can be the result of both behavioural causes (e.g. the road being blocked, making the EVSE inaccessible) and of technical causes (e.g. an EVSE having a technical defect making it unavailable for charging). This is represented in the figure below.

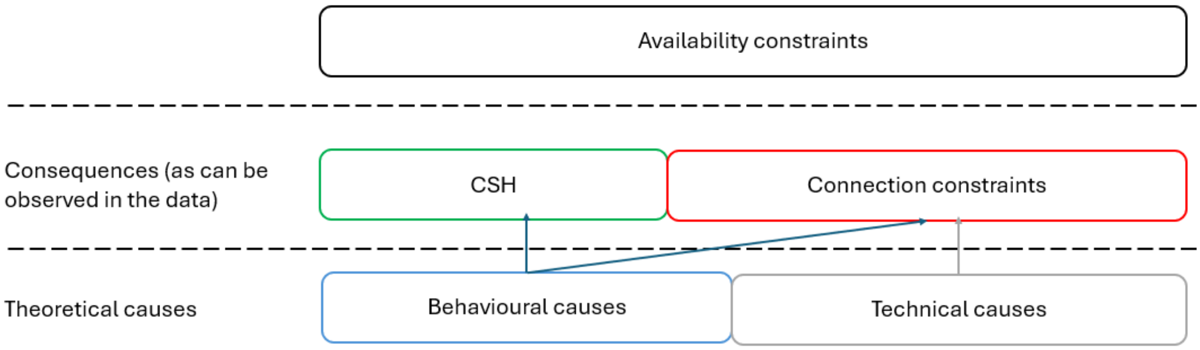


Figure 2. Terminology availability constraints (Own creation)

CSH corresponds thus to all periods for which the EVSE is not available for charging due to an EV being connected while not charging, which can be observed in the data as a period with an extremely long connection time. Connection constraints contains all periods for which the EVSE is not available for charging because EVs are unable to connect to the EVSE (e.g. the EVSE having a technical defect or being inaccessible). This distinction is needed since both cases require a different detection method. This will be discussed further in section *6.)*.

4.) Methodology

As already mentioned, the goal of this master thesis is to develop a model that can identify (for a given dataset about CEs gathered over a certain time period) how much charging time is being lost due to the EVSEs being unavailable for charging (see *3.) Overview table availability CI* for the different possible causes of CI unavailability). Since in the existing literature this approach of detecting potential availability constraints based on the charging behaviour at the EVSE itself and at the neighbouring EVSEs is not yet really been investigated upon, the methodology in this master thesis has a rather explorative character.

The framework for detecting these availability constraints in the dataset will be based on the machine learning technique anomaly detection. As defined by Chandola et al. (2009), anomaly detection is the process of "*finding patterns in data that do not conform to expected behaviour*" (Chandola et al., 2009, p.1). Since the dataset that will be used in this master thesis does not contain any labels (i.e. does not contain information about when a specific EVSE was defect or about other reasons causing the EVSE to be unavailable for charging), unsupervised anomaly detection will be applied in this master thesis. The specific unsupervised anomaly detection that will be used in this master thesis will be the Gaussian model-based parametric technique (Chandola et al., 2009).

The analysis will consist out of three major steps. In a first step, we will identify flags (i.e. discrepancies in the data which require further attention) for potential periods with connection constraints or for potential periods of CSH, by only looking at the charging behaviour of the EVSE we are considering. Then, in a second step, we will further investigate these identified discrepancies by analysing the charging behaviour (during the same time period as the identified discrepancies) at the adjacent EVSEs at the same CS. In a third step, we will further broaden our analysis by also looking at the charging behaviour at EVSEs at neighbouring CSs.

5.) Data

Before delving into the development of the framework, it is worth to first explore the data into more detail, in order to verify the assumptions that are inherent to the anomaly detection techniques used and to discuss the data-preprocessing steps that are needed. The data that will be used in this master thesis will be the publicly available dataset about CEs in Munich that is also being used in the research of Friese et al. (2021). In this dataset, every row corresponds to a unique CE for one of the 2409 EVSEs present in this dataset. Table 2 provides an overview of the attributes that will be used for the calculations in this master thesis, together with its explanation.

Column label dataset	Meaning
evse_id	Unique id for the EVSE
interval_id	Unique id for every recorded CE
state_start_ts	Starting time of CE
state_end_ts	End time of CE
interval_length	Total duration of CE
lat	Latitude coordinate of EVSE
lon	Longitude coordinate of EVSE
power_rating_kw	Capacity of EVSE (in kW)

Table 2. Used attributes (Own creation)

As stated by Chandola et al. (2009), unsupervised anomaly detection techniques make "*the implicit assumption that normal instances are far more frequent than anomalies in the dataset*" (Chandola et al., 2009, p.10). When plotting the number of CEs against the duration of the CEs (see Figure 1), we can clearly observe that the density of CEs between 0 and 4 hours (which is the connection time expected for a normal CE) is much higher than the density of CEs longer than 4 hours, which confirms this implicit assumption.

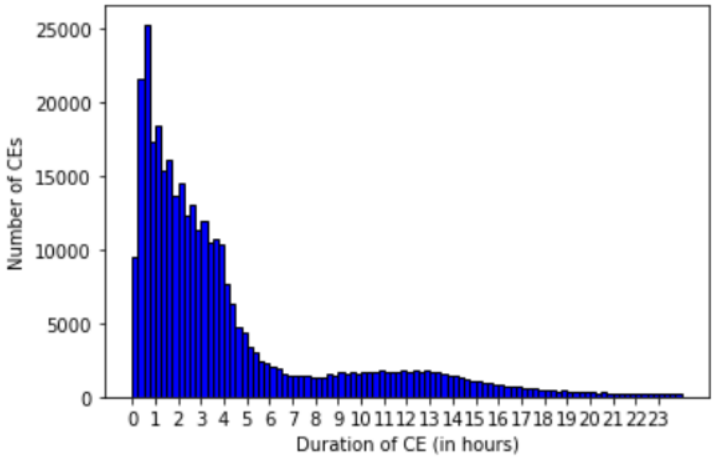


Figure 3. Histogram of number of CEs relative to duration of that CE (Own creation)

Besides the fact that most of the CEs have a duration between 0 and 4 hours, we can also observe in this histogram that no CE has a duration that exceeds 24 hours. This is because, in the dataset, all CEs that have a connection time of longer than 24 hours are cut off after 24 hours. The remaining connection time is then recorded as a new CE. This causes limitations, since for the periods of CSH (see section *6.2.1) Extremely Long Connection Times*) we will only be able to detect the periods between 16 hours and 24 hours, since the periods over 24 hours are recorded as a new CE and therefore, it will start again at connection time of 0 hours, which means that for the next 16 hours we will not classify it anymore as a period of CSH. Next to this, it also has an impact on the analysis of the idle time. Because the period longer than 24 hours is recorded as a new CE, the time between these 2 CEs will be 0. As will be discussed further in this section, the CEs for which the idle time is equal to 0 will be omitted from the analysis, and therefore these new recorded CEs will be removed from the dataset.

Besides analysing the duration of CEs, we will also analyse the idle time between 2 consecutive CEs at the same EVSE, in order to detect potential periods with connection constraints. When making a histogram of the time between 2 consecutive CEs at the same EVSE, we can clearly observe that shorter periods of idle time (i.e. periods of less than 24 hours) are by far more frequent than periods of very long idle time (see figure below).

