Drivers of Electric Vehicle Adoption
a Case Study at Public Chargers in the Netherlands

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Engineering & Policy Analysis**

Faculty of Technology, Policy and Management

by

Dewy Koopman

Student number: 5402549

To be defended in public on March 15th 2023

Graduation committee

Chairperson	: Dr. ir. M. Kroesen, Section Transport and Logistics
First Supervisor	: Dr. ir. M. Kroesen, Section Transport and Logistics
Second Supervisor	: Dr. ir. C. van Daalen, Section Policy Analysis
Advisor	: Ir. M.L. van der Koogh, Section Energy and Industry

Executive Summary

The damaging effects of climate change drive the transition to a more sustainable world. In recent decades, advances in technology integrate our daily life with more sustainable techniques and products. The introduction of these products is frequently supported by policies that enforce or support the choice of the sustainable alternative above the prevailing choice. However, in most cases, these transitions do not occur overnight. This holds as well for the adoption of the electric vehicle. This modal shift is in full swing, where the usage of the electric vehicle over the 'old' internal combustion engine-powered vehicle will reduce the emission of green house gases enormously. This modal shift has a long time span, since the electric vehicle depends on the charging infrastructure that can charge these vehicles. The implementation of the charging infrastructure is a time-consuming task, and therefore an established problem in research.

In the Netherlands, municipalities are responsible for the placement of public charging points. Public chargers are charging points that can be used by anyone with the correct registration identifier and are placed (mostly) on public land. To determine the locations of these chargers municipalities make use of two strategies. They are placed either on demand or by strategy. Placement-on-demand occurs when a citizen requests the placement of a public charger at a certain location and the municipality approves either this or an adjusted version of the proposal. Additionally, the capacity can be expanded when the municipality sees that the existing infrastructure cannot meet demand. Strategic placement means that the charger locations are determined by the municipalities itself, for example, at locations to expose the public to electric driving.

Research showed the importance of the feedback loop between the placement of the charging infrastructure and the adoption of electric vehicles for the modal shift. After 2030, only new cars with zero emissions can be sold in the Netherlands. This will cause an increase in the demand for public chargers, which all have to be placed in the coming years. This makes one wonder how and where these public chargers need to be placed. This is the research problem that is addressed in this study.

It is important that the placement of public chargers takes place in an optimal way. Not only to meet the capacity demand later, but also to stimulate adoption during the near future. This makes it important to understand the adoption of electric vehicles at public charging points and its relations with the characteristics of its surroundings. Furthermore, research on the relationship between the adoption of electric vehicles and socioeconomic characteristics is suggested in the literature due to the existence of a research gap. Combining this with the social relevance, a clear research question arises: what are the relations between electric vehicle adoption at public charging points and neighbourhood characteristics? This study focusses on answering this research question to reduce the academic research gap by identifying relevant relations between the adoption of electric vehicles and the characteristics of the neighbourhood through a case study in the Netherlands. Furthermore, these relations will be used to identify policy recommendations with a focus on improving the rollout strategy of public chargers in the Netherlands, if possible.

To achieve such results, the status quo of electric driving, charging infrastructure, the relevant actor arena, and existing policy in the Netherlands must be established in order to contribute. This study does such analyses in such a way that a framework is constructed in which later results can be evaluated. Furthermore, an overview of relevant variables for the modelling of electric vehicle adoption on neighbourhood level will be created. This was done through a systematic review of the literature. Furthermore, during the literature review process, an analysis of mathematical modelling techniques used in the field of EV adoption was performed. Together with an analysis of the modelling of EV adoption in the Netherlands, this created a scope for the quantitative study. Based on the literature review, statistical modelling was evaluated as an appropriate modelling approach, with regression analysis as the most promising technique. Furthermore, a relevant case study in the Netherlands showed how EV adoption at public chargers could be measured, namely by making use of transaction data.

The quantitative study started with creating such measurements of EV adoption in the transaction data of public chargers in the MRA-Elektrisch region. This region includes municipalities in the provinces Noord-Holland, Flevoland, and Utrecht, without the cities of Amsterdam and Utrecht. The

first statistic is the EV users measurement. This describes the number of unique charge RFID's used in a neighbourhood, which show behaviour in their charging sessions as one would expect when they live in that neighbourhood. The second measurement is the average occupancy rate of public charging points in a neighbourhood, where an occupied charger was defined as a charger connected to a vehicle. This measures the demand for public chargers in a neighbourhood.

Subsequently, the relevant variables found in the literature study on the drivers of electric vehicle adoption have been translated into a set of neighbourhood characteristics of the MRA-Elektrisch region using open data sources. These variables were used, as independent variables, to explain both the EV adoption measurements, the dependent variables. For the occupancy rate, a multiple linear regression model was created to find relationships with the characteristics of the neighbourhood. For the EV users measurement, this was done by making use of a zero-inflated negative binomial regression model. With these models, significant relations could be detected between the independent and dependent variables. For significant relationships, confidence intervals were constructed for the coefficient that describes the relationship. In this way, more robust conclusions could be drawn on these relations.

The relations found between the adoption of electric vehicles at public charging points and neighbourhood characteristics can be best described by the following description. EV adoption at public charging points is higher in neighbourhoods with more rental homes, smaller households, and more urban characteristics such as a high number of facilities. The residents in these neighbourhoods are younger in age, well-educated and have high incomes. In addition, residents are more likely to be self-employed and have green views. These relations are in accordance with literature, except for the relations between of household size and rental homes, which have an opposite relation with what one would expect based on previous studies.

On the basis of the found relations, a further analysis was conducted in order to develop policy recommendation for the MRA-Elektrisch region. Using the significant relationships that emerged from the models, the potential number of EV adopters for all neighbourhoods was estimated based on the neighbourhoods characteristics. The neighbourhoods were evaluated as high potential when the potential number of EV users was greater than the current number of EV users. The identified neighbourhoods with high potential and high demand for current charging infrastructure are suggested to be prioritised in the placement of public chargers in the near future. Identified neighbourhoods with high potential and a ratio of (potential) EV user per charger that is high under current conditions were suggested to prioritise in the placement of public chargers in the future.

Based on these findings, recommendations were made for the general rollout strategies of municipalities for public chargers, and for the rollout strategies of public chargers of municipalities in the MRA-Elektrisch region. The first general recommendation is to recognise the target audience of public charging stations and that they differ from the more recognised EV user who uses private chargers. Furthermore, EV adoption at public charging stations was found to be heterogeneous, making the placement strategy on demand appropriate. The strategy can be improved by adding a potential analysis to each request so that capacity can be increased when a location has a high potential. This will save resources in a later stage. For the strategic rollout strategy, the recommendation for a similar potential analysis was made to establish whether the location has sufficient demand. Furthermore, based on the identified relations between the adoption of EVs and public points, it appears that the strategic placement strategy is also appropriate.

For municipalities in the MRA-Elektrisch region, additional recommendations are made. It is suggested that these municipalities increase the capacity of public chargers in high-populated regions, as both the relative and absolute numbers of residents were positively correlated with the occupancy rate. This indicates that public charging points are not yet evenly distributed over the population. Furthermore, it is recommended to prioritise the neighbourhoods with high potential and high demand on the current charging infrastructure for the placement of additional chargers in the near future. Furthermore, it is recommended to increase the capacity of planned chargers in the neighbourhood with high potential and a high potential EV user per charger ratio.

Although the relationships found between EV adoption and neighbourhood characteristics are based on confidence intervals of the coefficients of significant relations, this does not imply that these relationships are actual drivers in EV adoption. The relationships found are based on correlations, which are not causations. Furthermore, the results of the quantitative study are obtained using models and no model is correct. This study simply presents the results found using two statistical models. Different choices in models and assumptions could have led to different results. Moreover, this study only focusses on the adoption of electric vehicles at public charging points and neglects the major part of the electric vehicle driving population who uses private chargers. This is because no data are available for this group. Furthermore, one of the key assumptions of this study is that every EV driver uses only one unique registration key. Another point of discussion is the analysed region, which covers only a part of the Netherlands. This MRA-Elektrisch region was found to be an unrepresentative sample for the rest of the Netherlands, which makes that the conclusions of this study have to be put into perspective. However, this study still provides useful insights for understanding the adoption of electric vehicles at public charging points. Therefore, this study must be perceived as the best possible approximation of the relations at this moment in time.

Preface

This master thesis concludes an intense period of study at the higher education institutions in the Netherlands and Australia. This is therefore both a moment to celebrate and a moment of reflection on the past years. During my studies, I got the opportunity to study at three different universities and pursue three different degrees. Each of them, unique in its own way and contributing to knowledge, skills, and personality. Although I did not think this during exam weeks or on the eve of larger deadlines, I will definitely miss this period. I think back on this period with warm feelings. The opportunities we have as students in the Netherlands are almost endless. This makes one grateful for the country one grew up in.

Of course, my study time was a period of ups and downs. During the more intense periods, I could always turn to the people around me for help, support, or the necessary distraction. Therefore, I would like to make use of this opportunity and thank my family and friends for their contribution not only in these study related moments but even more for the other moments in life. Thank you!

Then about this thesis. Although a thesis is perceived as an individual work, the result depends for a great deal on the cooperation and interaction with the thesis committee. I can say that when this team is approachable and supportive, this will have a positive influence on the course of events and on the end product. Therefore, I would like to thank my thesis committee consisting of Maarten Kroesen, Els van Daalen, and Mylène van der Koogh for their level of support and involvement. Each of them with their own expertise and knowledge, but all with the shared factor of willingness to make time. Furthermore, the combination of this committee made that at any moment at least one of them could think along on the research process. I really appreciated this. In special, I would like to thank Mylène van der Koogh for the time she made available to discuss the progress of the study and the opportunities given to make this thesis a success.

Dewy Koopman
Amsterdam, 28-02-2023

Contents

List of Figures

List of Tables

1 Introduction

Climate change drives governments and organisations around the world to implement more sustainable initiatives. To encourage these initiatives, policies are developed at international and national levels to enforce the reduction of greenhouse gas emissions. At the international level, the Paris Climate Agreement of 2015 states the goal of limiting global warming to a maximum of two degrees Celsius, with the aim of 1.5 degrees Celsius (Agreement, 2015). Participants in this United Nations treaty were obligated to develop a national climate policy. Based on this treaty, the European Commission approved the Green Deal of 2020 as the European reaction to this UN treaty. This Green Deal enforces the reduction of emission levels compared to 1990 by 55% in 2030 and climate neutrality in 2050 (European Commission, 2020). The implementation of these goals is described in the Fit For 55 plan (European Commission, 2021).

Based on the Paris Climate Agreement, the Dutch government introduced the Klimaatwet of 2019. This Dutch law enforces a reduction of 49% of the exhaust levels of 1990 in 2030, and 95% in 2050 (Rijksoverheid, 2019b). To achieve this, the Klimaatakkoord of 2019 states the plan for the five sectors; electricity, industry, traffic and transport, agriculture, and built environment (Rijksoverheid, 2019a). The first goal has already been achieved by meeting the 25% reduction required in 2020. With a reduction of 25.4%, partially helped by the lockdowns of the Covid-19 epidemic, the Netherlands is on its way (Rijksoverheid, 2021).

However, big steps must be taken. Moreover, there are plans to update the Dutch Klimaatwet to the levels pursued by the European Union (Planbureau voor de Leefomgeving, 2021). In the coming years, great steps must be taken in each of the five sectors. One of these sectors is that of traffic and transport, in which 7,3 megatons of CO_2 has to be reduced in 2030 (Rijksoverheid, 2019a). One of the pillars in this sector's Klimaatakkoord is the modal shift towards electric driving. Zero emission should become the new standard in all forms of road transportation. This means that future cars, fans, buses, and trucks will be powered by either electricity or hydrogen. For each of these transport methods, different policies will apply. The biggest polluter in the transportation sector is the personal car. In 2019, 60.6% of the emissions from the EU transport sector could be allocated to cars (European Environment Agency, 2022a). Furthermore, in the Netherlands, about a fifth of total emissions can be allocated to cars (Rijksoverheid, 2022b). This makes the car a significant polluter and an important link in the modal shift. The Dutch government shares this vision by forcing zero emission for all new cars bought for personal use after 2030 (Rijksoverheid, 2019a). This would mean that after 2030 all new cars will be driven by hydrogen or electric power, where the latter has the predilection.

1.1. Research Problem

The modal shift to electric driving presents challenges. One of these challenges is the creation of a sufficient capacity charging infrastructure. It is expected that by 2030 already 1.9 million electric vehicles (EVs) will be driving in the Netherlands, and a total of 1.7 million charging stations will be needed (Rijksoverheid, 2022b). This means that in the coming years, the charging infrastructure in the Netherlands will have to undergo a severe reshaping. The plan for the rollout of this additional charging infrastructure has been written in the 'Nationale Agenda Laadinfrastructuur'(NAL, 2019).

In August last year, it was estimated that approximately 370 thousand charging points were active (NAL, 2022b). An estimated 270 thousand of these are private chargers that are not registered. The other hundred thousand charging points are (semi-)public chargers. These are chargers available for everyone. Since semi-public chargers are placed on private terrain, these will not be accessible 24/7 in most cases. In 2021, roughly 21 thousand of these (semi-)public charging points were placed. This comes down to eighty charging stations placed per working day. According to NAL, this number is expected to increase to eight hundred public charging points per working day in 2030 (Vermeij & Veger, 2019) .

So big steps need to be taken in the coming years for the (semi-public) charging infrastructure. Furthermore, the charging infrastructure has an additional role in achieving the modal shift. The available charging infrastructure can stimulate the adoption of electric vehicles (Javid & Nejat, 2017). This feedback loop between the electric vehicle adoption and the placement of charging infrastructure can be a well desired catalyst in the modal shift. Therefore, the optimal placement of the public charging infrastructure plays an important role.

Therefore, thousands of chargers need to be installed in the coming years to support the modal shift. Furthermore, the locations of these chargers have an effect on the progress of the modal shift. This leaves policy makers with the task of determining the locations of these public charging points and which to prioritise. The substantiation of this task is an important research problem in the stimulation of the modal shift.

1.1.1. Research Gap

In the Netherlands, the placement of public charging infrastructure is done by municipalities with the help of supporting actors. Municipalities decide on the locations of these public chargers by a strategic strategy or by a demand strategy. In the latter case, citizens can request the placement of a charging point through their municipality. When such requests meet the requirements, the placement will be planned. The requirements are for the municipalities to determine. The second way these chargers are placed on demand is supported by data. When a municipality sees that the demand for a charging point is high, it can decide to place another charging point nearby or increase its capacity.

Determining the location of these charging points is an important and demanding activity. The modal shift would benefit from optimising these locations (Gupta et al., 2021). However, for such optimisation, a correct understanding of electric vehicle adoption at public charging points is crucial. This demands insights into the relations between the adoption of electric vehicles at public charging points and the characteristics of the neighbourhood. Research on such relations in the literature is limited, and a case study does not exist for the Netherlands. Furthermore, a recent systematic review of the literature on understanding EV adoption concluded that more information is needed on the socioeconomic factors driving EV adoption (Stockkamp et al., 2021).

Combining these two findings, it is clear that there is a research gap on the adoption of electric vehicles at public charging points and the relationship with the characteristics of the neighbourhood. This gap is illustrated in the literature by the demand for more insights in socio-economic factors driving the EV adoption, as in practise, where more understanding is desired for the optimal placement of public charging points. This research gap is the basis for this thesis study.

1.1.2. Research Objective

This master thesis will focus on reducing this research gap. It will explore the relationships between the characteristics of the neighbourhood and the adoption of electric vehicles at public charging points in the Netherlands. This will be done by answering the following research question:

> *"Which relations are there between neighbourhood characteristics and the electric vehicle adoption at public charging points in the Netherlands?"*

The results of this master thesis will be twofold. First, a contribution to the academic literature with relations found between neighbourhood characteristics and the adoption of electric vehicles around public charging points. Second, policy advice on the placement of public charging points for Dutch municipalities based on a further analysis using found relationships. The goal of this policy advice is to contribute to improving public charging point placement strategies.

1.1.3. Research Sub-Questions

The research questions will be answered by combining the answers of the following subquestions:

1. What is the status quo of electric vehicle adoption, charging infrastructure, actor arena, and its policy context in the Netherlands?
2. What are relevant neighbourhood characteristics for the understanding of electric vehicle adoption on neighbourhood level in the Netherlands?
3. How can electric vehicle adoption be measured on neighbourhood level?
4. What insights can be gathered about the relationship between electric vehicle adoption and neighbourhood characteristics in the Netherlands?

1.1.4. Research Approach

This study will begin by creating a relevant framework for the study, based on background information on the research problem. This framework will guide the rest of the study toward relevant insights and conclusions. It will hereby start by describing the status quo; of electric vehicle adoption, the charging infrastructure, actor arena, and the policy context in the Netherlands. This answers Sub-Question 1. This will be done by a comprehensive analysis of policy documents and research outlets.

Next, a literature study will follow to find relevant neighbourhood characteristics for understanding the adoption of electric vehicles. This will answer Sub-Question 2. Furthermore, techniques used in the field on modelling of EV adoption will be analysed in terms of suitability. And, as a third, the relevant research landscape on modelling of EV adoption in the Netherlands will be described. The literature study will be carried out using a systematic literature review approach, where possible. With the insights of this study, a conceptual model will be created that can be used to answer the research question.

Following this conceptual model, is operationalisation. The conceptual model will be translated into an operational model. One of the key challenges is the creation of a method to measure electric vehicle adoption at public chargers on neighbourhood level, which is Sub-Question 3. This will be done by creating a statistic on the adoption of electric vehicle users and a statistic on the occupancy of chargers using transaction data from public charging points. Both these statistics measure the adoption of electric vehicles in another dimension, where the first focusses on the adoption of citizens in the neighbourhood and the second focusses on adoption at the charging points in that neighbourhood. The combination of these two statistics makes it possible to make conclusions on adoption of electric vehicles at the neighbourhood level. The statistics will be created using definitions from the literature and data analysis approaches. The transformation will be clearly documented and every step taken will be explained.

The statistics will be combined with open data on the drivers of adoption of electric vehicles identified by the literature study of Sub-Question 2. This will complement the operational model, which will be used to find relations between EV adoption and neighbourhood characteristics, that is, Sub-Question 4. Furthermore, to answer this fourth sub-question, the results will be further analysed for the improvement of the rollout policy of public chargers. For this analysis, the results of all the sub-questions will be used.

1.1.5. Research Scope

This study focusses on the municipalities in the Netherlands as problem owner. Therefore, the conclusions in this study are based on the perspective of municipalities. However, most of the findings will also be relevant to other parties.

Furthermore, the adoption of electric vehicles analysed in this study is that of electric vehicle drivers who charge at public charging points near their home. This selection must be made to map the adoption of electric vehicles at public charging points to the characteristics of the neighbourhood of EV adopters who live in that neighbourhood. It should be clarified that this scope only includes part of the total adoption of electric vehicles in the Netherlands.

Furthermore, the adoptions of electric vehicles used in this study, as we shall see later, are those of charging points in the MRA-Elektrisch region. This includes the municipalities of Noord-Holland, Flevoland, and Utrecht without the cities of Amsterdam and Utrecht. Nonetheless, this study focusses on the Netherlands as a whole. The results of the analysis for the improvement of current policy will be more relevant for municipalities in the MRA-Elektrisch region, where the results of the relations will be more relevant for the academic world.

1.2. Social and Academic Contribution

This thesis is based on a clear academic and social research gap. The purpose of the study is to contribute to reducing this gap by finding the relations between the adoption of electric vehicles and the characteristics of the neighbourhood. It will thereby contribute to the scientific literature by exploring these relations in a case study in the Netherlands. The results will benefit the academic world by providing information on the adoption of electric vehicles at public charging points in one of the leading countries in electric driving. Furthermore, this study will contribute by providing a method to measure the adoption of electric vehicles at charging points.

A good application of the results of this research would be the neighbourhood prognoses of the Nationaal Agenda Laadinfrastrctuur (NAL). The NAL provides estimates for electric vehicle adoption and charging infrastructure at the neighbourhood level in the Netherlands for 2025 and 2030. These predictions are a combination of various modelling techniques, research, and data sources. The relevant output of this thesis could be used to further tweak these models and thus improve predictions.

Furthermore, this thesis will contribute by reflecting on the current policy in the Netherlands with the new information on the adoption of electric vehicles. These insights consist of both the found relations as the results of the further analysis for the improvement of policy, stating suggestions for the prioritising of neighbourhoods in the (near) future. This contributes to the rollout strategies of public charging points. And the presence of charging infrastructure can stimulate the adoption of electric vehicles, which makes the contributions of this study significant in the modal shift. And the sooner this modal shift is fulfilled, the less damage is done to nature.

1.3. Connection to Master Programme

Climate change is one of the grand societal challenges of our time (Voegtlin et al., 2022). Its complex and interconnected character makes this a wicked problem (Hermans et al., 2010). The magnitude and devastating effects of its consequences make the problem interwoven in most decisions made today. International and national policy arenas are occupied with the creation and implementation of policies to prevent climate change. The focus is on reducing emissions and using new techniques as clean alternatives. As in the modal shift towards electric driving.

This thesis research is related to this wicked problem and will try to contribute by improving the policy at the municipality level in the Netherlands. Using data analysis and (statistical) modelling techniques, new insights are created and transformed into policy recommendations. The combination of the research objective, the research problem, and the methods used position this study in the field of Engineering & Policy Analysis.

This thesis belongs to the Energy & Industry Section (E&I) of the Faculty of Technology, Policy and Management (TPM) of Delft University of Technology. This research was carried out in close collaboration with the Future Charging project. This is a research project hosted by the Amsterdam University of Applied Sciences (HvA) in collaboration with research institutes, companies, and governing organisations and has the objective of contributing in the breakthrough of electric driving in the Netherlands.

1.4. Reading Guide

This report follows the structure of the research approach and sub-questions as discussed. The report will start with background information necessary to understand the research problem in Chapter 2. It will analyse the status quo in different dimensions. Thereafter, Chapter 3 will state the findings of the reviews of the literature on factors influencing the adoption of electric vehicles, the techniques used in the field, and the research landscape focused on the adoption of electric vehicles in the Netherlands. On the basis of these findings, the chapter will sketch a conceptual model. Chapter 4 will transform this conceptual model into an operational model. It will explain this transformation and the relevant methodology used. Following, Chapter 5 will summarise and analyse the main findings and answer the fourth sub-question. Furthermore, it will use these findings to improve current policy through a further analysis of potential charging station locations. Chapter 6 will discuss the research approach and reflect on the usability and relevance of these findings. Finally, Chapter 7 will conclude this research by reflecting on the research question and stating the findings. Furthermore, this chapter will give suggestions for future research.

2

Background Information

This chapter describes the relevant background information for the research problem in this study. The focus will be on the status quo of electric vehicle adoption, charging infrastructure, actor arena, and policy context in the Netherlands. These analyses create a guiding framework for the rest of this study so that the contribution of this study can be placed in context.

2.1. Electric Vehicle Adoption

The Netherlands is one of the leading countries in electric mobility. In September 2022, more than 300000 battery electric vehicles (BEV) were registered, which represents 3.04% of all cars in the Netherlands (Netherlands Enterprise Agency, 2022). Together with the 175000 plug-in hybrid electric vehicles (PHEV), 5.04% of Dutch passenger cars had a form of electric engine. This means that one in 20 cars in the Netherlands can use the public charging infrastructure. Worldwide, this ratio is around one in two hundred and fifty (EV-Volumes, 2022).

In the sales of new cars, the electric engine is even better represented. In September 2022, 24% of the new cars were a BEV, making the cumulative average for the year 20.7 % (Netherlands Enterprise Agency, 2022). The highest rate observed so far. PHEV models represented 11.9%, of total sales, bringing the share of electric engines to almost a third. These statistics belong to the best of Europe, where the Netherlands has the sixth highest share of electric engines in car sales behind the five Scandinavian countries (European Environment Agency, 2022b). Focussing only on the BEV, the Netherlands would take the third place after Norway and Iceland. However, in absolute numbers, these six top countries lose to France and Germany within Europe. Worldwide, China, Japan, and the United States are at the top of the list (EV-Volumes, 2022).

Not only passenger cars are converted to electric engines, so are other road transport vehicles. In the Netherlands, 1.17% of commercial vans have an electric engine. Of the commercial trucks, only 0.21% has an electric engine. These numbers are behind those of passenger cars, although the modal shift has begun. This study will not go into these categories.

The adoption of electric cars differs by region in the Netherlands. Figure 2.1 shows the adoption of electric vehicles per thousand citizens based on the residence-corrected electric vehicle registrations per province by CBS (2022b). It can be seen that Utrecht is the frontrunner, followed by its surrounding provinces.

Every year, a survey of electric vehicle drivers is conducted in the Netherlands: the Nationaal Laadonderzoek (Wolterman et al., 2022). This survey is voluntary and, therefore, is not a good reflection of the EV driving population. However, this survey provides some insight. According to the survey of 2022 the electric vehicle driver could be best described by the following description. Electric vehicle drivers are mostly male and on average 50 years old. More than half of them are between 40 and 60 years old. Respondents came from all over the Netherlands, however, some correlation could be found with Figure 2.1. One-third of the participants obtained their electric vehicle through a personal purchase, and another third of the population leases their electric vehicle through their employer. The rest obtain their electric vehicle via other lease or purchase constructions. Positioning themselves on the Roger's adoption curve, almost 59% find themselves an early adopter (Rogers et al., 2014).

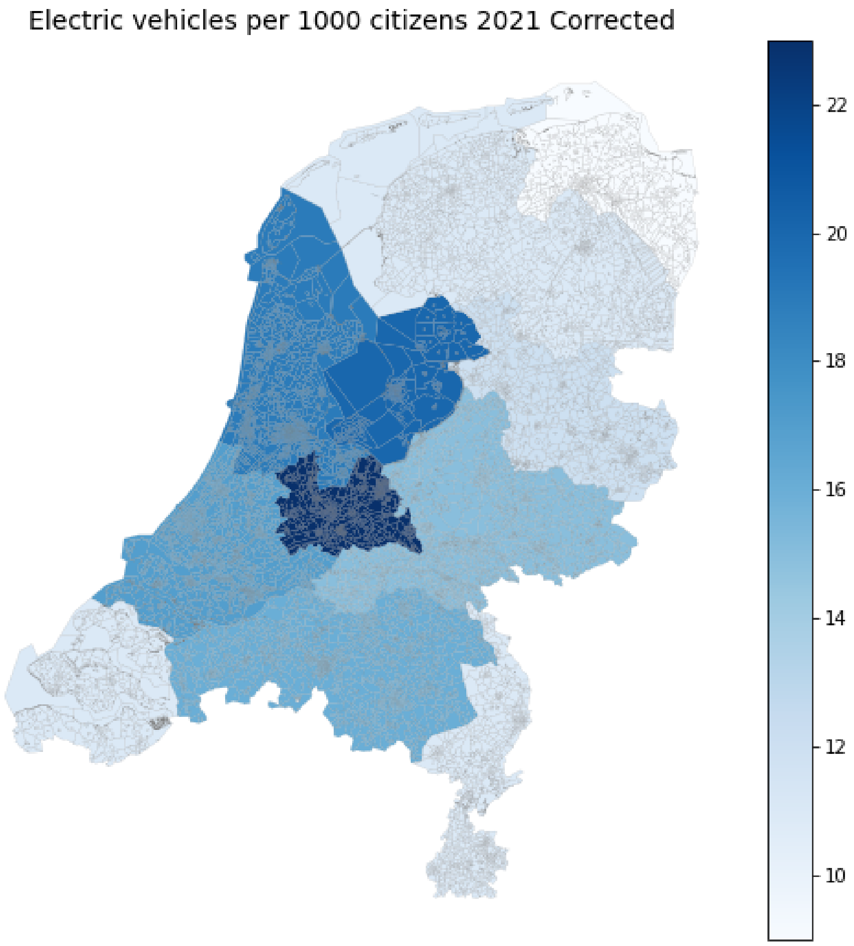

Figure 2.1: Electric vehicles per thousand citizens 2021, data source: CBS (2022b)

Another 22% identify themselves as innovators, and 13% would be part of the early majority.

2.2. Charging Infrastructure

The Netherlands has one of the best charging infrastructures for electric vehicles in the world (Funke et al., 2019). In absolute numbers, it was among the top four countries with the most public charging points after China, the United States, and South Korea in 2021 (International Energy Agency, 2022). In September 2022, the public charging point counter in the Netherlands reached 107,318 regular and 3,382 fast chargers (Netherlands Enterprise Agency, 2022). This means an average of 4.3 electric vehicles per charging point. Moreover, it is estimated that there are around 277 thousand private chargers (Wolterman et al., 2022). This would reduce the electric vehicle per charging point ratio below 1.5.

According to the National Laadonderzoek, two-thirds of the (observed) electric vehicle drivers have a private charger at home (Wolterman et al., 2022). In addition, having the opportunity to install a private charging point is a driving force behind the adoption of electric vehicles. The percentage of electric vehicle drivers who use public charging points is around 37% of which 71% occasionally have problems with charging at a these points (Wolterman et al., 2022) . These problems mostly concern unreachable charging points or an insufficient number of connections. However, the situation appears to improve, with 89% of electric drivers experiencing problems the year before (Wolterman et al., 2022).

Another point of concern for public charging points is the long queues for placing these points. Of the successful applications, it takes more than three months to place the charging point in 77% of the cases. In 45% of the applications, placement takes between 7 and 12 months. One in three applications is rejected by municipalities.

Succeeding Figure 2.1, the public charging points per thousand citizens are visualised at the provincial level in Figure 2.2 based on data from NAL (2022b). Figure 2.2 shows similarities with Figure 2.1 where most of the provinces with the highest adoption of electric vehicles also have a higher number of charging points. An exception to this is the province of Zeeland, where a relatively high number of public chargers per thousand citizens is available.