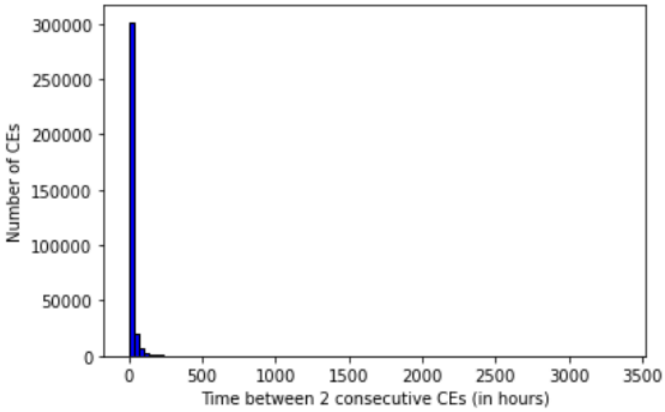
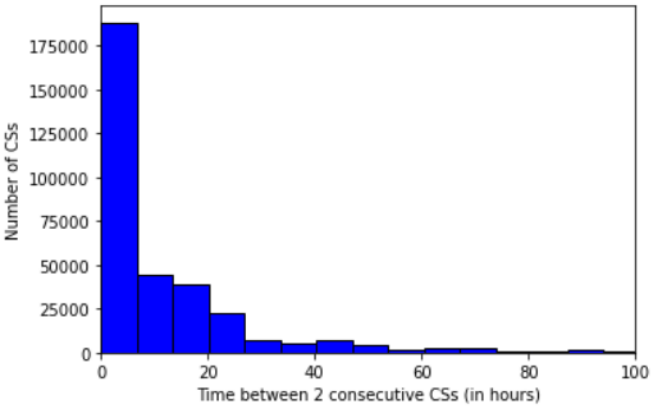


Figure 4a. Histogram of number of CEs relative to time until previous CE (Own creation)



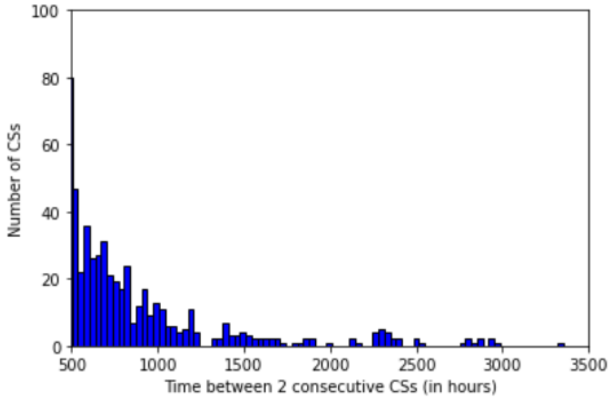


Figure 4b. Histogram of number of CEs relative to time until previous CE (zoom on small x values) (Own creation)

Figure 4c. Histogram of number of CEs relative to time until previous CE (zoom on big x values) (Own creation)

For the Gaussian model-based parametric technique, that will be applied on the distribution of the idle time between 2 consecutive CEs at the same EVSE, first of all we have to have sufficient observations for making the distribution of historical charging behaviour at a specific EVSE. Therefore, in our framework, we only will do the analysis on EVSEs with a sufficiently high number of CEs. In the framework, the threshold is set at 3 CEs a week, so a total of 66 CEs during the 22 weeks for which we have data available. For 1481 EVSEs out of the 2409 present in the dataset, this condition is satisfied, which means that we will do our analysis on these 1481 EVSEs. This first pre-processing step leads to an omission of 18 680 CEs out of the dataset.

Next to this, before being able to apply the Gaussian model-based parametric technique, the distribution has to comply with the assumption that "*the data is generated from a Gaussian distribution*" (Chandola et al., 2009, p.30) (i.e. a normal distribution). However, when looking at the distribution of the idle time between 2 consecutive CEs, we observe an exponential distribution instead of a normal distribution. Therefore, before being able to apply the Gaussian model-based parametric technique, we first have to transform the data such that the idle time between 2 consecutive CEs at the same EVSE follows a Gaussian distribution. A commonly used method for transforming exponentially distributed data into normally distributed data is by applying a log transformation on the data. However, since for some observations this time between 2 consecutive CEs is equal to 0, we cannot directly take the log value of it (since log(0) = -infinite). Therefore, we decided to omit these 0 values from the calculations, which results in a loss of 22 955 CEs. After applying the log transformation on the dataset for which the 0 values were omitted, we observe the following distribution, which closely resembles the Gaussian distribution.

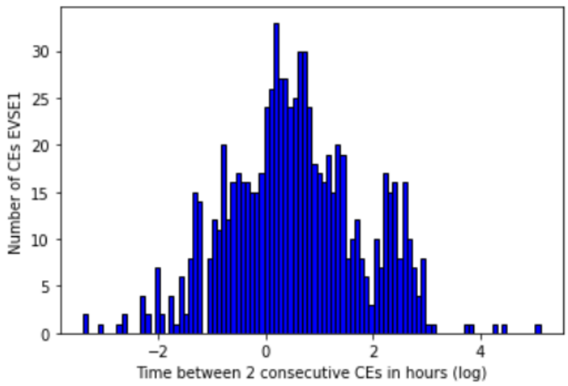


Figure 5. Histogram of number of CEs relative to time until previous session (log transformation, values 0 omitted) (Own creation)

This distribution is made for the EVSE with most recorded CEs in the dataset.

A final step that has to be taken in the data pre-processing part has to do with the way in which we can detect availability constraints in the dataset. As will be explained in more detail in section *6.1)*, one potential cause of an availability constraint is when the EV driver can successfully connect its EV to the EVSE, but no energy is being transferred to the EV. This is often referred to as a failed charging attempt (Rempel et al., 2022). However, since data on actual energy transferred to the EV is missing in the dataset, it is impossible to determine with certainty after how long connection time the EVSE can be considered as functioning. Therefore, in this master thesis, we will use a threshold instead and we will consider an EVSE as being functional if the connection time is at least 2 minutes. This is based on the article of Rempel et al. (2022), in which they classify an EVSE as being functional if it charges an EV for at least 2 minutes. Therefore, as a preparatory step, we exclude all the CEs with a connection time of less than 2 minutes from the dataset, since we assume that these CEs represent failed charging attempts. We will therefore treat these CEs as if they did not occur. This exclusion results in an omission of 125 CEs.

Below, the impact of all data pre-processing steps on the dataset is visualized.

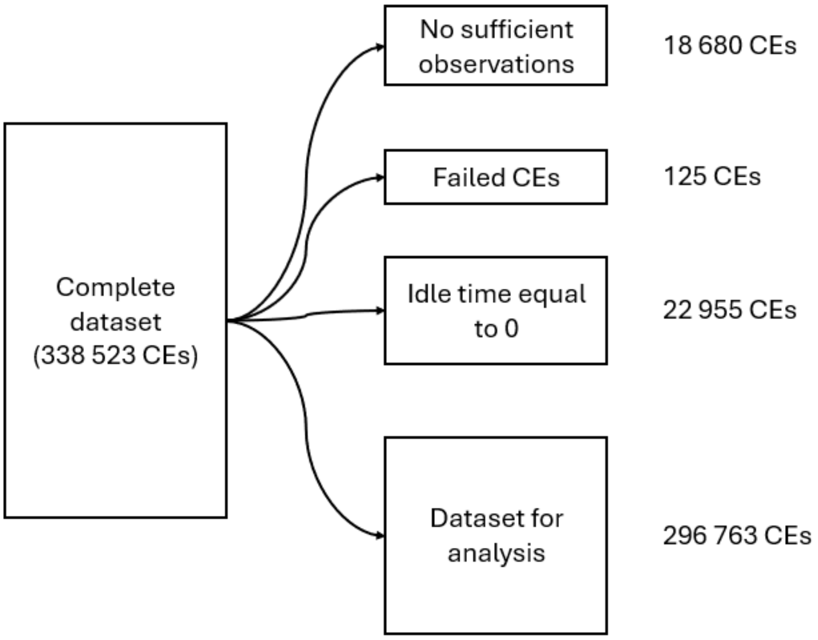


Figure 6. Visualization impact data pre-processing on dataset (Own creation)

The dataset after all pre-processing steps, which will be used in the development of our framework, contains information on 296 763 CEs. Below, an overview of the descriptive statistics for these 296 763 CEs is provided.

	Connection time (in hours)	Idle time (in hours)	Total connection time per EVSE (in hours)
Median	2,58	4,19	1101,2
Mean	4,26	12,24	1077,28
Std	4,67	26,13	503,99
Min	0,03	0	36,93
Max	23,99	2301,07	2229,51

Table 3. Descriptive statistics dataset after pre-processing steps (Own creation)

6.) Framework

In this section of the master thesis, we will explain the structure of the framework for identifying potential irregularities in the charging behaviour that may indicate a connection constraint or a period of CSH, which negatively impact the availability of the CI. As already mentioned, the framework is built up using three steps, corresponding to subsection *6.2), 6.3)* and *6.4)*. However, before being able to develop the framework for identifying potential CI availability constraints, we first need to understand how these emerge in the data. This will be explained in subsection *6.1)*.

6.1) Overview CI Availability Constraints

As can be seen in the overview table for CI availability (*Table 1*), CI being unavailable for charging can be attributed to different high-level causes (i.e. technical causes and behavioural causes). This has also implications for detecting these different causes in the dataset. The table below gives a more detailed listing of potential causes resulting in the CI being unavailable, and how these causes can be detected in the dataset. The table is a mixture of what already has been discussed in the existing literature, supplemented with own ideas for potential causes that have not yet been discussed in the existing literature.