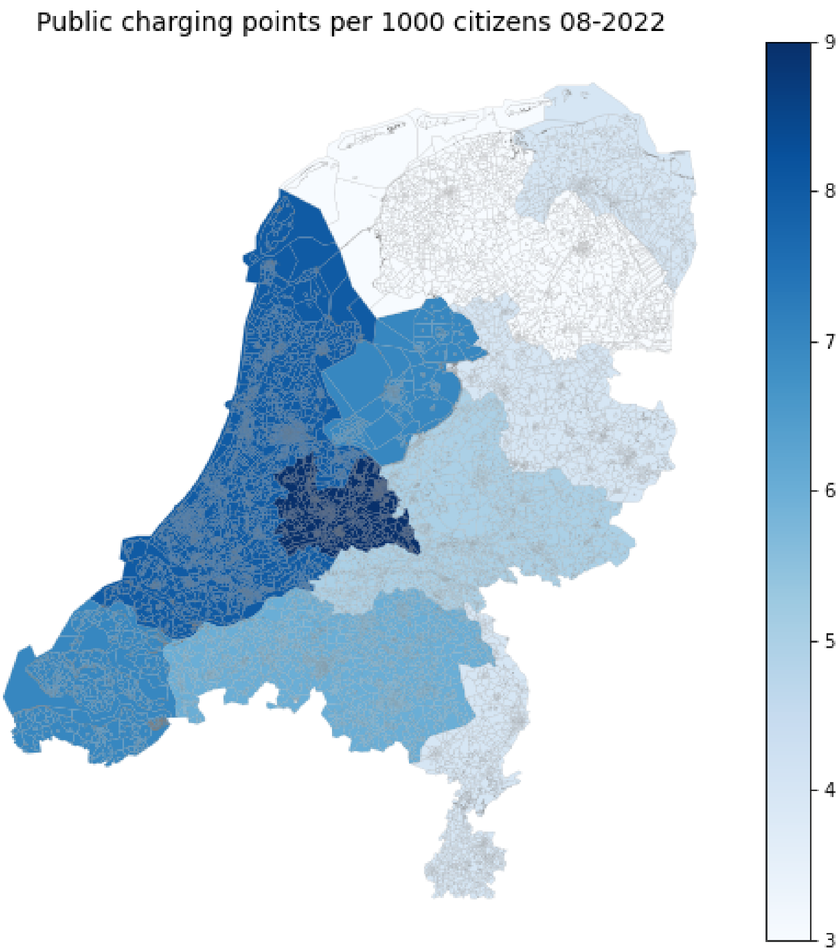

Figure 2.2: Public charging points per thousand citizens 08-2022, data source: NAL (2022b)

Public charging points are facilitated by municipalities in the Netherlands. Figure 2.2 shows the aggregated level of charging points. The same visualisation at the municipality level can be found in Figure A.1 in Appendix A. However, not all municipalities provide data which gives a distorted view. What can be concluded is that the ratio of charging points differs per municipality.

In most municipalities, it is possible to request a public charging point for citizens. This request is free, as is the placement and maintenance. Only the charging price per kWh can be higher than what one would pay using a private charger. Regulations apply to the request, which makes that not all of them are accepted, as we have seen in the previous section. More about the regulations will be provided in Section 2.3. When the municipality accepts a request, the municipality will find the optimal place to locate the charging point. The network operator (netbeheerder) will then create a connection to the electricity grid when this is not available. And a charging point operator (laadexploitant, CPO) will install and maintain the charging point.

Public charging points come in different types. One could distinguish the categories of normal and fast chargers. Fast chargers are charging stations designed for rapid charging of electric vehicles. Such charging sessions aim to charge quantities enough to complete a trip. These fast chargers have a higher power than normal chargers, which reduces charging time. Fast chargers are mainly found next to highways in the Netherlands. In the category of normal public chargers, one can distinguish the placement of the charger. The public chargers are located on public land, which means that they are accessible at any time. Semi-public chargers are located on private land, which means that they do not have to be available 24/7.

A charging location is defined as a charging pool where one or more charging stations are available (Netherlands Enterprise Agency, 2019). Therefore, a charging pool has only one address. Each charging station present in the charging pool is a physical place where one or more charging points are available for one or more electric vehicles. The connection between the electric vehicle and the charging point is made using a connector. A charging point can have multiple connectors, however, in practise the number of connectors is often equal to the number of charging points (Netherlands Enterprise Agency, 2019). The composition of a charging pool is shown in a visualisation in Figure 2.3.

Four different categories of connections are distinguished:

- An outlet of the charging point where a cable needs to be plugged in.
- A cable attached to the charging point that must be plugged into the vehicle.
- A pantograph that is used to connect to a charging point above the vehicle. This is not common for personal EV's.
- An induction plate, which is used to charge without conduction. This is also not common for personal electric vehicles.

The first two categories of connections use the same connectors/ plugs to make the connection between the cable and the EV. However, there are different versions of these connectors and they are used. The plug to be used depends on the brand, model, and type of charging. As stated before, this type of charging can be distinguished into regular charging (less than 22 kW) and fast charging (more than (22 kW). Charging with powers higher than 150kW is called ultra-fast charging. The following types of connector are used in the Netherlands:

- Type 1, which is the standard Japanese connector for charging.
- Type 2, which is appointed by the Commission of the European Union as the standard connector.
- Combined charging system (CCS), which is an enhanced version of the type 2 connector with built-in support for fast charging.
- Chademo/ Type 4, is used for DC charging only and is used for fast charging.
- Tesla supercharger, used exclusively by Tesla vehicles.

Another classification can be made in charging points. The electrical grid provides alternating current, where electric cars operate on direct current. Therefore, the current needs to be transformed. This transformation can be performed by either the charging station or a converter in the electric vehicle. When the transformation is made by the converter in the electric vehicle, one speaks of AC charging. This is mostly the case for regular charging. When the transformation is made at the charging station, one speaks of DC charging, which is mostly fast charging. As one could imagine, the demand for fast charging is different from that for regular charging. Since fast charging involves higher investments, the

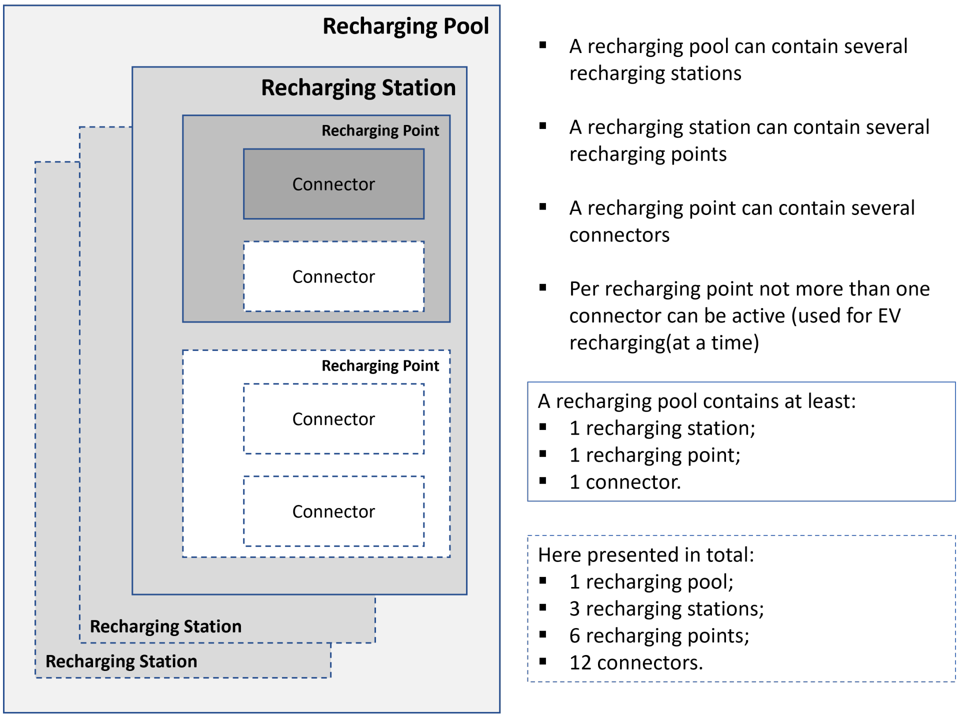

Figure 2.3: Public charger definitions, European Commission (2022)

locations of these charging stations are limited. In practise, this means that fast chargers are mainly located on highways.

The situation for private chargers is different. Here, the EV driver has to pay for the charging point and installation themselves. However, this gives the user the freedom to place the charger where they want in their own terrain, use them alone, and use electricity rates lower than those of public charging points. Private chargers are beyond the scope of this research and will not be considered further.

2.3. Policy

This section will focus on the policy applicable in the Netherlands. It will thereby scope down from international agreements to the policy on municipality level.

2.3.1. International Treaties

The Paris Agreement of 2015 is the first legal document that states the desire to limit temperature increases to a maximum of two degrees Celsius (Agreement, 2015). It states that participating countries must come up with national plans to combat climate change. As a result, both the European Union and the Dutch government developed a climate policy. The EU did this in 2020 with the Green Deal. The accomplishments of these goals are described in the Fit for 55 report (European Commission, 2021). In terms of personal transportation, this comes down to the following:

- In 2030 the emissions of new cars must be at least 55% lower than those of 1990.
- Complete phaseout of internal combustion engines in new cars from 2035.
- The introduction of a new tax system for fossil fuels.
- Creation of frameworks for the rollout of charging infrastructure on national levels where:
  - public charging points must be located every 60 kilometres on high ways,
  - the capacity of the infrastructure should provide at least 1 kW per electric vehicle.
- Improve the user experience of charging electric vehicles.
- Require smarter charging.

2.3.2. Klimaatakkoord

The Dutch government introduced its climate policy in 2019 under the name Klimaatakkoord (climate agreement). In this agreement, the Dutch government states the goal of reducing emissions in 2030 by 49% compared to 1990 levels (Rijksoverheid, 2019a). And in 2050, reducing 95% of the emmisions compared to 1990 levels. These goals differ from the aspirations of the European Union since the Klimaatakkoord was published before the Green Deal. However, the Dutch government is thinking about aligning its goals with those of the EU (Planbureau voor de Leefomgeving, 2021).

The Klimaatakkoord focusses on five different sectors. One of them, the mobility sector, will be responsible for the modal shift. The following regulations are applicable:

- In 2030 the use of cars for business purposes should be reduced by 8 billion kilometres.
- In 2030 all new cars sold should have zero emission.
- In 2030 around 1.8 million charge points will be installed.

2.3.3. Regeerakkoord

The Regeerakkoord is the coalition agreement of the Dutch government. In this agreement, the coalition parties present their plans for the coming reign. The latest Regeerakkoord is from 2021 and is constructed by VVD, CDA, D66, and ChristenUnie. In this coalition agreement, the following plans are stated focussing on electric driving:

- In continuation of the European Green Deal, reduce emissions to at least 55% compared to 1990 levels in 2030, with the aim of a 60% reduction.
- Support from the government to get the electric occasions market on the ground.
- The introduction of a pay-per-use system for cars after 2030.

2.3.4. Nationale Agenda Laadinfrastructuur

The Nationale Agenda Laadinfrastructuur (NAL) is the agenda for the implementation of the necessary charging infrastructure. This agenda is responsible for the rollout of the needed charging infrastructure in the Netherlands and is the Dutch response to the request for such a framework by the Green Deal. The NAL is a collaboration between (local) governments, research institutes, and grid providers. This collaboration is divided into six regions, which are:

- Noord-Holland, Flevoland, and Utrecht (MRA-Elektrisch).
- Groningen, Friesland, and Drenthe.
- Overijssel and Gelderland.
- Noord-Brabant and Limburg.
- Zuid-Holland and Zeeland.
- G4-cities: Amsterdam, Den Haag, Rotterdam, and Utrecht.

In these six regions, local governments and research institutes work closely together to implement the charging infrastructure. For example, think of joining tenders or discussions of guidelines. Furthermore, the NAL provides the agenda for the implementation of charging points. This plan is focused on the years 2025 and 2030. It provides prognoses of the number of charging points installed and the number of (semi-)public charging points that need to be placed per working day. These can be seen in Figures 2.4 & 2.5. The numbers provided are not a goal since the implementation of charging points depends on innovations (Vermeij & Veger, 2019). However, for comparison, the number of semi-public chargers in September 2022 is plotted in Figure 2.4, and the average number of charging points placed per working day in 2021 is plotted in Figure 2.5. It should be noted that there is a lag between these two periods of time.

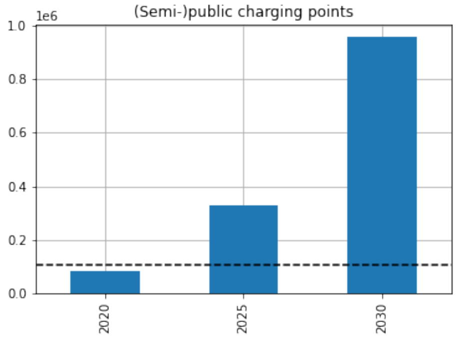

Figure 2.4: Semi(-public) chargers prognoses, data source: Vermeij and Veger (2019)

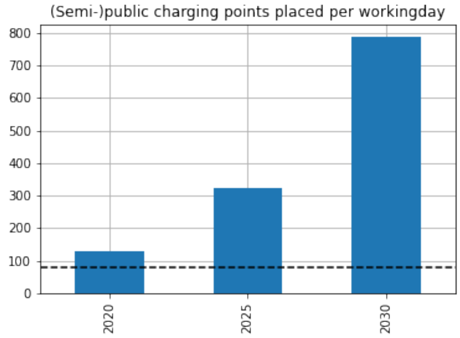

Figure 2.5: Semi(-public) chargers prognoses placement, data source: Vermeij and Veger (2019)

2.3.5. Policy of Municipalities

Municipalities are responsible for the rollout of public charging infrastructure. They are free to come up with their own policy, however, guidelines exist as we have seen in Section 2.3.4. The six regions assist municipalities in the placement of public charging points by preparing guidelines, managing offers, and creating discussions. The guidelines of the MRA-Elektrisch region are considered in this section, since this region is most relevant for the modelling, as we will see later on. Furthermore, the guidelines of this region are mentioned as an example in the NAL.

The 'Ladder van laden' (framework for charging) is a well-known concept within the implementation of the charging infrastructure. This concept creates hierarchy in types of charging infrastructure where it gives preference to private charging points, then semi-public charging points, and thereafter public charging points. The MRA-Elektrisch references to the newer 'Ladder van plaatsen' (framework for placement) which creates hierarchy in the rollout strategies for public charging points as in the following sequence:

1. Placement by demand based on requests of citizens.
2. Placement by demand based on data insights.
3. Placement by municipality strategy.

Although the requirements differ by municipality, the policies mention the same restrictions. This was evaluated by examining the policy of a selection of municipalities in the MRA-Elektrisch region, see Appendix A.1. These requirements can be summarised as follows, where 'x' indicates a value that changes per municipality:

- The applicant lives in the municipality or works at least x hours a week in the municipality, or does this within x months.
- The applicant drives an electric vehicle with a range of at least x kilometers.
- The applicant has no private parking place and has to make use of public parking spaces.
- Within a range of m metres from the application point no other public charging station is available or does not have enough capacity.

Some of the municipalities in the MRA-Elektrisch region work with preferred locations, which means that after the application of a citizen, the proposed location will be the location closest to the preferred locations requested by the citizen. Municipalities that do not use such locations search for preferable locations after receiving a request.

2.4. Actor Network Scan

Although municipalities determine the policy for the rollout strategies of public charging points, there are more actors in the context of the placement of a public charging point. This section will provide a

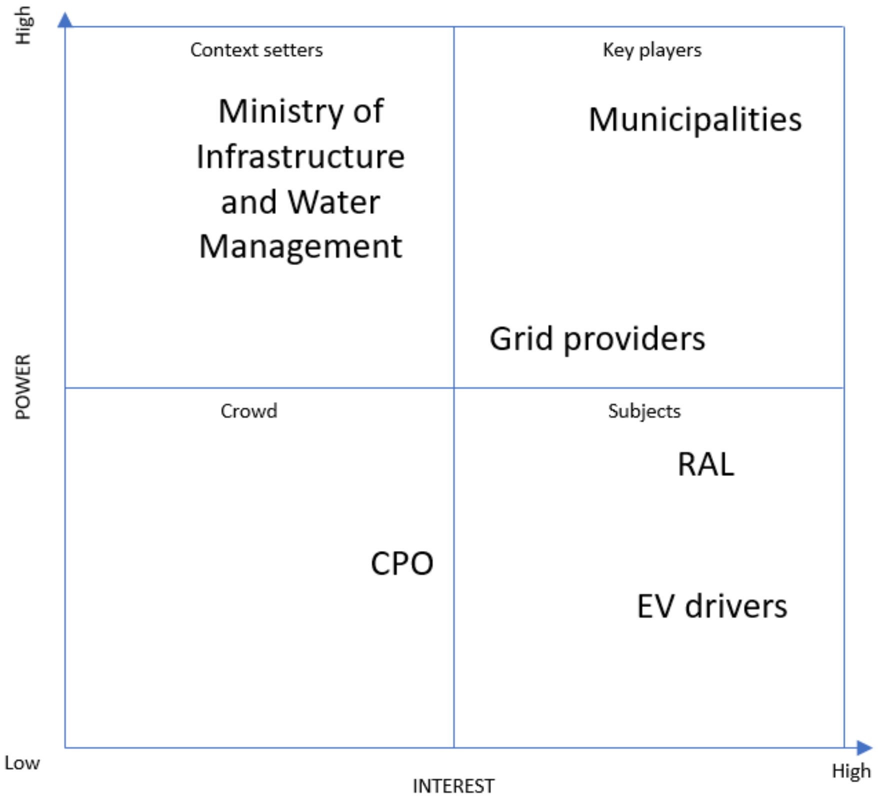

Figure 2.6: PI-grid of actors involved in the placement of public charging points

brief actor network scan to sketch the actors present in the field. This will be done by stating the actors and their interests and placing them in a power versus interest (PI) grid in Figure 2.6

Municipalities are the actors who are in charge of the public charger rollout policy. Furthermore, they decide whether or not a charging point is placed on public land. This also makes them the actor with the most power in the context arena. Furthermore, they have a high interest in this problem, since they are responsible for the rollout and want to provide its citizens. This is why municipalities are a key player and are located in the upper right corner. Since the 24^{th} of March 2022, the Netherlands has 324 municipalities.

Another actor in this context is the Regionale Agendas Laadinstrastructuur (RAL). As we have seen in Section 2.3.4, these regional collaborations of the NAL work together to contribute to the adoption of electric vehicles. This makes them a player with a high interest in the rollout strategies. The RAL has no significant power, however, they are there to guide and support the municipalities. Therefore, in this way, they help the keyplayer and have some indirect power.

Electric vehicle drivers are the end users of charging points. They will make use of the public infrastructure and therefore will be of great interest in the placement of these chargers. In most municipalities, drivers can request placement so that they have some influence on placement. However, this depends on the municipalities' regulations and whether the municipality works with preferred locations. Therefore, electric vehicle drivers are labelled as subjects as can be seen in Figure 2.6

Another actor class is the grid providers, also known as distribution system operators, who supply the energy on the electricity grid. A charging point without power will not work. These grid providers have as interest to secure the grid stability (Helmus & Van den Hoed, 2016). Therefore, they have low interest, since charging points must be placed and the energy usage will not change. Furthermore, they have no influence on the outcome other than not granting access to the grid due to capacity problems. However, in the Netherlands, capacity problems exist in large parts of the current grid infrastructure (Netbeheernederland, 2023). And these capacity problems make the stated goals as in earlier in this section unrealistic (Veeger, 2021). This makes grid providers a key player, since they have the power and motivation to block new connections. The grid providers active in the MRA-Elektrisch regions are Liander and Stedin.

The public charging points are placed and served by charging point operators. These operators try to monetise the investments in establishing and operating these charging points. The interest lies in the facilitation of a positive business model (Helmus & Van den Hoed, 2016). Therefore, they are interested in the placement of charging points but from a quantitative perspective. They have no power other than to not make a bid. They fall into the crowd, as can be seen in Figure 2.6. Some examples of CPO's in the Netherlands are Vattenfal and Eneco.

Not directly involved but an actor with great power is the national government represented by the Ministry of Infrastructure and Water Management. They have the power to change regulations around the placement of charging points that municipalities must follow. They have minor interest in the interpretation of the placement of the charging equipment, as long as it succeeds. Therefore, the Ministry of Infrastructure and Water Management is a context setter, as can be seen in Figure 2.6.

2.4.1. Conclusions Actor Network Scan

From this brief actor scan, it can be concluded that municipalities are key players in the placement of public charging points. Furthermore, under the current conditions of a grid capacity shortage, grid providers can also be marked as key players. Since the capacity of the grid is limited in very large parts of the Netherlands, the seven grid providers stated above have great power to deny new requests for grid connections. And this is one of the essential steps for the rollout of new charging points. Therefore, it is suggested that the close cooperation with the grid providers in the rollout is continued (as in the NAL) and that the municipalities actively support the upgrading of the electricity grid.

Although the municipalities are dependent on CPOs for the placement of public charging points, market forces will create a lot of supply under current conditions. This can change when the circumstances to exploit charging points change and the business model becomes less financial attractive. Therefore, it is important for the municipalities to acknowledge the financial motives of the CPO's and create circumstances where these CPO's are willingly to exploit.

As can be seen in Figure 2.6 the other high-power actor is the Ministry of Infrastructure and Water Management. However, since the interests of the Ministry and the municipalities are mainly aligned, this actor is viewed as a context setter. Therefore, in the rest of this research, it will be assumed that the actions taken by the municipalities will not create friction with the Ministry. For municipalities, it is important to stay up-to-date with the interest of the Ministry and create good feedback loops for when the motives of these two actors shift.

The other two remaining actors are the EV drivers and the RAL. Both actors have low power but high interest in this arena. Since municipalities are part of the RAL, it is assumed that motives and actions will align with the interests of municipalities. For electric vehicle drivers, this is different. Although electric vehicle drivers can request places for public charging points, the optimal placement in such a request does not always have to align with the optimal placement for the municipality. Therefore, some municipalities use preferred locations. Although this is a good way to safeguard the interests of the municipality, it is important to acknowledge the request made by the EV drivers and to assess the request with the motives of the citizen.

The conclusions made above will be taken into account in the further analysis to improve the rollout policy of public chargers.

3 Literature Review

This chapter will describe and analyse the relevant literature for this study using the systematic literature approach of Siddaway et al. (2019). First, the relevant modelling techniques will be discussed using the systematic review of the literature of Maybury et al. (2022) to obtain insight into the applicable approaches. Thereafter, relevant literature on the modelling of electric vehicle adoption in the Netherlands will be discussed. For this analysis, no systematic literature review was found to be available. Therefore, a review was conducted using mostly grey literature. Hereafter, the focus was placed on identifying relevant neighbourhood characteristics for the understanding of electric vehicle adaption on neighbourhood level, that is, Sub-Question 2. This analysis will be performed based on the systematic literature review of Austmann (2021) on the drivers of electric vehicle adoption used in empirical research. The chapter will conclude with a conceptual model in which the findings of the previous sections will be taken together.

The review of the literature of this study is based on the approach suggested by Siddaway et al. (2019). One of the first steps is clarifying whether systematic reviews have been done already. This was established by making use of the query ”systematic literature review” AND ”electric vehicle adoption” for the title, abstract, and keywords in the Scopus database and in general for the Google Scholar database. This is because the minimum number of online search engines consulted should be at least two (Siddaway et al., 2019). As an additional selection criterion, only articles after 2012 were included, since systematic literature reviews are valid for a period of 10 years (Siddaway et al., 2019). The search approach was carried out the latest on 27-12-2022 and included the systematic literature review of Maybury et al., 2022 which could be used to identify modelling techniques and that of Austmann (2021) on electric vehicle adoption drivers used in modelling approaches. This provides two of the three analyses with a foundation by existing systematic literature reviews.

For the second analysis, focussing on the modelling of the adoption of electric vehicles in the Netherlands, no sufficient results were encountered during this first attempt. Therefore, a further literature study has been conducted. During the phase of ’becoming familiar with the literature’, it became clear that the desired literature in this field was difficult to obtain through formulated search queries in the English language using the two mentioned search engines (Siddaway et al., 2019). This came as no surprise, since the systematic review of the literature Maybury et al. (2022) already showed that the results of studies focussing on the Netherlands are limited. Therefore, an alternative approach was used, with a focus on the grey literature and studies reported using the Dutch language. As a consequence, Google Search Engine was added to the selection of search engines. In addition, the information channels of the actors were considered as stated in Section 2.4 were considered.

3.1. Modelling of Electric Vehicle Adoption

The systematic review of the literature on the modelling of electric vehicle adoption has been based on the fifth research question of the systematic literature review of Maybury et al. (2022) focussing on the techniques used for modelling. Due to the recent publication date, this systematic review of the literature was found to be suitable for this analysis, that is, the systematic review of the literature was performed within ten years and no flaws could be observed (Siddaway et al., 2019). To be complete, the complete study landscape of articles citing this systematic literature review has also been analysed. This was done for the last time on 27-12-2022.

Let us first analyse the place of this study in the field of mathematical modelling of electric vehicle adoption. Most mathematical modelling in the field of electric driving nowadays focusses on finding factors driving EV adoption (Maybury et al., 2022). The focus of these articles, however, can be on a variety of stakeholders, such as car sellers (Shankar & Kumari, 2019) or car manufacturers (Mulholland et al., 2018). The majority focusses on the problems of policymakers, as does this study. The scope of this range of papers is still very broad. A subgroup focusses on implemented policies and their effect on the adoption of electric vehicles in and between countries. Yao et al. (2020) found that charger density among other policy instruments is a good predictor of the adoption of electric vehicles in 13 countries, including the Netherlands, using a multiple linear regression analysis. Another subgroup focusses on identifying barriers to the adoption of electric vehicles. Relevant articles are those of Smith et al. (2017) and Tal et al. (2018) that match barriers in EV uptake to parameters as sociodemographic variables using either a logistic regression or a discrete choice model, respectively. These papers show that the variables that are taken into consideration do not have to be positively correlated with the adoption of electric vehicles, as negative correlations also describe characteristics. The last subgroup is focused on the implementation of public charging points. Focussing on both the quantity aspect; the total number of charging points needed, calculated by the predictive models in Dong (2018). As the quality aspect of the rollout; the optimal placement of charging points as in, e.g., Tian et al. (2019) using prediction models and a multi-objective programming model for optimisation. This study will focus on both these aspects by identifying the relations between EV adoption and neighbourhood characteristics.

Now the focus will be on the more frequent modelling techniques in the field. A common method of modelling used for the understanding of electric vehicle adoption is that of system dynamics (SD) (Maybury et al., 2022). SD lends itself for the modelling of complex systems and the interaction between variables and the environment (Richardson, 2011). In this category of studies, a similar classification of the stakeholder can be made; SD papers are either 'automotive industry-oriented' or 'public policy-oriented' (Vilchez & Jochem, 2019). Within the public policy-orientated field of SD, research is focused on policy questions to support the adoption of electric vehicles. In SD models, sociodemographic variables at low aggregated levels are less represented, although their importance is observed (Ben-Akiva et al., 1985). Therefore, finding relations between EV adoption and sociodemographic variables at the neighbourhood level uses less of the SD potential. This method will not be considered in this study.

Another more widely used technique is agent-based modelling (ABM) (Maybury et al., 2022). In ABM models, the systems are disaggregated into individual components, agents, which have their own characteristics and rules of behaviour (Crooks & Heppenstall, 2012). In this way, experiments can be done on the individual choice level with interaction of (potential) EV adopters. This level of aggregation is too detailed for this study, since the focus here lies on finding relations on the neighbourhood level. Therefore, this technique will not be considered for modelling in this study.

Business models are another class of modelling techniques distinguished by Maybury et al. (2022). Business models are expected more in corporate context, however, the referred business models use scientific substantiation to contribute to the greater good. For example, Nian et al. (2019) provide a business model to encourage the adoption of electric vehicles in the absence of policy support in Singapore. Although the final conclusions of this study could be relevant for a business model, this kind of modelling is not suitable for finding relations between neighbourhood characteristics and the adoption of electric vehicles. Therefore, business models will not be considered in the rest of the study.

The most frequent modelling technique for electric vehicle adoption is a discrete choice model according to Maybury et al. (2022). This type of modelling lends itself for the prediction between two or more choices. In the field of electric vehicle adoption, this contains mostly the choice for electric adoption. Combining this technique with other statistical techniques distinguished by Austmann (2021) we can describe the last and most promising group of mathematical models for this study; statistical models. Although these models are static and, therefore, not suitable for describing dynamic processes, they are appropriate for making conclusions using significance levels and correlations between data sets.

The field of statistical models used in the study of the adoption of electric vehicles can be divided into five subclasses (Austmann, 2021). These classes are the field of spatial analysis, panel regression, sample difference, ordinary least squares regression, and a rest group focused mainly on correlation and covariance statistics. The most used methods come from the fields of spatial analysis, such as

T. D. Chen et al. (2015) or panel regression, such as Vergis and Chen (2015). Spatial analysis lends itself for the modelling of data with spatial characteristics, such as locations, making it extremely suitable for the modelling of regional differences. Panel regression is the application of regression techniques such as logistic regression or a generalised method of moments regression. The choice of statistical technique depends mainly on the available data. In this case, most of the techniques described in the five subclasses can be used to investigate the relationships between variables and will apply.

In conclusion, many modelling techniques are used in the field of electric vehicle adoption. Frequent used techniques are SD modelling and ABM. Both of these do not match the aggregation level of this study. Furthermore, business models are not relevant for this study as well. On the basis of the literature review, it is suggested to make use of a statistical method to answer the research question, since this method lends itself to find significant relations in static data. A reoccurring technique in distinguishing drivers of electric vehicle adoption is regression. However, which statistical model to use depends on the context will be further discussed in Chapter 4.

3.1.1. Electric Vehicle Adoption Research in The Netherlands

For the analysis of the modelling of electric vehicle adoption in the Netherlands, the choice was made to take an alternative approach, as that of the systematic literature review approach of Siddaway et al. (2019). Initial attempts did not give the desired results. These desired studies on the adoption of electric vehicles in the Netherlands belong mainly to the so-called grey literature, which is the literature not published by commercial publishers (Siddaway et al., 2019). The scattered field of researchers and the lack of a central publishing hatch in the Netherlands make the envisioned systematic approach not applicable. Therefore, the literature review conducted is not systematic. However, this does not mean that it has to be written off. On the contrary, the literature review has been performed by an extensive evaluation of research sources in the Netherlands. This search has been performed in both Dutch and English languages. A relevant selection has been made, which will be stated in the following. Although this research has been carried out with the greatest care, the unavoidable conclusion is that it did not analyse the entire study landscape due to the scattered character of this. However, the best attempt has been made.

The Netherlands is one of the leading countries in the adoption of EV, as we have seen in Chapter 2. It is not surprising that much research has been done on the adoption of electric vehicles due to the presence and interest in electric vehicles in the Netherlands. However, the acknowledgement of research focused on the Netherlands in literature reviews is limited. In the overview of (mathematical) modelling on EV adoption, only two the Netherlands focused papers are analysed. The first study provides a multilevel perspective on low-carbon transitions in mobility, with electric vehicles as a technology replacement (Köhler et al., 2020). The other provides a business case model for electric vehicles in the Netherlands (Wesseling et al., 2020). In other reviews, no special emphasis was placed on the Netherlands. This does not mean that such research is not available. On the contrary, the Netherlands is a perfect research case for the adoption of electric vehicles. One of the most relevant studies for this research is that of the ”buurtprognose” of the NAL, see Chapter 2, in which predictions are made for the number of electric vehicles and the demand for charging points per neighbourhood (NAL, 2022a). This model is built on outlooks for the future combined with input from municipalities and some regional characteristics. However, direct relations between neighbourhood characteristics and historical EV adoption are not included in this model.

Another relevant predictive model on EV adoption is that of van Montfort et al. (2016). They used data from charging sessions at public chargers in the city of The Hague to evaluate the needed charging infrastructure in the city. They created a definition for an electric vehicle user that could be measured in the charging data. Based on this definition, they provided new insights for the placement of public chargers. Which has a similar scope as that of this study. This definition will be used as a starting point for creating a EV adoption statistic relevant for this study.