Failure	Classification of failure	Detection of failure
Design failure (i.e. cable EVSE too short to reach EV) (Rempel et al., 2022)	Technical cause	Long idle time
Broken connector/plug (Rempel et al., 2022)	Technical cause	Long idle time
Payment sytem failure (Rempel et al., 2022)	Technical cause	Long idle time
Non-responding screen (Rempel et al., 2022)	Technical cause	Long idle time
EV plugged while not charging (CSH) (Wolbertus & Van Den Hoed, 2017)	Behavioural cause	Long connection time
Unexpected shut off during charging (Rempel et al., 2022)	Technical cause	Short connection time
Adjacent parking spot occupied without charging	Behavioural cause	Long idle time
Blocked road (e.g. road constructions)	Behavioural cause	Long idle time
Incorrect usage instructions	Technical cause/ Behavioural cause	Long idle time/ Short connection time

Table 4. Overview CI availability constraints (Own creation)

Note that this list is not an exhaustive list, since it is not feasible to consider every potential event causing CI unavailability without extensive field research. However, since the dataset is not detailed enough and since other useful data (e.g. a list of road constructions for the period we will analyse) is not available for this period, it would anyway not be possible to assign the identified availability constraints in the dataset to a cause in such great detail. Instead, the main purpose of this table is to understand how to look for availability constraints in the dataset in the further development of the framework. Assigning a specific, low-level cause to every identified anomaly is out of scope for this master thesis. Instead, we will focus on assigning high-level causes (at the level discussed in *Table 1*).

6.2) Step 1: Identifying Discrepancies in Individual Charging Behaviour

In the first step of the framework, the objective is to identify (for every EVSE separately) anomalies in the charging behaviour. As can be seen in *table 4*, there are 3 different ways in which CI availability constraints can be observed in the dataset, namely extremely short connection times, extremely long connection times and extremely long idle times (i.e. the time between 2 consecutive CEs at the same EVSE). In the data pre-processing part, we already treated the CEs with an extremely short connection time (i.e. a connection time of less than 2 minutes) as if they not occurred, by omitting them. These failed charging attempts are thus incorporated in the idle time of an EVSE, and are therefore included in the analysis of the long idle times. The other 2 remaining ways in which availability constraints can arise in the data (i.e. extremely long connection times and extremely long idle times) will be discussed in the following 2 subsections below.

6.2.1) Extremely Long Connection Times

For identifying potential CSH in the dataset, we have to analyse CEs with a long connection time. However, as research has already shown (*see also section 2.2.3.2) Behavioural Causes*), whether CSH occurs during a CE is dependent on a numerous of different factors, such as the SoC at arrival, charging rate, battery capacity of the EV etc. Since this information is lacking in the dataset, it is impossible to perfectly determine when CSH occurs for every individual CE in the dataset, but instead we will apply a general threshold, based on the existing literature. In the research of Wolbertus & Van Den Hoed (2017), they give different possible definitions to determine CSH. One of these definitions is to categorize all CEs with a connection time of over 16 hours as CSH. When the CE exceeds this threshold of 16 hours, we will assume that CSH took place during this CE.

Therefore, in the first step of the framework, we will indicate all CEs with a connection time of over 16 hours as potential periods of CSH.

6.2.2) Extremely Long Idle Times

The final way of observing potential availability constraints in the dataset is by analysing periods for which the idle time between 2 consecutive CEs at a given EVSE is extremely long. These types of connection constraints will be identified by applying the commonly used statistical technique of selecting all data points which are more than 1,96σ away from the mean value (in appendix, a table can be found containing all parameters that are used in this master thesis). For every EVSE, we will first build a distribution of all idle times and then we will categorize all the periods for which the time between the 2 consecutive CEs exceeds this 1,96σ as potential connection constraints.

6.3) Step 2: Analysing Charging Behaviour at Adjacent EVSE at same CS

In the second and the third step of the framework, we will now further analyse the identified anomalies out of the previous step. Since for every CE in the dataset the maximum connection time is 24 hours and CSH is only identified when the CE exceeds 16 hours, these periods of CSH are too short (i.e. maximum 8 hours) for analysing the charging behaviour at nearby EVSEs during these periods, because they are too sensitive for single observations in CEs. Therefore, we will not further analyse the identified periods of CSH out of the first step, also because these 16 hours are a generally accepted cut-off to indicate CSH.

The CEs with a very short connection time were omitted before identifying the periods with a very long idle time and are therefore automatically included in the calculations for detecting these long idle times (see section *5.) Data*), since omitting these failed charging attempts out of the data lengthens the idle time. However, for the periods in which we identified in step 1 anomalies in the idle time between 2 consecutive CEs at the same EVSE, analysing the charging behaviour of neighbouring EVSEs will provide additional confidence and at the same time additional information to our framework.

In this second step of our analysis, we will limit our analysis by only examining the charging behaviour at adjacent EVSEs at the same CS as the EVSE we are considering (i.e. the EVSE for which we identified the anomaly in the first step). For every period in which we identified a potential connection constraint for an EVSE in the first step of the framework, we will thus analyse the charging behaviour at his adjacent EVSE at the same CS. There are 2 possible scenario’s we are looking for. On the one hand, the adjacent EVSE can also have an anomaly with regards to idle time during the same period. On the other hand, the other EVSE can have an occupancy rate that is significantly higher than one would expect (based on the historical charging behaviour at that specific EVSE). This additional information about the charging behaviour at the adjacent EVSE at the same CS will not only provide a higher degree of certainty about the identified period with a connection constraint, but it will also provide more insights about the possible causes of the connection constraint. In the first scenario, the cause of the connection constraint for the EVSE we are considering does not only affects that specific EVSE, but also the other EVSE at the same CS. This can be for example the CS having a technical defect, but also the CS being inaccessible due to road constructions etc. In the second scenario, the CS is still accessible and not defect, but the cause of the connection constraint for the EVSE we are considering is specific for that EVSE, which can be due to for example the plug being broken.

The first scenario of detecting overlaps in the anomalies is straightforward, since we already identified for every EVSE the periods with connection constraints in the first step of the framework. Therefore, in this second step, we just need to identify those periods for which there is an overlap in connection constraints for a pair of adjacent EVSEs.

However, in order to test whether the occupancy at the adjacent EVSE is significantly higher during the period of the observed anomaly, we first have to gain insights about the variability of the occupancy rate for each individual EVSE. Therefore, for each EVSE, we calculate the monthly occupancy rate. This results in multiple occupancy rates for the same EVSE, which allows us to calculate the mean and the standard deviation of the occupancy rate per EVSE. We then use these mean and standard deviation to build a confidence interval of occupancy rates for every EVSE. We want to verify whether the observed occupancy rate is significantly higher than the expected occupancy rate (i.e. the mean of the occupancy rate). For this one-tailed hypothesis test, we use again 1,96σ (which is the same threshold as in step 1). This 1,96σ corresponds to a confidence level of 97.5% (since it is a one-tailed hypothesis test). When the occupancy rate of the adjacent EVSE is higher than the upper limit of this confidence interval (i.e. 1,96σ), we conclude that it is significantly higher than the expected value.

The outcome of step 2 is a subset of the periods for which we detected a potential connection constraint in step 1, for which we have a higher degree of certainty because of simultaneous anomalies in the charging behaviour at the adjacent EVSE at the same CS (i.e. anomalies observed in a multivariate context (Boesel et al., 2023)).

6.4) Step 3: Analysing Charging Behaviour at EVSEs at nearby CSs

This third step of our analysis is analogue to the second step of the framework. However, now we do not analyse the charging behaviour of the adjacent EVSE at the same CS (i.e. with a distance equal to 0), but we analyse the charging behaviour at EVSEs nearby the EVSE we are considering. For determining which EVSEs are nearby for a certain given EVSE, we calculate the pairwise distance between all pairs of EVSEs in the dataset. For calculating the distance, we use the formula of the Euclidian Distance, as described in the article of Crozier et al. (2018).

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_i (q_i - p_i)^2}.$$

Figure 7. Formula for the Euclidean Distance between 2 datapoints (p and q) (Crozier et al., 2018)

We define a nearby EVSE as an EVSE with a distance of 250 meters or less to the EVSE considering. The reason to take 250 meters is because, as stated by Csiszár et al. (2019), this is a "*widely accepted walking distance between the parking place and the destination*" (Csiszár et al., 2019, p.176), which they base on the research of Daniels & Mulley (2013). As already discussed, EV drivers often use CI not only to charge their EV but also as a parking spot, especially in cities (Wolbertus & Van Den Hoed, 2017). Therefore, we take these 250 meters as the maximum distance that an EV driver wants to search further for another EVSE when his first intended EVSE is unavailable for charging.

By also analysing the charging behaviour at nearby EVSEs, we strengthen our framework in various ways. First of all, it increases our confidence about the anomalies identified in step 1 and reinforced in step 2 even further. Secondly, for some EVSEs, there are no adjacent EVSEs at the same CS, which makes the second step of the framework impossible. Therefore, by broadening the analysis by also including nearby EVSEs, we can still enhance the confidence of the anomalies identified in step 1. A final way in which this strengthens our model is because this analysis of the nearby EVSEs provides us with some additional insights about the categorization of the possible causes of the connection constraint. When for example, after performing step 2 in our framework, we have some periods in which there are simultaneous connection constraints for 2 adjacent EVSEs at the same CS, by also analysing the charging behaviour of the nearby EVSEs, we can find out whether the cause of these connection constraints is specific to the CS or whether the other nearby CSs are also affected by this cause (which can for example be the road that is being inaccessible).

7.) Results

In this section, we will discuss the results of our analysis. In the first subsection, we apply the framework on the publicly available dataset of CEs in Munich. In subsection *7.2)*, we make a categorization of the connection constraints identified in the first subsection. In subsection *7.3)*, we describe one specific identified connection constraint, which allows us to also visualize it. Then, finally, we also analyse the impact of taking into account periods with availability constraints on the occupancy rates of the EVSEs (which is often used when building predictive models for the deployment of new CI).

7.1) Application Framework

For the first step of our framework, from the cleaned dataset (i.e. the dataset containing all EVSEs for which there are at least 66 CEs and containing all CEs with a connection time of at least 2 minutes), we first select all CEs with a connection time of more than 16 hours, which represents all periods of CSH. This gives us 11 273 CEs for which we observed CSH on a total of 319 718 CEs present in the cleaned dataset. On average, 2,6% of the total connection time of an EVSE is longer than 16 hours. However, there are some outliers for which this percentage is over 10%, which is a considerable amount. Below, a histogram can be found of the distribution of these percentages across the different EVSEs.

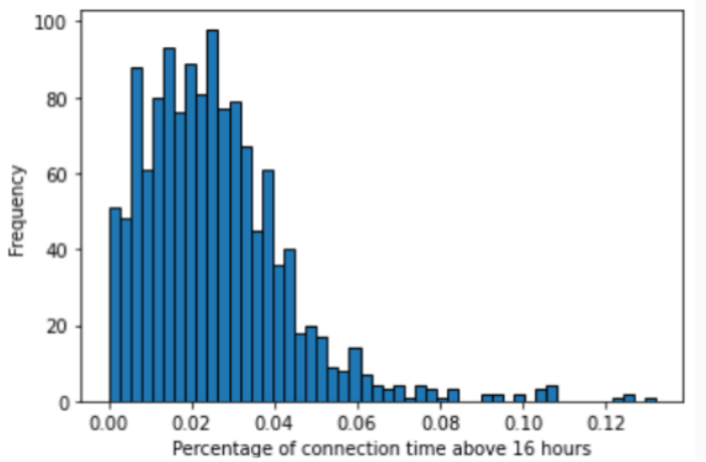


Figure 8. Histogram percentages of connection time above 16 hours per EVSE (Own creation)

Next, we want to identify all periods for which the idle time between 2 consecutive CEs is significantly longer than expected, based on historical charging behaviour at that particular EVSE. For doing this, we base our analysis on the mean and the standard deviation of the log(idle time) for every EVSE. However, remember that before being able to take the log of the idle times, we first had to omit all CEs for which the time until the previous CE was equal to 0. All the CEs for which the log(idle time) until the previous CE is more than 1,96σ from the mean of the logs of the idle times, we classify as periods with a potential connection constraint. This results in a total of 3118 periods (on a total of 296 763 CEs), which will be further analysed in the second and in the third step of the framework, in order to gain an increased confidence about these anomalies.

After performing the second and the third step of the framework (i.e. analysing the charging behaviour at adjacent EVSEs and at EVSEs with a distance of less than 250 meters as the crow flies), we also observe unexpected charging behaviour at adjacent or nearby EVSEs for 2343 of the 3118 periods, which thus gives us increased confidence that these detected anomalies in the data represent real availability constraints for that specific EVSE (see increased confidence by detecting multivariate outliers (Boesel et al., 2023)).

These 2343 periods together cause a total downtime of nearly 265 000 hours. However, this is the upper bound and is not equal to the connection time that is being lost because of these connection constraints, since the EVSEs are not connected for 100% of the time. Based on the average occupancy rate of every EVSE (corrected for the identified periods containing availability constraints), this corresponds with a lost connection time of 88 685 hours. The 11 273 periods of CSH (identified in the first step of the framework), cause a downtime of nearly 33 000 hours, with a corresponding lost connection time of around 11 000 hours. These two combined cause a total downtime of around 296 300 hours, with a corresponding lost connection time of almost 100 000 hours. Expressed as percentages, we observe that the analysed CI was unavailable for charging for around 5,45 % of the time. Connection constraints were responsible for the biggest part of the downtime (4,85% of the 5,45%), whereas CSH was responsible for 0,60% of the 5,45% of the total downtime. However, this 5,45% has to be nuanced, since it is only an average percentage. When we look at the distribution of downtimes for every EVSE (histogram below), we observe a large variability in these percentages between the EVSEs in the dataset. For some EVSEs, we observe a very small or even no availability constraints, meaning that these EVSEs are most of the time charging correctly or available for charging. However, for some EVSEs, this percentage is very high, meaning that these EVSEs are not very reliable. Furthermore, the impact of these downtimes also depends on when they appear. When it is for example during night, it will have a smaller impact on lost charging time and lost revenue than when it happens during the busiest hours.

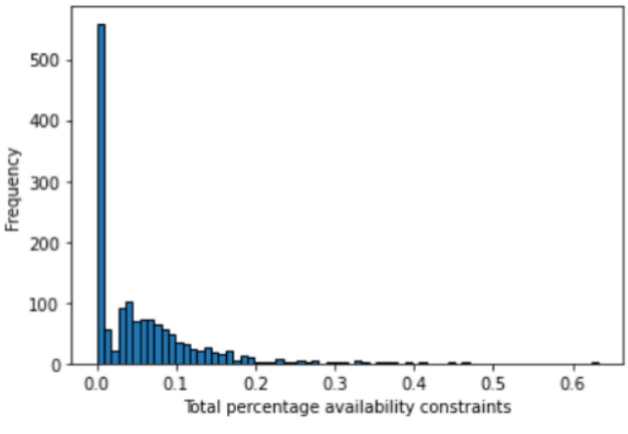


Figure 9. Histogram % availability constraints (Own creation)

We can also express this lost connection time in economic terms. However, since we do not possess any information about real energy transferred during a CE in the dataset, nor about the tariffs of the different CPOs, we are only able to make a rough approximation of it in order to get insights about the magnitude of the economic loss. We conduct the following steps for calculating the approximate lost revenue for the CPOs due to the availability constraints. Our starting point is the total amount of lost connection hours due to availability constraints, which equalizes 99 761 hours. However, when an EV is connected, it does not mean that it is charging the whole time. As stated in the research of Wolbertus et al., 2017, on average, an EV is only charging between 12% and 18% of its connection time. We use this as an approximation (we take 15%), which we multiply with the amount of lost connection time. This, we then multiply with the capacity of each EVSE (which is available in the dataset) to come to the total amount of energy that could be transferred, if there were no availability constraints. Finally, we multiply this with the price of charging 1 kWh at a public EVSE. Since the price of charging depends on various different factors, such as the time of charging, the differences in prices between different CPOs etc. (information that we do not possess), we take an average price of €0,50/kWh, which is a good approximation for the average price of charging 1 kWh at a public EVSE in Germany (Cardino, 2023). This leads to an approximated total lost revenue of around €155 000. Per EVSE, this corresponds to an average lost revenue of around €105 (over a period of 5 months). However, as the variability of the downtime percentage between the different EVSEs is high, this is also true for the lost revenue, where for some EVSEs the lost revenue exceeds €1800.

All this information can be found below in the schematic overview.