Another paper showing the fore-running character of EV infrastructure in the Netherlands is that of Helmus and Van den Hoed (2016) on key performance indicators (KPI) of charging infrastructure. Based on stakeholders' interest and goals, they create relevant KPI's for measuring the performance of charging infrastructure. For the stimulation of electric vehicle adoption they suggest the following KPI's: growth in capacity utilisation, number of frequent users per charging station, percentage of long chargers, and charge time ratios (Helmus & Van den Hoed, 2016).

Another relevant study is the annual National Laadonderzoek ('national charging research'), which is a survey sent to members of the Vereniging Elektrisch Rijden Association, a Dutch association for electric vehicle drivers (Wolterman et al., 2022). This survey provides information on the population driving electric vehicles in the Netherlands. As discussed in Chapter 2, this population can best be described as being male, older in age, more educated, and having a higher income. These characters of Dutch electric vehicle users were found to be valid by additional research (Hoekstra & Refa, 2017).

3.2. Variables Used for Modelling

The systematic literature review for the identification of neighbourhood characteristics that drive electric vehicle adoption has been performed using the systematic literature review of Austmann (2021). This study focusses on the drivers of the electric vehicle market used in empirical studies. This makes it relevant for the identification of such relevant neighbourhood characteristics, since the literature review takes all the variables into account. However, what should be taken into consideration is that this systematic literature review by Austmann (2021) focusses on variables where this study is interested in finding characteristics, as stated in Sub-Question 2. A characteristic can be expressed using various variables, so, this dependency should be considered. Furthermore, this research is focused on adoption at public charging points, where the systematic literature review focusses on adoption as a whole.

The systematic literature review is less than ten years old and no other flaws have been detected, which gives no reason to redo this systematic literature review (Siddaway et al., 2019). To complete the study, all articles citing this review of the literature have also been taken into account. This was done for the last time on 27-12-2022, when sixteen additional papers have been reviewed.

The variables used for the modelling of electric vehicle adoption can be classified into seven categories (Austmann, 2021). These are the classes; automobile sector, incentives, socioeconomic / sociodemographic, geography, energy prices, development of electric vehicles, and psychological. This section will focus on identifying the relevant variables in these categories for this study. An overview of all the variables found in the literature can be found in Table 3.1.

3.2.1. Automobile Sector

The first category is the automobile sector, which includes variables that describe the state of the auto sector, for example, the number of sales, the presence of automobile headquarter(s), or the number of sales of (hybrid) electric vehicles. The purpose of this research is to find relevant neighbourhood characteristics. Therefore, this class of variables falls mostly out of scope, since such variables are dependent variables.

3.2.2. Incentives

The second class of modelling variables is that of incentives. This class contains variables that describe the financial and non-financial incentives taken by the governing actors to stimulate the adoption of electric vehicles. This category falls out of scope as well, since most incentives are on a national scale in the Netherlands, as found in Chapter 2. Furthermore, the policy and incentives on the municipal scale follow the same guidelines, which makes these characteristics uniform. However, the presence of environmental zones in inner cities is one of the incentives present in some of the neighbourhoods in the Netherlands and has been evaluated as beneficial for the adoption of electric vehicles (Hardman, 2019).

3.2.3. Socio-Demographic

The third category is that of socioeconomic / sociodemographic variables that describe the social, economic, and demographic state of a region. This class of variables is very relevant for this study, however, the significance of these variables can differ per neighbourhood due to cultural differences (Kumar & Alok, 2020). One of the most recurring variables in the modelling of electric vehicle adoption is income (Austmann, 2021). EV drivers today are characterised by having a relatively high income, which makes them capable of purchasing these cars. Electric cars tend to be more expensive than cars powered by an internal combustion engine (Coffman et al., 2015). This income dependency has been found by research in Europe (Nayum et al. (2016), Morton et al. (2017)), America (Axsen et al. (2016), Liu et al. (2017), Shom et al. (2022)) as well as Australia (Vidyattama et al., 2022). Although Yang and Tan (2019) found that survey participants with lower household income in Beijing were more likely to purchase electric vehicles. A possible explanation of this contradictory finding was not given, but could be due to local policy that stimulates the purchase of electric vehicles.

Another variable related to the adoption of EVs is the level of education. Higher-educated people tend to drive more in electric vehicles. This observation was made in Europe (Plötz et al. (2014), Nayum et al. (2016), Sovacool et al. (2018), Mukherjee and Ryan (2020), Haidar and Rojas (2022)) as in the United States (Liu et al., 2017). Furthermore, the rate of unemployment appears to have a negative relationship with EV adoption (Haidar & Rojas, 2022).

Younger and middle age groups tend to have a preference for electric vehicles, as found in Germany (Plötz et al., 2014) and in Nordic countries (Nayum et al. (2016), Sovacool et al. (2018)). This contradicts the findings of Wolterman et al., 2022 which found that in the Netherlands most electric car drivers are older. The same holds for Ireland, where people aged 19 to 34 years were negatively correlated with the adoption of electric vehicles (Mukherjee & Ryan, 2020). What can be concluded is that age can be a driving force for the adoption of electric vehicles.

Another variable considered is that of the size of the household (Liu et al. (2017), Morton et al. (2017)). Larger households tend to use electric vehicles more often (Nayum et al., 2016).

In the sociodemographic class, gender is another variable. It seems that electric vehicle owners are predominantly male (Plötz et al. (2014), Sovacool et al. (2018), Secinaro et al. (2022)). The same observation was made for the Netherlands (Wolterman et al., 2022) .

Another promising variable is that of the travel-to-work pattern (Liu et al. (2017), Morton et al. (2017)). This travel-to-work pattern describes the means by which people travel to work, for example, by car, by carpooling or by public transport. It appears that commuters travelling for more than one hour are positively correlated with the adoption of electric vehicles in Ireland (Mukherjee & Ryan, 2020). The travel to work pattern can be expressed in many ways as commuting distance/time or availability of public transport. Furthermore, the ratio of self-employed people is also used in the modelling and was found to be positively correlated (He et al., 2022).

Another variable correlated with the adoption of electric vehicles is that of house ownership. People who bought their house instead of rent it, tend to drive an electric car more frequently (Shom et al., 2022). This variable could be correlated with an income variable, as people who buy a house tend to have higher incomes.

The population density of an area is another factor which is used for modelling electric vehicle adoption (Haidar & Rojas, 2022). This variable could also be considered as a variable in the geography category. The population density could measure at least two common underlying characteristics. One of them, the number of inhabitants. Since studies also use population quantities for modelling (Kumar and Alok (2020), Austmann (2021)). Some specify this even more by, for example, taking the percentage of citizens with a driving licence (Haidar & Rojas, 2022). Another variable that falls under this characteristic is the size of a household, as we have seen before.

Another variable used in some studies on electric vehicle drivers is that of racial origin of citizens (Austmann, 2021). However, this study will exclude such a variable, since race has no characteristic value in this context. Race can be used as a variable to measure the underlying characteristics; however, this creates a discriminating tendency.

Car usage is an additional characteristic used in the modelling (Austmann, 2021). This characteristic is modelled using variables as vehicle distance travelled or the number of cars per household (Priessner et al., 2018). Households with multiple vehicles also appear to have more often an electric vehicle (Melliger et al., 2018).

3.2.4. Geography

The fourth category is that of variables that describe geography / infrastructure. This class is very relevant for this study, as these variables can be quite heterogeneous between neighbourhoods. One of the most important variables in this category is the availability of the charging infrastructure. The feedback loop between public charging points and adoption of electric vehicles may be a desired catalyst in the modal shift. The literature is mostly united that the placement of public charging points is one of the actions of the governing bodies that favours the adoption of electric vehicles (Austmann, 2021). Empirical research showed that early investments in public charging points contributed to the adoption of electric vehicles in the United States (Narassimhan & Johnson, 2018). Other research came to the same conclusion (Egnér & Trosvik, 2018), (Wee et al., 2020),(Sierzchula et al., 2014), (Z. Chen et al., 2017), (Kumar & Alok, 2020). Stockkamp et al. (2021) sharpens this observation by naming the placement of public charging points crucial for the modal shift. Berkeley et al. (2018) see the lack of placement of this infrastructure as a barrier to the uptake of electric vehicles.

Not all research agrees with this, however. Using time series analysis, no significant relation was found between public charging point placement and EV adoption by Kaufmann et al. (2021). However, this conclusion was based on Granger causality, so hard conclusions cannot be drawn.

Another variable used in the modelling of EV drivers is the distance to the city centres (T. D. Chen et al., 2015) or major city (Mersky et al., 2016). Furthermore, the availability of grocery stores has also been used (Li & Zhou, 2015). The same holds for the presence of shopping, catering, sports, and trade-related venues (Straka et al., 2020). It seems that public points can be good indicators as well.

In this trend, public facilities are well-used variables and more such facilities could also be of interest. For example, the distance to the nearest hospital and school. The presence of a school in a neighbourhood could indicate more households with children in the area, which is already a variable used in modelling the adoption of electric vehicles (Nayum et al., 2016). Hospitals fall into the line of sports and catering venues used by Straka et al. (2020). These variables could be an additional indicator for more urban areas. However, these two variables were not mentioned in the systematic literature review and have been added based on domain understanding.

3.2.5. Energy Prices

Variables in the energy price class are less relevant for this study, since these variables measure characteristics at higher levels of aggregation, such as fuel costs and electricity prices on national level. Therefore, in this study no variables in this class as stated by Austmann (2021) will be used in this study. However, a characteristic related to this class is that of energy consumption and production. Research has been done on the relationship between EV adoption and the presence of solar panels (Vilchez & Jochem, 2019). This characteristic can be measured on the neighbourhood scale and is therefore relevant to this study. According to Wolterman et al. (2022) 68% of the Dutch EV drivers has solar panels at home which makes it a good indicator for the overall group of electric vehicle drivers. However, the majority of the survey respondents are not dependent on public chargers.

3.2.6. Development of Electric Vehicles

The class of development of electric vehicles contains variables describing the supply of electric vehicles, which is homogenous for the whole country in the case of the Netherlands. Therefore, this class is not relevant for this study, as the variables explaining electric vehicles are aggregated at the country level and cannot explain the differences between neighbourhoods. This is the reason why in this study no variables of this class will be used.

3.2.7. Psychological

The last category contains psychological factors with variables about people's norms, attitudes, and ideas. Although this class has great interest in the literature, there is no empirical research of its connection to EV adoption, although it is positively related to intentions to buy electric vehicles (Austmann, 2021). The characteristics found in the literature are those of environmental awareness and social influence. Environmental awareness measured, for example, in political beliefs as by Zambrano-Gutiérrez et al. (2018) or by the use of existing environmental indices such as in Sierzchula et al. (2014).

Another characteristic found in the literature is that of social influence (Liao et al., 2017). When people come into contact with electric vehicles in daily life, the adoption step will become smaller. This reinforces the effect of the adoption of electric vehicles.

3.3. Conceptualisation

Based on the findings of the literature study, a conceptual model can be designed that provides the plan for the quantitative study of this research. This model will be some form of a statistical model, as suggested at the beginning of this chapter, that can be used to find relations between characteristics of the neighbourhood and the adoption of electric vehicles in public charging infrastructure. The implementation of this conceptual model will be described in Chapter 4.

This conceptual model will make use of as many of the variables found in the literature study, represented in Table 3.1. Furthermore, additional variables may be added when there is reasonable ground that these characteristics might be connected to the adoption of electric vehicles. These two groups of variables form the independent variables in the conceptual model, as can be seen at the top of Figure

Table 3.1: Variables used in the literature to model EV adoption

Variable	Class	Source(s)	Relationship
Environmental zone	Incentives	Hardman (2019)	Unknown
Income	Socio-economic	Nayum et al. (2016), Morton et al. (2017), Axsen et al. (2016), Liu et al. (2017), Shom et al. (2022), Vidyattama et al. (2022)	Mostly positive
Education	Socio-economic	Plötz et al. (2014), Nayum et al. (2016), Sovacool et al. (2018), Mukherjee and Ryan (2020), Haidar and Rojas (2022), Liu et al. (2017)	Positive
Unemployment	Socio-economic	Haidar and Rojas (2022)	Negative
Age	Socio-economic	Plötz et al. (2014), Nayum et al. (2016), Sovacool et al. (2018), Mukherjee and Ryan (2020)	Disputable
Household size	Socio-economic	Liu et al. (2017), Morton et al. (2017), Nayum et al. (2016), Melliger et al. (2018)	Positive (larger)
Gender	Socio-economic	Plötz et al. (2014), Sovacool et al. (2018), Secinaro et al. (2022)	Positive (male)
Commuting behaviour	Socio-economic	Liu et al. (2017), Morton et al. (2017), Mukherjee and Ryan (2020), Shom et al. (2022)	Positive (longer)
Self-employment	Socio-economic	He et al. (2022)	Positive
House ownership	Socio-economic	Shom et al. (2022)	Positive
Population density	Socio-economic	Haidar and Rojas (2022)	Positive
Number of residents	Socio-economic	Kumar and Alok (2020), Austmann (2021)	Positive
Car usage	Socio-economic	Priessner et al. (2018)	Positive
Charging infrastructure	Geography	Narassimhan and Johnson (2018), Egnér and Trosvik (2018), Wee et al. (2020), Sierzchula et al. (2014), Z. Chen et al. (2017), Kumar and Alok (2020), Stockkamp et al. (2021), Berkeley et al. (2018)	Mostly positive
Distance to city centre	Geography	T. D. Chen et al. (2015)	
Major city	Geography	Mersky et al. (2016)	Positive (closer by)
Distance to grocery shops	Geography	Li and Zhou (2015)	Unknown
Facilities	Geography	Straka et al. (2020)	Positive
Solar panels	Energy prices	Kaufmann et al. (2021)	Positive
Environmental awareness	Psychological	Zambrano-Gutiérrez et al. (2018), Sierzchula et al. (2014)	Unknown
Social influence	Psychological	Liao et al. (2017)	Expected positive

3.1. Furthermore, the conceptual model makes use of electric vehicle adoption statistic which describes the adoption of public charging points per neighbourhood. Furthermore, the conceptual model should also include some statistic for the occupancy of the chargers in a neighbourhood. This helps to provide insight in the effect of the independent variables on the actual adoption at the chargers and the usage of these chargers. Since these chargers are public, not only residents use these charging points. The two dependent variables combined with the independent variables will create two statistical models. Combining the results of these statistical models will help answer the research question. The conceptual model can be seen in Figure 3.1.

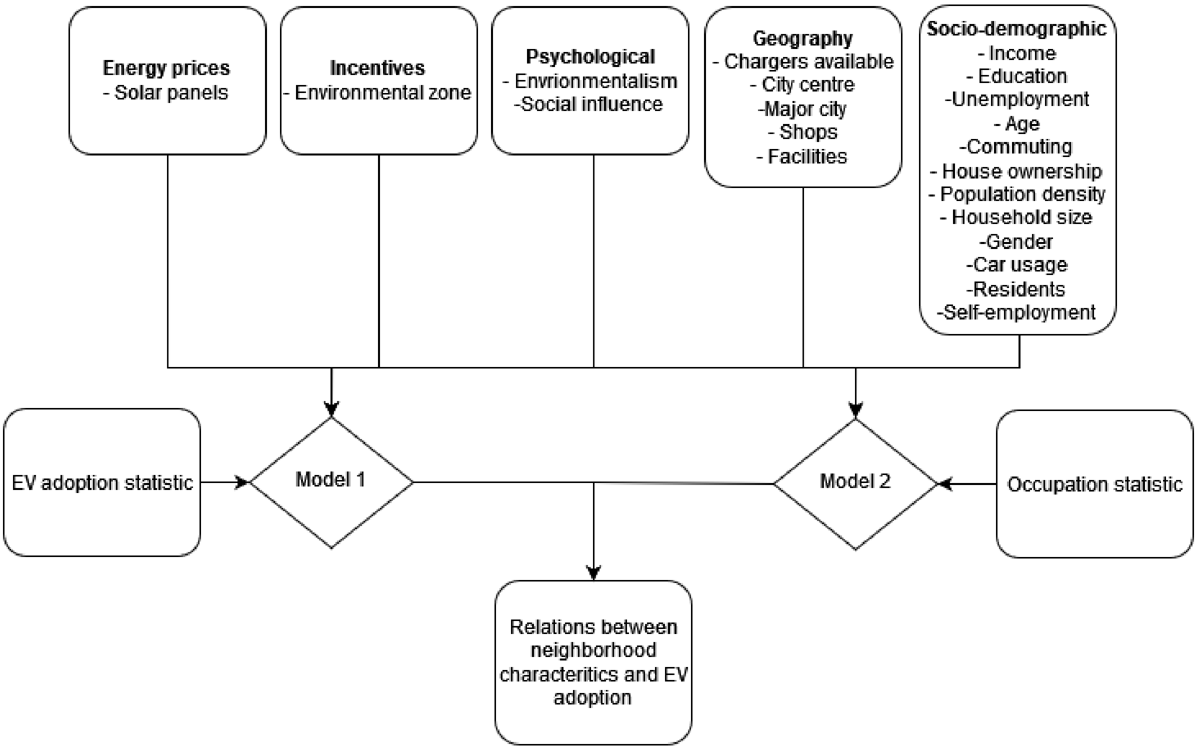

Figure 3.1: Conceptual model

4
Methods

This chapter describes the operationalisation of the conceptual model with which the previous chapter ended. Therefore, it will describe the methodology used for the statistical models. First, the overall scope and level of aggregation will be stated. Thereafter, the two KPIs for the adoption of electric vehicles will be constructed. First, that of the EV user per neighbourhood statistic. Then the occupancy rate of the public chargers per neighbourhood. Next, independent variables as stated in the literature study in Chapter 3 will be mapped to measurable data at the neighbourhood level in the Netherlands. The chapter will end with a description of the modelling techniques used and their assumptions.

4.1. Research Scope

This study focusses on the adoption of electric vehicles in the Netherlands. This adoption of electric vehicles will be measured from real-world data based on the charging transaction of public charging points. These data are supplied by IDO-LAAD (2022), which contains historical charging transactions of public chargers in the Netherlands. For this study data are available on the MRA-Elektrisch region, one of the Regional Agenda Laadinfrastructuur regions, as stated in Chapter 2. In practise, it means that data are available for the provinces of Noord-Holland, Flevoland, and Utrecht with the exception of the cities of Amsterdam and Utrecht. The MRA-Elektrisch region is visualised in Figure 4.1. It is important to note that although the uttermost effort has been made, the adoption at public charging points does not reflect the total adoption of electric vehicles in the Netherlands. As indicated by Chapter 2 most electric vehicles are charged by private chargers in the Netherlands. However, this study focusses on the adoption of electric vehicles on public chargers to improve the rollout policy for these chargers.

4.2. Aggregation Level

Relations between EV adoption and neighbourhood characteristics must be expressed at some level of aggregation. Since the intention of the study is to measure effects on neighbourhood level, this aggregation level should meet at least this level.

The Netherlands is divided into municipalities, each of which is divided into districts ('wijken'). These districts are divided into neighbourhoods ('buurten'), which is the most detailed regional level (CBS, 2022a). The distinction between districts and neighbourhoods is made by the municipalities and the 'Centraal Bureau voor de Statistiek' (CBS) is the organisation who keeps track of these statistics. However, smaller levels of aggregation are possible. Neighbourhoods consist of postal codes which can be distinguished into the postal-4, postal-5, and postal-6 levels. The number represents how many characters of the postal code are used for the aggregation level. An even more detailed level can be created using geographical grids, such as the 500x500 metre grid or the even more detailed 100x100 metre grid.

At most aggregation levels, CBS and other parties keep track of statistics and characteristics. However, not every level of aggregation has the same standard of degree of statistics. In practise, this means that the smaller the aggregation level, the smaller the degree. Therefore, this study will take the neighbourhoods ('buurt') as the level of aggregation since this is the highest suitable level of aggregation with the most data available. This research will use the neighbourhood division ('buurtindeling') of

Figure 4.1: MRA-Elektrisch region

2020 version 3 (CBS, 2022c). This version has been chosen since this is the most recent final version. It should be noted that these data are more than two years old.

Of these neighbourhoods, not all were found to be applicable for modelling. Neighbourhoods were excluded according to the following criteria:

- No registered citizens in the neighbourhood.
- No houses in the neighborhood.
- No registered households in a neighborhood.
- No public charging point is present which meets all the following conditions:
  - It is not a fast charger.
  - The charger was operational after 01-01-2022.
  - The charger has been operating for at least four weeks.
  - The charger has an operational status.
  - The charger should have at least one socket.
  - Had reliable location data, see Chapter 6 the Discussion.

These exclusion criteria are used to analyse the neighbourhoods where actual adoption of electric vehicles could be measured in charging transactions. Therefore, it was necessary to have at least one operable public charging point available in each neighbourhood that was analysed. Otherwise, no charging transactions are available. Furthermore, this study focusses on the relations of EV adoption with neighbourhood characteristics and therefore it is required that actual people live in the neighbourhood to create these characteristics.

Using the exclusion criteria, a total of 9,111 records of 11,074 records in *DIM_CHARGEPPOINT* were found to meet the requirements using IDO-LAAD (2022). Combining this with the 1326 neighbourhoods that were found to meet the requirements, a total of 5405 valid distinct chargers, *ChargePoint_ID*, were found. Based on the requirements, 42 chargers, *ChargePoint_ID*, did not make the selection, as the neighbourhoods were not valid. Later in this section, another cut will be made in the neighbourhoods based on the availability of data on neighbourhood characteristics, see Section 4.6.6.

The used number of charging points is significantly lower than the original number of 9111 chargers in *DIM_CHARGEPPOINT*. This is due to the fact that charge points with the same *ChargePoint_ID* can have multiple records in *DIM_CHARGEPPOINT*. This happened because many of the chargers changed from provider, *Provider*, during the year 2022 which caused multiple records. An overview of the number of public charging points per neighbourhood can be found in Figure 4.2. The code to create this spatial region can be found in Appendix A.8.1.

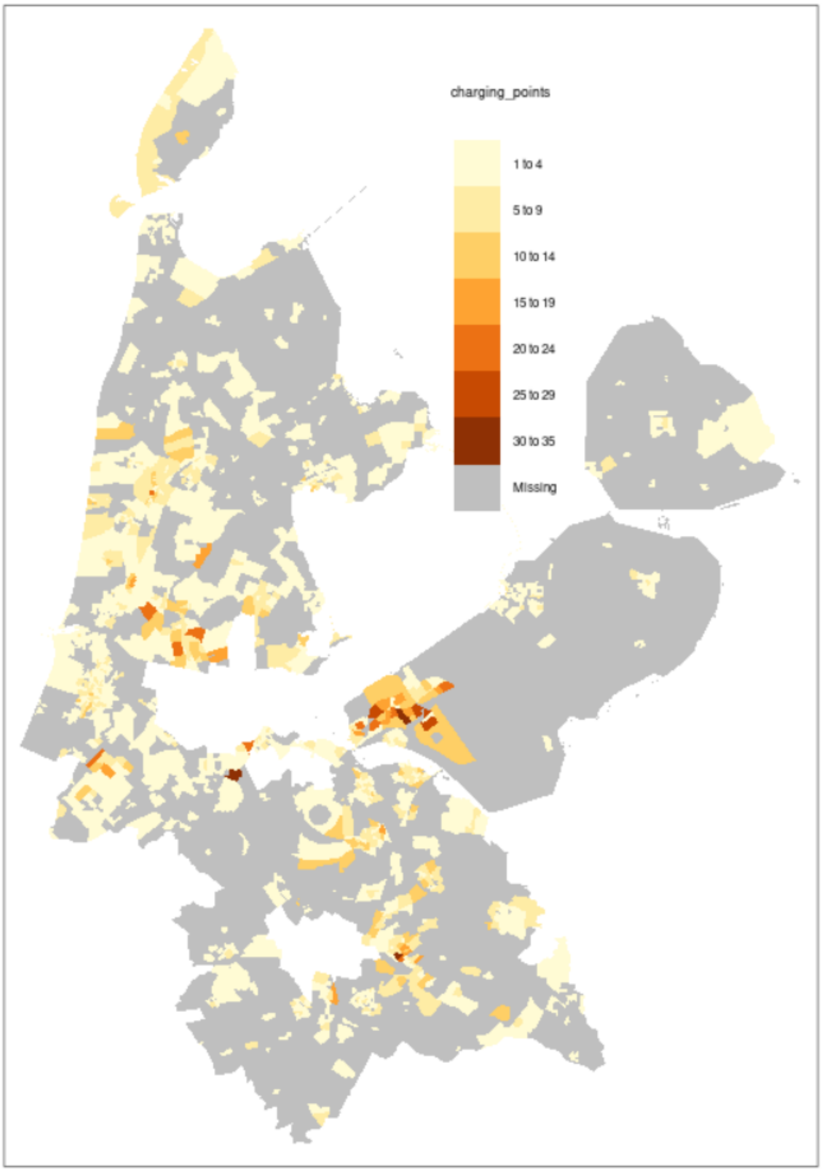

Figure 4.2: Public charging points

4.3. EV Users Statistic

4.3.1. Definition

A model is as good as the data you put in it. Therefore, it is important that the variables are of the highest quality. In this case, this means a good statistic for the adoption of electric vehicles at public charging points on neighbourhood level. Since no such statistic is available, this study creates one using the CHIEF datamart (IDO-LAAD, 2022). This datamart consists of charging transactions at public charging points in the Netherlands. The datamart is owned by the IDO-LAAD project that analyses charging behaviour and is a collaboration between the Amsterdam University of Applied Sciences, the University of Amsterdam, and ELaadNL (IDO-LAAD, 2022). The data is made available for the datamart by consortium partners and is not publicly accessible. However, for this research, access was granted to part of this datamart by the data owners.

As mentioned in Chapter 3 there are special KPI's for the measurement for stimulation of the electric modal shift (Helmus & Van den Hoed, 2016). These performance indicators consist of growth in capacity utilisation, the number of frequent users per charging station, the percentage of long chargers, and the charge/connection time ratio. Furthermore, in the Netherlands, earlier research created an adoption measurement based on public charging data in which EV users charge at public charging points close to their home, described by the following definition (van Montfort et al., 2016):

- The driver charges more than five times a month at a charging point.
- Or at multiple charging points in a 200-meter scale radius, based on a maximal walking distance of 250 meter.
- These charging sessions start between 16:00 and 04:00.

This definition of electric vehicle adoption based on public charging transactions makes use of KPI 'frequent users per charging station' as in Helmus and Van den Hoed (2016). This measurement is also applicable for this study since it recognises EV users and maps them to the location where they are measured for neighbourhood characteristics as well; the neighbourhood in which they live.

This study extends the definition by reducing the time-varying component of these data. As one might think, not every month the same number of EV adopters will be found using this definition, due to time-varying factors such as holidays and unequal days in a month. Therefore, this study extends the definition with the following criteria:

- At least four of the 12 months in 2022 the driver has to meet the criteria above.

The addition of this criterion ensures that drivers who spend part of the year at a different location than their home neighbourhood will be excluded. The four-month criterion has been chosen, since this is the period after which one should report a move to another address in the largest municipality of the Netherlands (Gemeente Amsterdam, 2022). The effect of this criterion is that electric vehicle adoptions in the last three months of 2022 will be excluded. But quality is chosen above quantity in this case.

4.3.2. Creation Statistic

The EV users' statistic is created using the CHIEF datamart. Through this datamart, the CHIEF.DM database is accessible containing the *FACT_CHARGESESSION* table with the data of charging sessions provided directly by the consortium partners. In this table, every instance registers a charging transaction made by a public charging point. Since these data are supplied directly by the providers, the data can contain flaws or missing values. Foreign keys direct to other mapping tables containing data on the RFID tag, location, date, charging point, and car type. These mapping tables are assumed to be correct. The relation scheme with all the available tables, foreign keys, and columns is shown in Figure 4.3. Furthermore, the lookup table *LOOKUP_REMARK* describes whether there are comments with the transaction. In Appendix A.3 a more detailed description of the magnitude of the *FACT_CHARGESESSION* records can be found to put the findings into perspective.

Using the datamart, the definition of Section 4.3, and the aggregation level, as explained at the beginning of this chapter, the adoption statistic of the users of electric vehicles was created. This was done by coding in the R language (R Core Team, 2021), which can be found in Appendix A.2. The following approach has been followed:

1. Transactions with NA values in the columns named hereafter were filtered out.
2. A selection was made on the transactions:
    - In the *UseType* column only 'regulier' is selected.
    - In the *IsValid* column only 1 values are selected, so only select transactions without remarks.
    - In the *kWh* only values greater than 0 are selected.
    - In the *RFID_skey* only values greater than 0 are selected.
3. The remaining transactions are grouped by month, year, and *RFID_skey*. The first two variables are obtained from *StartConnectionDateTime*. The *RFID_skey* is removed from the data set if it exceeds 50 charging sessions per month. This excludes non-personal usage of RFID charging tokens, which is not in the scope of the study.
4. Using the *StartDate_skey* only transactions between 16:00 and 04:00 are selected as in van Montfort et al. (2016).
5. The *DIM_LOCATION* is added to the table and transactions with missing *PostalCode* or *Location_skey* are dropped.
6. For the remaining *RFID_skey's* in the dataset all the unique *RFID_skey* values are placed in a *RFID_list* and are checked one by one if they have charged more than 5 times in a 200 metre radius that month. This is checked using the Algorithm 1.
7. The six steps above are repeated for every month in 2022. For each month, the tuples of (*RFID_skey*, *PostalCode*) are stored. When this tuple appears at least four times throughout the year, *RFID_skey* is counted as EV users for its *PostalCode*.
8. The sum of the EV users per *PostalCode* is brought back to the neighbourhood level based on a mapping table supplied by the CBS (CBS, 2020).

Algorithm 1 Finding sessions within 200 metre radius

$EVusers \leftarrow$ list
for $RFID$ in unique $RFID\_list$ **do**
  $Locations \leftarrow$ unique(*Location_skey[RFID]*)
  $Sessions \leftarrow$ unique(*ChargeSession_skey[RFID]*)
  if $|Sessions| > 5$ **then**
    $Locations\_grouped \leftarrow$ sort(groupby(*Location_skey*))
    $Most\_used\_charger \leftarrow Locations\_grouped[0]$
    if $|Most\_used\_chargersessions| > 5$ **then**
      $EVusers^\frown$Postcode[most_used_charger]
    else
      $dist\_matrix \leftarrow Locations$ within 200 meter radius of each charger
      if $max(dist\_matrix) > 5$ **then**
        $EVuser^\frown Postcode[max(dist\_matrix)]$
      end if
    end if
  end if
end for

The reason the mapping is made on postcode level is that many of the charging points changed provider during the year 2022, which has as consequence that the *ChargrePoint_skey* changed of these chargers. By mapping on postcode level, this does not result in data loss in the transfer period. However, during the mapping at postcode level, some incorrect values appeared that have been excluded. See Chapter 6 the Discussion.