	% EVSE unavailable due to CSH	% EVSE unavailable due to connection constraints	Total % EVSE unavailable
Median	0,0045	0,0723	0,0371
Mean	0,0060	0,0485	0,0545
Std	0,0058	0,0651	0,0669
Min	0	0	0
Max	0,0307	0,6267	0,6317

Table 5. Descriptive statistics on % downtime (per EVSE) (Own creation)

	Downtime per EVSE (in hours)	Lost connection time per EVSE (in hours)	Lost energy transfer per EVSE (in kWh)	Lost revenue per EVSE (in €)
Median	298,49	32,45	95,74	47,87
Mean	200,13	67,36	211,29	105,65
Std	245,48	100,39	333,31	166,65
Min	0	0	0	0

Max	2319,67	1111,12	3637,3	1818,65

Table 6. Descriptive statistics on downtime, lost connection time, lost energy transfer and lost revenue (per EVSE) (Own creation)

	Amount
Total downtime due to connection constraints (in hours)	263 689
Total downtime due to CSH (in hours)	32 709
Total downtime (in hours)	296 398
Total lost connection time connection constraints (in hours)	88 687
Total lost connection time CSH (in hours)	11 074
Total lost connection time (in hours)	99 761
Total energy lost (in kWh)	312 924
Total revenue lost (in €)	156 461

Table 7. Total amount downtime, lost connection time, lost energy and lost revenue (Own creation)

7.2) Categorization Periods with Connection Constraints

Next to providing additional confidence about the anomalies identified in the first step of the framework, the information about the charging behaviour at neighbouring EVSEs allows us to make a more detailed categorization of the possible causes of the identified connection constraints. For doing this, we make a pivot table in which every row corresponds to a specific period in which we identified a connection constraint, whereas the columns display the possible charging behaviour at neighbouring EVSEs. Below (in *Table 8*), we have provided a selection of rows out of the full pivot table, which we will now use to further explain how this can help us for a further categorization of the periods with connection constraints. When we take for example the period with ID 1, we can see that during the period for which we identified a connection constraint at the EVSE we are considering, there was also an exceptional long idle time at the adjacent EVSE at the same CS, while no overlaps were identified at neighbouring CSs. Therefore, the cause of this identified connection constraint will probably be at CS level (since the other nearby CSs seem to be not affected). This can for example be caused by the CS having a technical defect, making all the EVSEs at this CS unavailable for charging. However, for the period with ID 2, the charging behaviour we observe at the nearby CSs is different than for the example discussed above. Besides an overlap in exceptional long idle time at the adjacent EVSE at the same CS, we also observe an overlap at a nearby CS. This means that the cause of this connection constraint is probably broader than 1 CS (since it seems to affect multiple CSs). However, there are also 2 nearby CSs for which the occupancy during this period is significantly higher. This indicates that not all neighbouring CSs are affected by the cause, and thus that EV drivers who are looking for a CS to charge their EV will, during this period, more charge at these 2 CSs with a higher occupancy rate due to the other CSs being unavailable for charging. The cause can be for example road constructions making several (but not all) neighbouring CSs inaccessible and thus unavailable for charging. In the research of Weekx et al. (2024) they find similar overflow dynamics between different CSs, namely that when the preferred CS of an EV driver is occupied, that this EV driver will look for other CSs nearby. In our analysis, we find that these overflow dynamics are also present when the EVSE is unavailable for charging because of connection constraints (in addition to when a CS is occupied).

Connection constraint ID	Adjacent EVSE (at same CS)		Nearby EVSE (at nearby CS)	
	Connection constraint	Higher occupancy	Connection constraint	Higher occupancy
1	1	0	0	0
2	1	0	1	2

Table 8. Subset table counts charging behaviour at neighbouring EVSEs for periods with connection constraints (Own creation)

This logic, we can apply in an automated way on all identified periods of connection constraints, by applying the following logic:

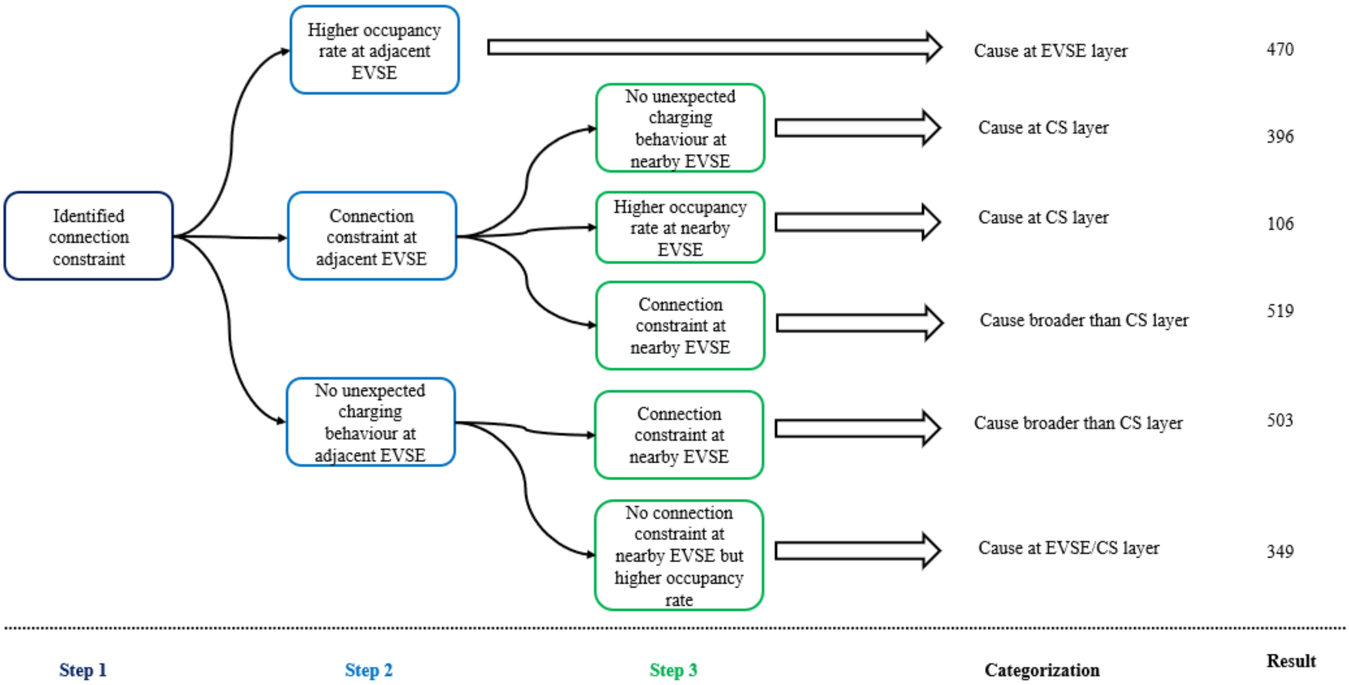


Figure 10. Categorization logic (Own creation)

Applying this logic on all periods for which we identified connection constraints results in the following categorization. For 1022 out of the 2343 identified periods with a connection constraint, the cause was broader than CS level, because we identified overlapping periods of connection constraints at neighbouring CSs. For 502 periods, the cause was at CS level, which means that we identified for an adjacent EVSE at the same CS also an exceptional long idle time during the same period. Finally, for 819 periods, the cause was at EVSE level, which means that we did not identify an overlapping period with a connection constraint at the adjacent EVSE or that we identified a significantly higher occupancy rate at the adjacent EVSE at the same CS during the same period.

Categorization	Count
Cause at EVSE level	819
Cause at CS level	502
Cause broader than CS level	1022

Table 9. Categorization identified periods with connection constraints (Own creation)

7.3) Application on Individual EVSE

In order to make this categorization clearer and more tangible, we will now provide a visual example of the logic. For a certain EVSE in the dataset, we find in the first step of the framework an outlier in the idle time between 2 consecutive CEs. This connection constraint is observed between 18/09/2021 and 04/10/2021. After analysing the charging behaviour at the adjacent EVSE (in step 2 of the framework) and at the nearby EVSEs (in step 3 of the framework), we find an overlapping connection constraint for the adjacent EVSE, while we observe a significantly higher occupancy rate at 4 nearby EVSEs. This is visualized in the following graphs.