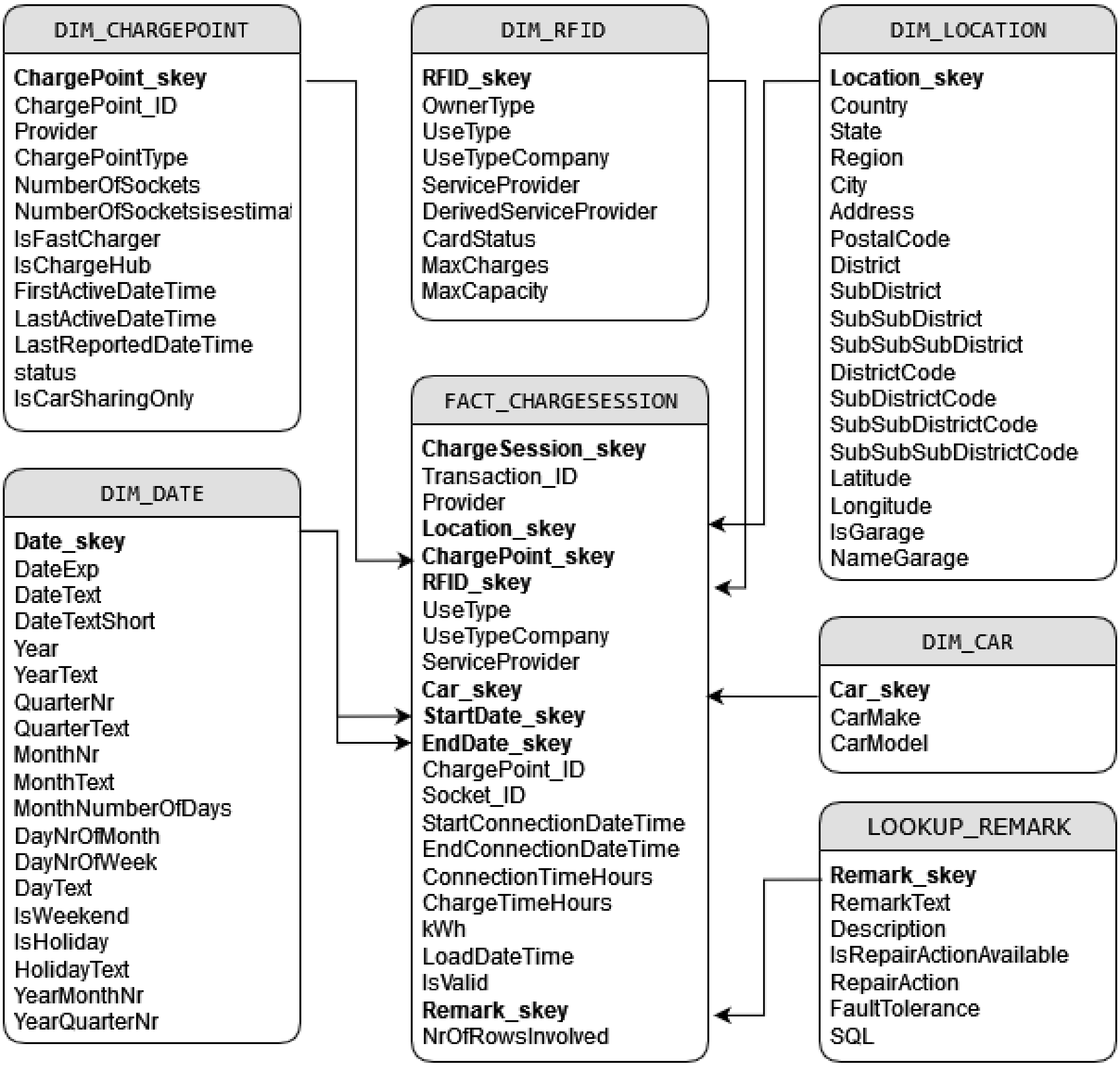

Figure 4.3: Relational scheme CHIEF datamart

4.4. Occupancy Rate

4.4.1. Definition

In addition to the number of EV users per neighbourhood statistic, this study also focusses on the occupancy rate of public chargers, as stated in the conceptual model in Section 3.3. This section will focus on defining this occupancy rate and creating it from the CHIEF datamart (IDO-LAAD, 2022).

The occupancy rate of a charging point is one of the KPI's that can be used to measure electric mobility (Helmus & Van den Hoed, 2016). However, the occupancy rate can be defined using two definitions. One of them is the part of time when a car is connected to a charging point. The other definition is the part of time that a charging point is used for charging. The latter definition is a subset of the first. In the first definition, cars that stand at a charging point but do not charge are included. Although this behaviour is not desired, they still use the charging point service. Therefore, this study defines the occupancy rate as the part of the time that a charging point is connected to a vehicle.

For this KPI the same aggregation level applies as for the previous dependent variable. Therefore, the occupancy rate of all charging points will be aggregated on neighbourhood level. This will be done by taking the average occupancy rate of all chargers in a neighbourhood where every charger gets the same weight. Chargers are excluded on the basis of conditions as in Section 4.2. The occupancy rate will be estimated on the basis of the year 2022. This is done by computing the monthly occupancy rates per neighbourhood and taking the average for the whole year. This method has been chosen to compensate for charging devices that have been installed throughout the year.

4.4.2. Creation Rate

Using the datamart, the definition of Section 4.4 and the aggregation level as explained above, occupancy rates were determined. This was done by coding in the R language (R Core Team, 2021), which can be found in Appendix A.8.1. The following approach has been followed:

1. Transactions with NA values in the columns named hereafter are filtered out.

2. A selection is made on the transactions:

   - In the *IsValid* column only '1' values are selected, so only the transactions are selected without remarks.

   - In the *kWh* only positive values greater than 0 are selected.

   - In the *RFID_skey* only positive values greater than 0 are selected.

3. Transactions are grouped by *ChargePoint_skey* and the sum is taken from *ConnectionTimeHours*.

4. The sum of *ConnectionTimeHours* is divided by *NumberOfSockets* per *ChargePoint_skey*.

5. Using the *PostalCode* and *Location_skey* columns in *DIM_LOCATION* table the postal codes are joined to the data.

6. The data is now grouped by *PostalCode* and again the average is taken of the *ConnectionTimeHours*.

7. The sum of *ConnectionTimeHours* per *PostalCode* is brought back to neighbourhood level based on a mapping table provided by the CBS (CBS, 2020).

8. The sum of *ConnectionTimeHours* is transformed into an occupancy rate by dividing the sum of *ConnectionTimeHours* by the total time available.

9. The steps above are repeated for every month of 2022 and the average occupancy rate is taken for each neighbourhood over these months.

4.5. Dependent Variables

The results of the above construction methods can be found in Figure 4.4 for the EV users per neighbourhood and in Figure 4.5 for the average occupancy rate per neighbourhood. In these figures, all neighbourhoods that were found to be suitable for this study have a colour other than grey. The more red the neighbourhoods, the higher the statistic. Note that the occupancy rate statistic makes use of more charging transactions than the EV users statistic, since the restriction of *UseType: Regulier* has been let go. However, the difference in the number of charging transactions is limited, as can be seen in Appendix A.3.

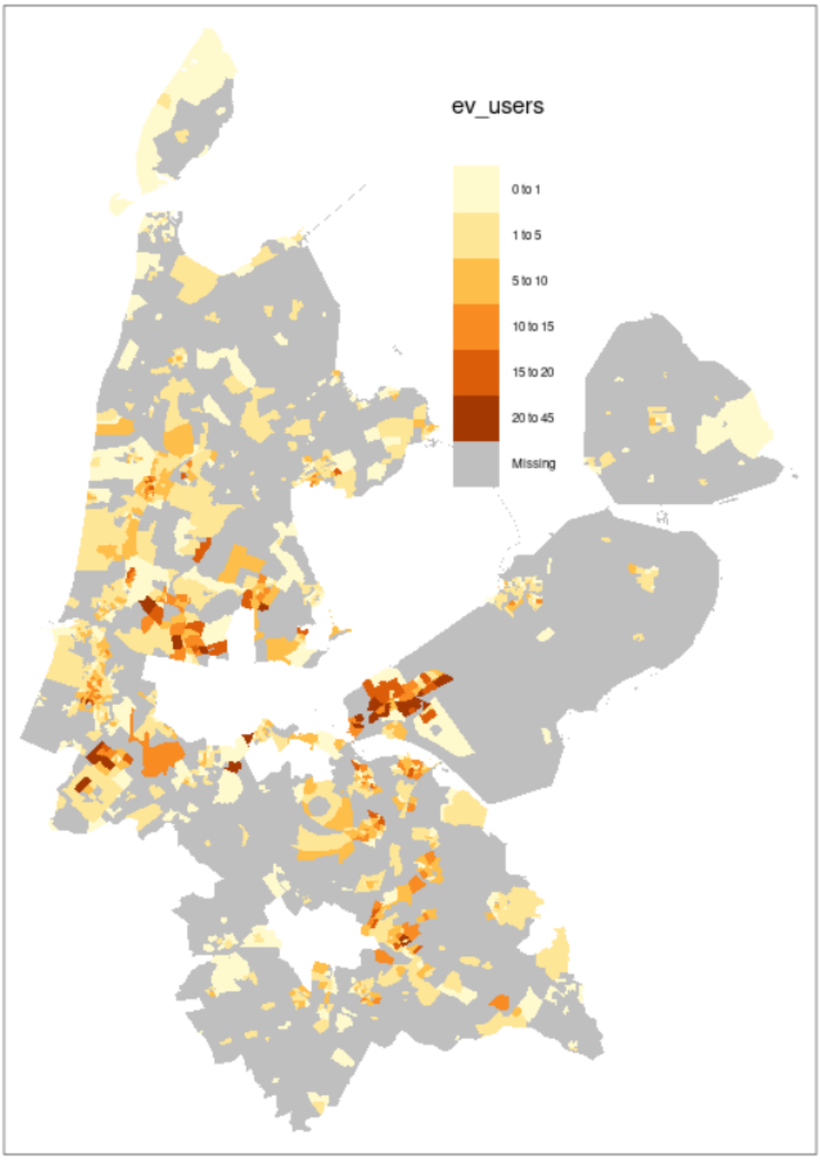

Figure 4.4: EV users

As can be seen in Figure 4.4 most of the neighbourhoods have between one and five EV users measured in the public charging data. From now on, this will be described as EV users. However, it is important to note that this does not reflect the total population of EV users in a neighbourhood. It is not surprising that the one-to-five group accounts for the major part of the neighbourhoods, with on average 4.19 EV users per neighbourhood. However, there are neighbourhoods that have far more EV users. Most of these neighbourhoods are located in Almere, around Amsterdam, or to the east of Utrecht. It seems that EV adoption thrives better closer to larger cities.

The distribution of EV users can be found in Figure 4.5 where it can be seen that the maximum number of EV users per neighbourhood is 42 in the MRA-Elektrisch region. The standard deviation of the distribution was found to be 5.34. The histogram shows an exponentially decreasing behaviour. The high peak at zero indicates that most of the neighbourhoods do not have electric vehicle users.

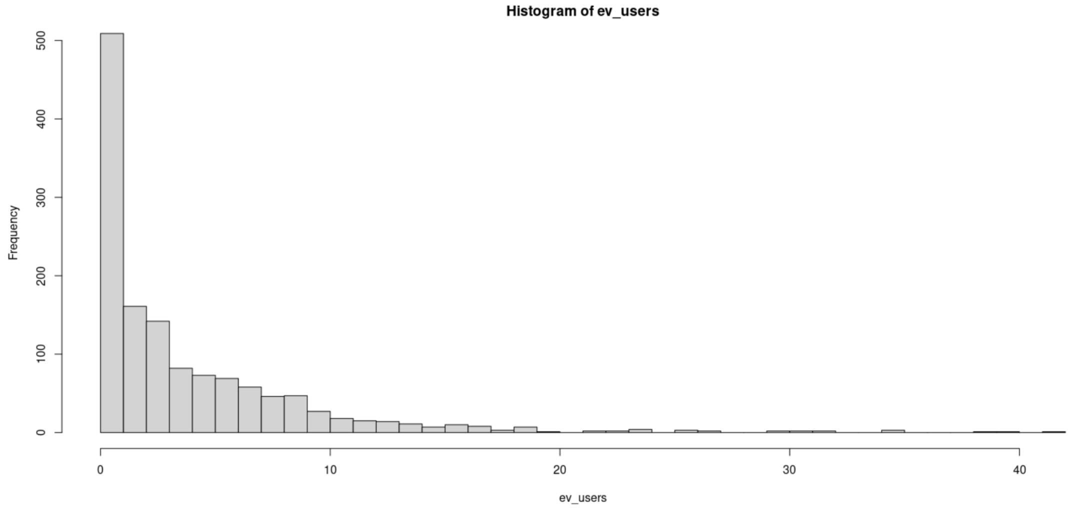

Figure 4.5: EV user histogram

The Figures 4.4 and 4.6 show similar behaviour in the two adoption KPI's over the spatial region. Neighbourhoods with higher occupancy rates can be found in the region around the cities of Amsterdam and Utrecht. This is in line with the behaviour of EV users. However, this relation does not hold for the neighbourhoods in the city of Almere. Where the city of Almere has one of the highest number of EV users, the occupancy rate of public chargers does not follow this trend. More public chargers appear to be available in this region to compensate. This is in line with Figure 4.2. Furthermore, the less urban regions scored less on the average charging occupancy rate.

On average, all charging stations in the neighbourhood had an occupancy rate of 0.270, resulting in approximately 6.5 hours of connections per day. The distribution of the occupancy rate can be found in Figure 4.7. This figure shows a distribution with characteristics of a skewed normal distribution. The standard deviation was found to be 0.126.

4.6. Selection of Independent Variables

This section will explain which and how the variables indicated by the literature study in Chapter 3 are translated into suitable data for modelling. An overview of the variables considered, their origin, and processing method can be found in Table 4.1. The distribution of the variables can be found in Appendix A.5. The code used for the creation of independent variables can be found in Appendix A.8.1.

4.6.1. Energy Prices

Solar Panels

In the category of energy prices, the presence of solar panels was found to be the only variable that is heterogeneous at the neighbourhood level. In the Netherlands, the CBS keeps track of the number of solar panel installations, the power of these installations in kW, and the energy generated by these installations on neighbourhood level (CBS, 2022d). The power of the installations in kW per neighbourhood was chosen as modelling variable. The code for the cleaning and processing of these data can be found in appendix A.5. The variable is named *Solar_power* in this study and its distribution can be found in Appendix A.24.

4.6.2. Incentives

Environmental Zones

In the Netherlands, 15 municipalities use environmental zones. Four of these municipalities (Amsterdam, Arhnem, Den Haag, and Utrecht) have restrictions on cars, while the others focus on other categories. The locations of these environmental zones are available at the postcode level (Rijksoverheid, 2022a). However, none of these cities fall into the observed region, which makes adding the variable redundant. Therefore, in this study, no variable in the incentive category has been included.