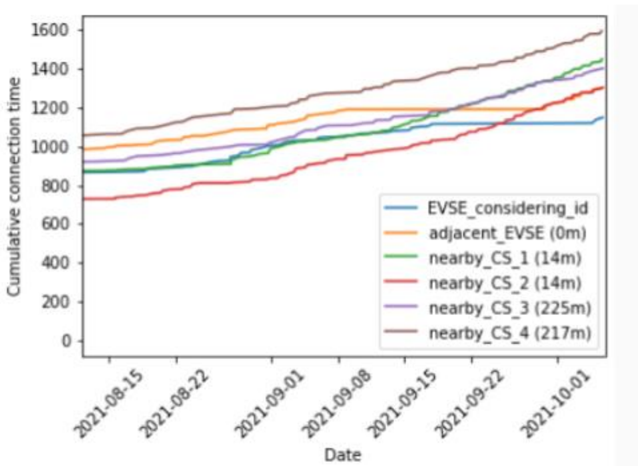


Figure 11a. Cumulative connection time for application (Own creation)

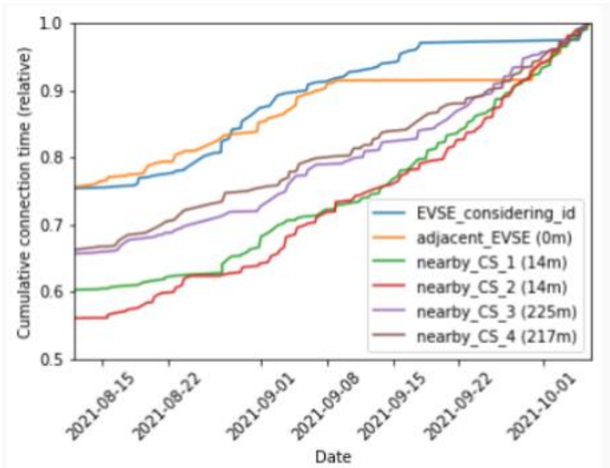


Figure 11b. Relative cumulative connection time for application (Own creation)

Note: the figures only display the cumulative connection time per EVSE for the period 15/08/2021 and 06/10/2021.

Figure 11a displays the cumulative charging behaviour for every EVSE in absolute terms, whereas figure 11b displays it in relative terms (which facilitates to see the increase in occupancy rate at the nearby EVSEs). These 2 graphs clearly show what happens at the adjacent and the nearby EVSEs during the connection constraint at the EVSE we are considering in this framework. For the yellow line (i.e. the adjacent EVSE), we observe a simultaneous horizontal line, which means that the EVSE is not used during the same period as the connection constraint of the EVSE we are considering. However, for the nearby EVSEs, we clearly observe a steeper slope for this period (i.e. between 18/09/2021 and 04/10/2021). This increase is the biggest at the EVSEs that are the closest to the EVSE with the connection constraint (i.e. for the green and the red line, both at approximately 15 meters from the EVSE we are considering in this application). For the EVSEs further away (i.e. above 200 meters), there is still a significantly higher occupancy rate, but the increase in occupancy rate is smaller.

The analysis of the neighbouring EVSEs provides us additional insights about the period for which the connection constraint occurs. Since, during this period, we also observe a connection constraint for the adjacent EVSE, the cause of the connection constraint goes beyond the EVSE layer, but instead the whole CS is being affected by it. However, the neighbouring CSs do not experience a connection constraint for this period, but instead they have an increased occupancy. This is especially the case for the closest EVSEs, for which the increase in the occupancy is the highest.

This observed charging behaviour can be explained as follows. The CS to which the EVSE of this example is attached is defect. Because of this, the other EVSE that is also attached to this CS is also unavailable for charging. Therefore, people who normally charge their EV at this CS, will look for nearby CSs for charging their EV, which explains the increased occupancy at these nearby EVSEs. This is a clear example of overflow dynamics that are present between nearby CSs (Weekx et al., 2024).

7.4) Impact Availability Constraints on Occupancy Rates

In this section, we will examine the impact of the availability constraints on the occupancy rates of the EVSEs in the dataset. For doing this, we calculate for every EVSE both the occupancy rate excluding the periods in which availability constraints occurred and the occupancy rate including the periods with availability constraints.

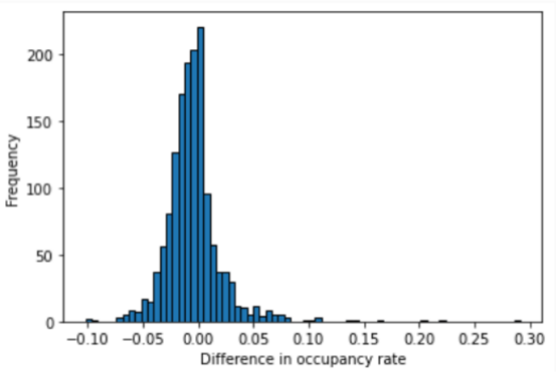


Figure 12. Histogram differences occupancy rate (occupancy rate without outlier - occupancy rate with outliers) (Own creation)

The first thing we can observe is that the difference in occupancy rate can be both negative as well as positive. This can be linked back to the different causes of the availability constraints. The difference in occupancy rate will be negative if we identified (using our framework) a substantial amount of CSH at that particular EVSE, meaning that the occupancy rate will be lower if we correct it for periods with availability constraints (by thus removing these periods of CSH). However, when we identified a considerable amount of connection constraints at that particular EVSE, the difference in occupancy rate will be positive, meaning that the occupancy rate will be higher if we correct it for periods with availability constraints (by thus removing these periods with connection constraints). The table below provides information on the statistics of the differences in the occupancy rates.

	Difference in occupancy rate
Median	-0.006
Mean	-0.005
Std	0.026
Min	-0.102
Max	0.291

Table 10. Descriptive statistics difference in occupancy rate (Own creation)

We observe that the differences in occupancy rates for all the different EVSEs seem to outweigh each other. However, this is not what we are interested in. As already mentioned earlier in this master thesis, predictive models for optimizing the roll-out of new CI often use the occupancy rates of the different EVSEs for building their model, such as for example in the research of Friese et al. (2021). However, when using these occupancy rates, they do not take into account potential periods of availability constraints for the EVSEs. For determining the implications of this, we will analyse whether, for every EVSE, both occupancy rates differ significantly from one another by performing a paired t-test.

T-statistic	-7.15
P-value	$1.36*10^{12}$

Table 11. Results from paired t-test (Own creation)

When comparing the 2 occupancy rates for every EVSE using a paired t-test, we find that both occupancy rates are significantly different from each other, with a p-value of approximately $1.36*10^{12}$.

The impact of this finding on existing research will be further discussed in section *8.) Link to the Existing Literature*.

8.) Link to the Existing Literature

In this section, we reflect our results on the existing literature. In the research of Friese et al. (2021), they use the same dataset about CEs as the dataset used in this master thesis. In the first step of their research, they extract usage patterns out of the data, which they then use in a second step to build a predictive model (see the quantitative aspect of increasing the availability of CI). For extracting these usage patterns, they use the occupancy rate calculated for every EVSE in the dataset. However, they do not take into account potential availability constraints of the CI when calculating these occupancy rates. This has implications on their model since, after conducting the paired t-test, we found that correcting a dataset for periods in which there are availability constraints has a significant impact on the calculated occupancy rates of the EVSEs. Therefore, when building a predictive model, the occupancy rates should be corrected in order to prevent the results from these models from being biased because of availability constraints that occurred during the period of observations.

This reflection also reveals the link there exists between the quantitative aspect of increasing the availability of CI and the qualitative aspect. By having insights on the qualitative aspect of CI, we can increase the quality of the deployment decisions of new CI by incorporating these additional insights into the development of predictive models (which is part of the quantitative aspect of increasing the availability of CI), by enabling the optimization models to be trained on richer datasets.

The framework allows thus to counter the limitation stated in the research of Hecht et al. (2020), namely that because the research in their article is based on data from a publicly available dataset, that they cannot verify the accuracy of the dataset and therefore consequently also cannot guarantee the correctness of their results. However, by applying the framework, we can gain additional insights in the dataset enabling to make research using these publicly available datasets more precise.

9.) Assessment Framework

In the existing literature, when conducting research about CI, researchers often split up the CI into 2 different parts (i.e. AC chargers and DC chargers) in their analyses, such as for example in the research of Friese et al. (2021) and Almutairi (2022). This is done because the charging behaviour at AC chargers differs substantially from the charging behaviour at DC chargers (i.e. fast chargers). However, since the starting point of the analysis in our developed framework is the historical charging behaviour of each individual EVSE, there is no need to first make the separation between AC chargers and DC chargers, since the difference in charging behaviour between these 2 types of EV-chargers is automatically taken into account in the analysis. This makes the framework more robust, since it eliminates the intermediate step of splitting up the dataset according to the type of EV-charger.

Now, we come to the limitations of our framework. A first limitation is that, in our analysis of detecting availability constraints in the dataset about CEs in Munich, we are not able to exactly identify CSH as well as failed charging attempts, since this requires additional data which is not present in the dataset. Therefore, we used a threshold (which we set at 16 hours). Although this is a general accepted threshold level (Wolbertus & Van Den Hoed, 2017), by applying this threshold, we will not correctly identify all periods of CSH that occurred in the dataset. This is also the reason why we set the threshold at such a high level, in order to ensure that the periods that are above the threshold are real periods of CSH. A consequence of this high threshold is that we underestimate the CSH that will have occurred during the time period, since CSH also happens when the connection time is less than 16 hours. This is also reflected in the results from our analysis, in which we found that around 2,6% of the connection times in our dataset were above 16 hours, whereas in the research of Van Den Hoed et al. (2013) about CSH in the city of Amsterdam, they come to a charging/connection ratio between 12% and 18%. This limitation can easily be solved when a dataset does contain information about the actual charging time, in addition to information about the connection time. Therefore, when in future research the framework is applied on a dataset that contains information about the real transfer of energy, the detection of CSH periods and of failed charging attempts can be further refined, which will make the results of the framework more accurate.

Another limitation is that, in order to be able to apply the Gaussian model-based parametric technique, we had to exclude some of the recorded CEs from the analysis. Since omitting data is inevitably accompanied with a loss of information, this omission will thus impact the results of our analysis. However, in order to mitigate this accompanied loss of information, we incorporated multiple steps in our framework in order to increase our confidence about the identified periods with an availability constraint.

A third limitation of our research is that, since the dataset about CEs in Munich that we use to develop our model does not contain any labels about when an EVSE actually had an availability constraint (this is also the reason why we apply unsupervised anomaly detection in this master thesis), we are not able to test the performance of our framework. However, this is something that can be done in future research, if this framework would be applied on a dataset that does contain this information.

Finally, we also need to be aware that in this framework, we make use of a static threshold for identifying availability constraints (i.e. we always work with 16 hours for identifying periods of CSH or with 1,96 standard deviations for detecting connection constraints). However, these thresholds can be refined further by using dynamic thresholds instead (i.e. thresholds that take into account the starting point of the period for which we identify a potential availability constraint). For example, a long idle time being observed during the day is more suspicious than when we would observe it at night. However, this is out of scope of this master thesis and is something that can be conducted in future research.

10.) Conclusion

In order to not slow down the penetration of EVs on the European roads, there is an urgent need to increase the availability of public CI for charging these vehicles. Since already a lot of research has been conducted on the quantitative aspect of the availability of CI, in this master thesis we try to give a more holistic overview about the qualitative aspect, since less research has been conducted so far on this qualitative aspect. More concretely, we developed a framework that allows us to identify potential availability constraints for a given dataset about CEs, by analysing the charging behaviour at the adjacent EVSE and at neighbouring EVSEs. After doing this analysis, we found strong relationships between the charging behaviour at nearby EVSEs in case of an availability constraint. This means that what happens at one EVSE can influence what happens at nearby EVSEs. For example, when we observe a significantly higher occupancy rate at an EVSE for a certain period, the reason for this change in charging behaviour may lie in a nearby EVSE having an availability constraint.

However, these gained insights about availability constraints for a given dataset do not only provide useful information concerning the qualitative aspect, but they can also be used in the development of predictive models for the deployment of future CI (which is part of the quantitative aspect). This because, when correcting for periods in which there were availability constraints, this has a significant impact on the occupancy rates of the EVSEs, which will thus have an impact on the results of the predictive models, since these are often based on the occupancy rates of the EVSEs and the CSs.

11.) Abbreviations

CI: Charging Infrastructure

CE: Charging Event

CS: Charging Station

CSH: Charging Station Hogging

CP: Charging Point

CPO: Charging Point Operator

EV: Electric Vehicle

EVSE: Electric Vehicle Supply Equipment

ICE: Internal Combustion Engine

ISM: Charging Infrastructure Status Map

SoC: State of Charge

12.) References

Ahn, S. W. (2019). EU member-states' reaction to the climate change in transportation: French case of the development of electric vehicles. *J. Contemp. Eur. Stud*, *37*, 103-127.

Almutairi, A. (2022). Impact Assessment of Diverse EV Charging Infrastructures on Overall Service Reliability. *Sustainability*, *14*(20), 13295. https://doi.org/10.3390/su142013295

Balasubramaniam, A., Balasubramaniam, T., Paul, A., & Seo, H. (2022). Electric Vehicle Usage Pattern Analysis Using Nonnegative Matrix Factorization in Renewable EV-Smart Charging Grid Environment. *Mathematical Problems in Engineering*, *2022*, 1–9. https://doi.org/10.1155/2022/9365214

Boesel, F., Schläpfer, L., Pozidis, H., & Gusat, M. (2023). Multivariate Anomaly Detection with Domain Clustering. *Proceedings of the 2023 ACM Symposium on Cloud Computing*, 193-199. https://doi.org/10.1145/3620678.3624655

Cardino. (2023, June 21). *Home vs Public Stations in Germany: Comparative Analysis*. Retrieved May 16, 2024, from https://www.cardino.de/en/blog-posts/home-vs-public-stations-in-germany

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A Survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

Crozier, C., Apostolopoulou, D., & McCulloch, M. T. (2018). Clustering of Usage Profiles for Electric Vehicle Behaviour Analysis. *IEEE PES Innovative Smart Grid Technologies Conference*. https://doi.org/10.1109/isgteurope.2018.8571707

Csiszár, C., Csonka, B., Földes, D., Wirth, E., & Lovas, T. (2019). Urban public charging station locating method for electric vehicles based on land use approach. *Journal Of Transport Geography*, *74*, 173–180. https://doi.org/10.1016/j.jtrangeo.2018.11.016

Daniels, R., & Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal Of Transport And Land Use*, *6*(2), 5–20. https://doi.org/10.5198/jtlu.v6i2.308

Delucchi, M. A., Yang, C., Burke, A. F., Ogden, J. M., Kurani, K., Kessler, J., & Sperling, D. (2014). An assessment of electric vehicles: technology, infrastructure requirements, greenhouse-gas emissions, petroleum use, material use, lifetime cost, consumer acceptance and policy initiatives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2006), 20120325. https://doi.org/10.1098/rsta.2012.0325

Dost, P., Spichartz, P., & Sourkounis, C. (2016). Cost efficiency of electric vehicles and extended range electric vehicles for various user groups. *International Symposium on Power Electronics, Electrical Drives, Automation and Motion*. https://doi.org/10.1109/speedam.2016.7526009

Egbue, O., & Long, S. (2012). Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions. *Energy Policy*, *48*, 717–729. https://doi.org/10.1016/j.enpol.2012.06.009

European Council. (2023, February 3). *Paris Agreement on climate change*. Retrieved February 21, 2023, from https://www.consilium.europa.eu/en/policies/climate-change/paris-agreement/

European Parliament. (2023, February 14). *Fit for 55: zero CO2 emissions for new cars and vans in 2035*. Retrieved February 21, 2023, from https://www.europarl.europa.eu/news/en/press-room/20230210IPR74715/fit-for-55-zero-co2-emissions-for-new-cars-and-vans-in-2035

Friese, P. A., Michalk, W., Fischer, M., Hardt, C., & Bogenberger, K. (2021). Charging Point Usage in Germany – Automated Retrieval, Analysis and Usage Types Explained. *Sustainability*, *13*(23), 13046. https://doi.org/10.3390/su132313046

Hecht, C., Das, S., Bussar, C., & Sauer, D. U. (2020). Representative, empirical, real-world charging station usage characteristics and data in Germany. *eTransportation*, *6*, 100079. https://doi.org/10.1016/j.etran.2020.100079

Karanam, V., & Tal, G. (2023, June 11-14). *Developing a Deep Learning Tool to Detect Electric Vehicle Supply Equipment Failures* [Paper presentation]. 36th International Electric Vehicle Symposium and Exhibition (EVS36), Sacramento, California, USA.

Lévay, P. Z., Drossinos, Y., & Thiel, C. (2017). The effect of fiscal incentives on market penetration of electric vehicles: A pairwise comparison of total cost of ownership. *Energy Policy*, *105*, 524–533. https://doi.org/10.1016/j.enpol.2017.02.054

Li, X., Wang, Z., Zhang, L., Sun, F., Cui, D., Hecht, C., Figgener, J., & Sauer, D. U. (2023). Electric vehicle behavior modeling and applications in vehicle-grid integration: An overview. *Energy*, *268*, 126647. https://doi.org/10.1016/j.energy.2023.126647

Liu, Y., Francis, A., Hollauer, C., Lawson, M. C., Shaikh, O., Cotsman, A., Bhardwaj, K., Banboukian, A., Li, M., Webb, A. K., & Asensio, O. I. (2023). Reliability of electric vehicle charging infrastructure: A cross-lingual deep learning approach. *Communications in Transportation Research*, *3*, 100095. https://doi.org/10.1016/j.commtr.2023.100095

Metais, M., Jouini, O., Pérez, Y., Berrada, J., & Suomalainen, E. (2022). Too much or not enough? Planning electric vehicle charging infrastructure: A review of modeling options. *Renewable & Sustainable Energy Reviews*, *153*, 111719. https://doi.org/10.1016/j.rser.2021.111719

Mortimer, B. J., Bach, A., Hecht, C., Sauer, D. U., & De Doncker, R. W. (2022a). Public Charging Infrastructure in Germany–A utilization and Profitability analysis. *Journal of Modern Power Systems and Clean Energy*, *10*(6), 1750–1760. https://doi.org/10.35833/mpce.2021.000181

Mortimer, B., J., Hecht, C., Goldbeck, R., Sauer, D. U., & De Doncker, R. W. (2022b). Electric Vehicle Public Charging Infrastructure Planning Using Real-World Charging Data. *World Electric Vehicle Journal*, *13*(6), 94. https://doi.org/10.3390/wevj13060094

Pevec, D., Babic, J., Carvalho, A., Ghiassi-Farrokhfal, Y., Ketter, W., & Podobnik, V. (2019). Electric Vehicle Range Anxiety: An Obstacle for the Personal Transportation (R)evolution? *2019 4th International Conference on Smart and Sustainable Technologies (SpliTech)*. https://doi.org/10.23919/splitech.2019.8783178

Pevec, D., Babic, J., Kayser, M. A., Carvalho, A., Ghiassi-Farrokhfal, Y., & Podobnik, V. (2018). A data-driven statistical approach for extending electric vehicle charging infrastructure. *International Journal of Energy Research*, *42*(9), 3102–3120. https://doi.org/10.1002/er.3978

Philipsen, R., Brell, T., Brost, W., Eickels, T., & Ziefle, M. (2018). Running on empty – Users' charging behavior of electric vehicles versus traditional refueling. *Transportation Research Part F-traffic Psychology and Behaviour*, *59*, 475–492. https://doi.org/10.1016/j.trf.2018.09.024

Razmjoo, A., Ghazanfari, A., Jahangiri, M., Franklin, E., Denai, M., Marzband, M., Astiaso Garcia, D., & Maheri, A. (2022). A Comprehensive Study on the Expansion of Electric Vehicles in Europe. *Applied Sciences*, *12*(22), 11656. https://doi.org/10.3390/app122211656

Rempel, D., Cullen, C., Bryan, M. M., & Cezar, G. (2022). Reliability of open public electric vehicle direct current fast chargers. *Human Factors*. https://doi.org/10.2139/ssrn.4077554

Salah, K. B., & Kama, N. (2016). Reducing Range Anxiety by Unifying Networks of Charging Stations. *MATEC web of conferences*. https://doi.org/10.1051/matecconf/20167004003

Sankaran, G., Venkatesan, S., & Prabhahar, M. (2020). Range Anxiety on electric vehicles in India -Impact on customer and factors influencing range Anxiety. *Materials Today: Proceedings*, *33*, 895–901. https://doi.org/10.1016/j.matpr.2020.06.423

Statista. (2023, February 2). *Belgium: number of electric passenger cars on the road 2008-2020.* Retrieved February 21, 2023, from https://www.statista.com/statistics/698294/number-of-electric-passenger-cars-in-use-in-belgium/

Statista. (n.d.). *Electric Vehicles - Worldwide*. Retrieved February 25, 2023, from https://www.statista.com/outlook/mmo/electric-vehicles/worldwide

Statista. (n.d.-a). *Electric vehicles - Belgium*. Retrieved February 25, 2023, from https://www.statista.com/outlook/mmo/electric-vehicles/belgium

Straka, M., De Falco, P., Ferruzzi, G., Proto, D., Van Der Poel, G., Khormali, S., & Buzna, Ľ. (2020). Predicting popularity of electric vehicle charging infrastructure in urban context. *IEEE Access*, *8*, 11315–11327. https://doi.org/10.1109/access.2020.2965621

Stosur, M., Kuniewski, M., Sowa, K., Blaszczyk, P., & Uramek, A. (2020). Analysis and modelling of failure states in electric vehicle charging infrastructure. *Przeglad Elektrotechniczny*, *96*(7), 13-18. https://doi.org/10.15199/48.2020.07.03

United Nations. (n.d.). *Net Zero Coalition.* Retrieved February 24, 2023, from https://www.un.org/en/climatechange/net-zero-coalition

Van Den Hoed, R., Helmus, J., De Vries, R., & Bardok, D. (2013). Data analysis on the public charge infrastructure in the city of Amsterdam. *World Electric Vehicle Journal*, *6*(4), 829–838. https://doi.org/10.3390/wevj6040829

van der Kam, M., van Sark, W., & Alkemade F. (2020). Multiple roads ahead: How charging behavior can guide charging infrastructure roll-out policy. *Transportation Research Part D: Transport and Environment*, *85*, 102452. https://doi.org/10.1016/j.trd.2020.102452

Weekx, S., Tal, G., & Vanhaverbeke, L. (2024). How many more public charging stations do we need? A Data-Driven approach considering charging station overflow dynamics. *Transportation Research Record*. https://doi.org/10.1177/03611981241242081

Wolbertus, R., & Van Den Hoed, R. (2017). Charging station hogging : a data-driven analysis. *The 30th International Electric Vehicle Symposium & Exhibition*.

Zhang, J., & Wozniak, C. (2017). EV charge station system reliability Modeling and Analysis. *SAE Technical Paper Series*. https://doi.org/10.4271/2017-01-0198

13.) Appendix

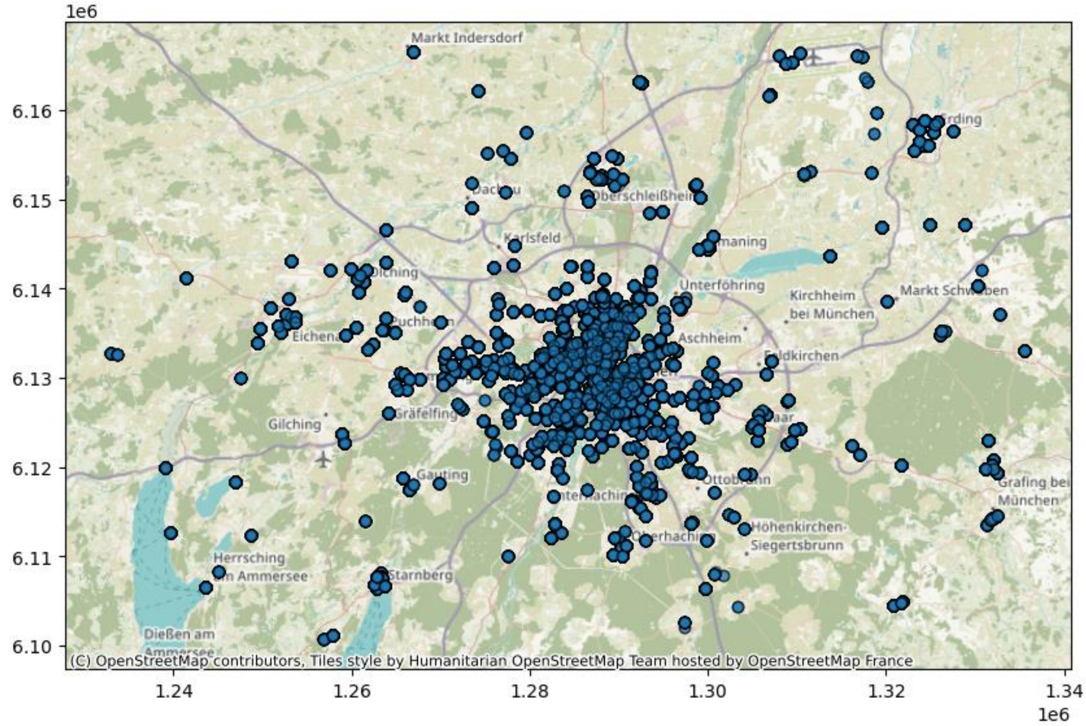


Figure 13. Geographical location of the observed EVSEs (Own creation)

Parameter	Description	Value	Source
max_d	Maximum distance nearby EVSE	250 meters	(Csiszár et al., 2019)
z_score	Deviations from mean	1,96	Own decision
thld_CSH	Threshold after which CE is categorized as CSH	16 hours	(Wolbertus & Van Den Hoed, 2017)
thld_charging_attempt	Threshold before CE is categorized as a successful CE	2 minutes	(Rempel et al., 2022)
thld_weekly_nr_CEs	Threshold average weekly CEs at EVSE	3 CEs	Own decision

Table 12. Overview chosen parameters in the framework (Own creation)