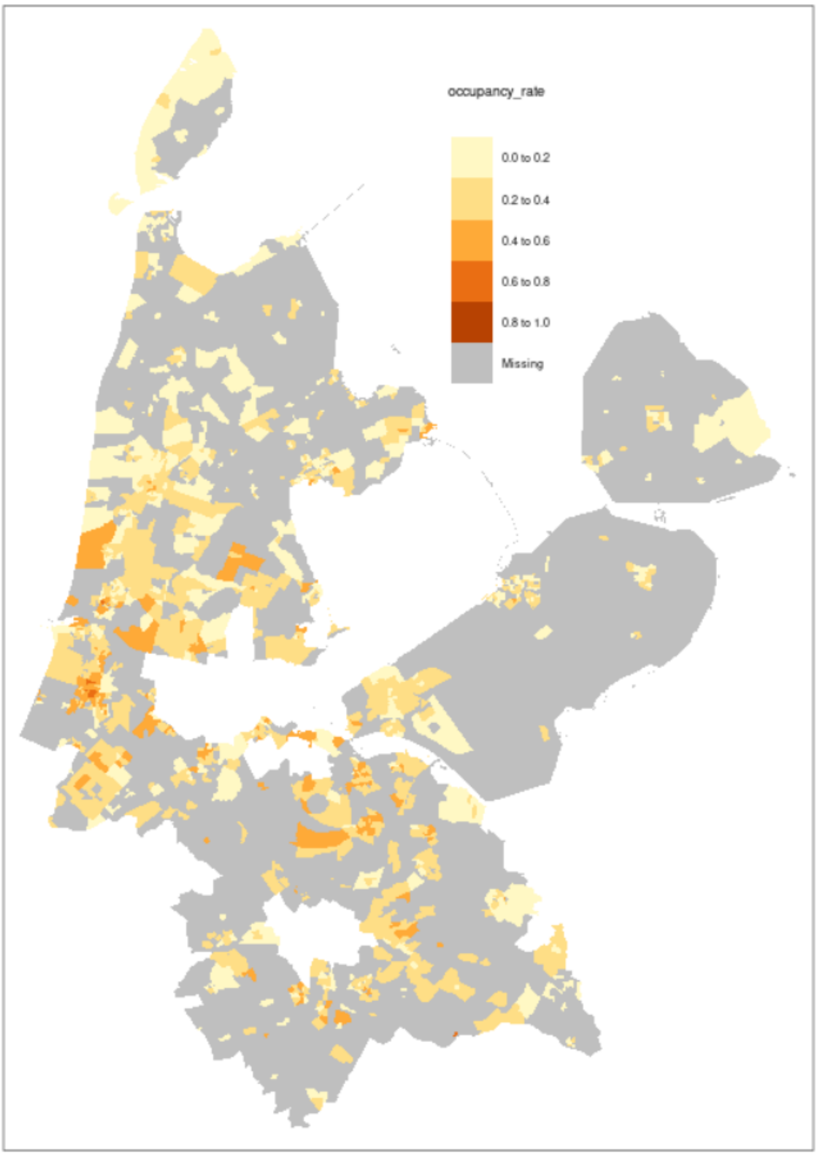

Figure 4.6: Occupancy rate

4.6.3. Psychological

Environmentalism

Environmentalism is a character difficult to express, as there is no clear scale to measure this at the neighbourhood level. In this study, environmentalism will be modelled as the percentage of votes on green parties in the latest parliamentary elections in the Netherlands. This was the most recent election where the voting options were homogeneous for the whole of the Netherlands. The turnout of this election was 78.7% (Parlement.com, 2022) which makes it relatively reliable. The number of votes per polling station was used to calculate this percentage at the neighbourhood level (Open State Foundation, 2021). Using the Python script in Appendix A.10 the JSON files were mapped to the neighbourhoods (Van Rossum & Drake, 2009). The following political parties were considered green choices during this election: GroenLinks, PvdD, SP, PvdA. This was based on the voting behaviour of these parties in the three years before the election (Kiesklimaat, 2021). The variable is named *Perc_green_votes* and its distribution can be found in Appendix A.15.

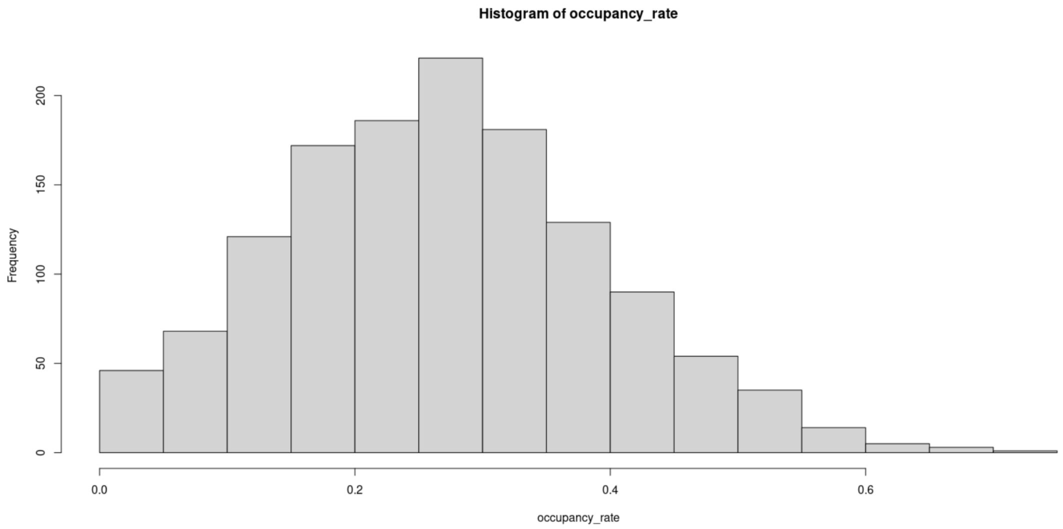

Figure 4.7: Occupancy rate histogram

Social Influence
The choice has been made to exclude social influence from the modelling dataset although it was one of the variables that occurred in the literature study in Section 3. The social influence in this case can only be measured by some sort of spatial lag in the number of EV users. However, EV adoption is a modal shift that makes the number of EV users will increase the coming years and has increased in the last years. Including a spatial lag would of course be found significant due to the increasing effect of the modal shift. However, it does not explain the characteristic as targeted here. Therefore, this variable was not taken into account.

4.6.4. Geography

Chargers Available
According to the findings of the literature study in Section 3 the available chargers affect the adoption of electric vehicles in a positive way. However, the choice has been made not to include the number of chargers as independent variable for the explanation of both the EV adoption KPI's. This choice was made based on the insights of Chapter 2 where it was found that most public chargers in the Netherlands are placed on demand. Therefore, including the number of chargers for the explanation of EV adoption would lead to incorrect causality.

Distance Centre
Many close proximity statistics exist at the neighbourhood level in the Netherlands. However, the distance to the city centres is not available through any data outlet. An approximation used in this city is the distance to a train station since these are mostly located in central places. Furthermore, this variable will also be used for commuting characteristics. The distance to the nearest train station is retrieved from the CBS CBS (2022c). The variable is named *Dist_train_station* and the distribution can be found in Appendix A.9. As for all the variables coming from this data set, the data has been filtered on -99999999 values which are either unknown, not reliable, or secret. For the city centre approximation, a new variable was created which is either one or zero, depending on whether the distance to the nearest train station is within one kilometre. This variable is called *Has_city_centre*.

Shops
As stated in the previous section, there are many proximity statistics at the neighbourhood level. The CBS provides data on the distance for the closest provider of many services. The approach of these characteristics is a measure of the proximity of services. Therefore, it has been chosen to create a statistic of the average road distance to the services mentioned below. The same method has been used to handle incorrect values as in *Distance Center* has been used. The variable is called *Dist_shops* and its distribution can be found in Appendix A.8.

- Department store.
- Supermarket.
- Pharmacy.
- Restaurant.

Public Points
The CBS provides distances from the neighbourhoods to various public points. With a selection of these points, the presence of a public point can be estimated. Using the minimum distance to the public points stated below and the binning of the distances, a variable was created to determine whether a public point was present in the neighbourhood. This is done by dividing the continuous distance variable into bins and selecting the first bin as the estimator for the public point present. This means that if the distance to one of the facilities below is less than one kilometre, it is modelled that a public place is present. The variable is called *Has_public_point*. Only the binary variant of this characteristic is included in the vision of strategically placed public chargers near public points. This variable was not found in the literature study of Chapter 3 but found relevant based on the analysis of the policy of Chapter 2.

- Amusement parks.
- Zoo.
- Indoor playground.
- Museum.
- Swimming Pool.

Public Facilities
Other public facilities that were taken into account for the modelling are the distance from primary schools and hospitals. The CBS provides statistics for these variables (CBS, 2022c). The variable is called *Dist_school* for schools and its distribution can be found in Appendix A.7. The distance to hospitals is named *Dist_hospital* with its distribution in Appendix A.6. The same approach has been followed as in previous sections to create a variable to determine whether such a facility is present in a neighbourhood. When the average distance from a neighbourhood is less than one kilometre to a facility, the facility is modelled as present. The variables are called *Has_hospital* and *Has_school* respectively. For these variables also holds that they were not found in the literature study of Chapter 3. However, they were included to investigate whether this was correct.

Parking Places
Parking places are connected to public charging stations, as these charging stations are located at public parking lots. Data on parking places are available in the Netherlands, however, this consists only of private (paid) parking places. Centralised data are not available at public charging places. There are incentives to start recording the so-called blue zone parking zones in the near future. For now, this kind of parameter could not be taken along in the modelling.

4.6.5. Socio-Demographic

Income
Income is one of the most used variables in modelling the adoption of electric vehicles (Austmann, 2021). In the Netherlands, the CBS keeps track of income statistics on neighbourhood levels. After an initial analysis, it was found that the statistics stating the average or median income in the neighbourhoods contained a high percentage of NA values (higher than 50%). Therefore, the choice has been made to take statistics about the percentage of incomes in the 40% lowest incomes and in the 20% highest. The percentage missing data is respectively lower, however, is still moderate since the CBS only reports this statistic for neighbourhoods of 100 people or more. The variables are called *Perc_high_income* and *Perc_low_income* and their distributions can be found in Appendix A.17 and A.20 respectively.

Education
Education is modelled using the number of high, medium, and low educated people statistics of the CBS (CBS, 2022c). For modelling, the statistics are normalised by the number of citizens per neighbourhood. Since the three classes complement each other, only the percentage of high- and low-educated people are taken. The variables are called *Perc_high_educated* and *Perc_low_educated*. The underlying distribution can be found in Figures A.16, and A.19 respectively in the Appendix.

Age
Age is modelled using CBS statistics on the number of citizens in certain age classes and are transformed into percentages. The classes used are the following:

- 15-24 years: *Perc_15_24_yr*, see Appendix A.11.
- 25-44 years: *Perc_25_44_yr*, see Appendix A.12.
- 45-64 years: *Perc_45_64_yr*, see Appendix A.13.
- 65 and older: *Perc_65_EO_yr*, see Appendix A.14.

The group 0-14 years has been excluded beforehand for modelling purposes. This, since this category is represented by variables in Section *Household size* and adding this category would mean the inclusion of the complete dimension.

Commuting
Data on travel distance are available in the Netherlands, however, at the municipal level. This means that for this study no data on commuting means is used. However, indirect data will be used. As stated above, the distance to the train station is used in the model: *Dist_train_station*. Another variable that will be taken into account in this study is the percentage of self-employed people in a neighbourhood. This variable is called *Perc_self_employed* and its distribution can be found in Appendix A.22.

House Ownership
Another house characteristics is who has the ownership of the house. Data are available on the neighbourhood level of the percentage of purchased and rented houses (CBS, 2022c). The variable included is the percentage of rental houses in a neighbourhood, *Perc_rental_house*, and can be found in Appendix A.21. The percentage of bought houses is not included since it is the complement of *Perc_rental_house*.

Density
The density, as described in Chapter 3, can measure multiple underlying characteristics. Therefore, these data were also added to the variable *Population_density*. The underlying distribution can be found in Appendix A.23. Furthermore, the number of citizens in a neighbourhood is also added in the variable *Num_residents* with the underlying distribution in Appendix A.10.

Household Size
The household size is another socio-demographic characteristic. Data on this are found via the CBS with statistics of the average household size per neighbourhood and the percentage of households with children. Both are used for modelling using variables *Avg_household_size* and *Perc_household_child*. Their distributions can be found in the Appendices A.4 and A.18.

Gender
Gender statistics are available, as well as the number of male and female residents in a neighbourhood (CBS, 2022c). However, its contribution is limited since, in practise, the population will be divided 50/50 on the neighbourhood level. Therefore, these variables are not taken into account for modelling.

Car Usage
The CBS provides data on the number of personal cars present in a neighbourhood (CBS, 2022c). In practise, the more cars there are in a neighbourhood, the more electric vehicles will be present. Therefore, the density of the car is used instead as the modelling variable. This statistic is created by dividing the number of cars by the number of residents. The variable is called *Car_density*. The distribution of this variable can be found in Appendix A.5.

4.6.6. Data Cleaning

The independent variables as described were processed to make them suitable for modelling. Since the data came from official data outlets, it was assumed that incorrect values could not occur. However, not available (NA) values were present in most of the data. These values are problematic for the modelling and therefore had to be dealt with. In the process of cleaning neighbourhoods with more than 25% of the independent variables as NA values, they were deleted from the data set. This resulted in a reduction of 34 neighbourhoods. The final selection of neighbourhoods can be found in Figure 4.8. The deleted neighbourhoods due to the high rate of NA values can be seen in Appendix A.3. The remaining percentage NA values per independent variable can be found in Table 4.2.

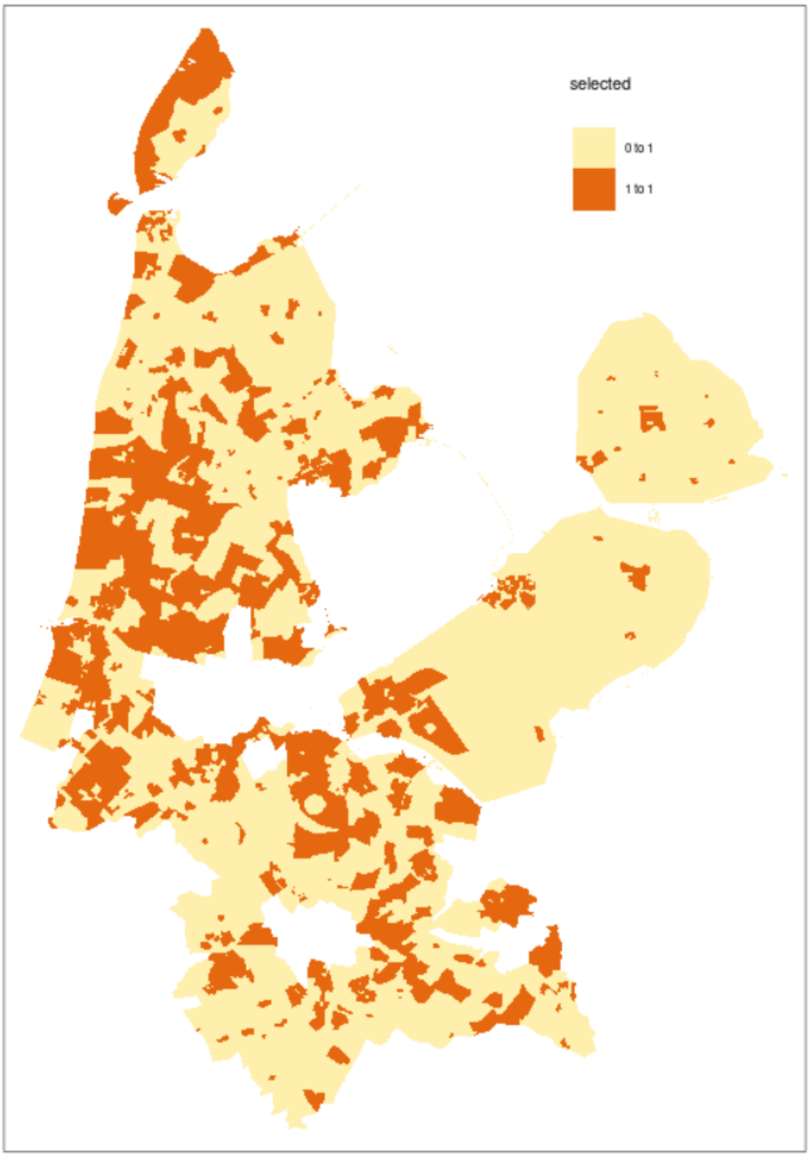

Figure 4.8: Final selection of neighbourhoods

Table 4.1: Independent variables

Variable name	Description	Origin	Year	Preprocessing
Avg_household_size	Average household size	CBS (2022c)	2020	Appendix A.8.1
Car_density	Density of cars per resident	CBS (2022c)	2020	Appendix A.8.1
Dist_hospital	Distance to closest hospital	CBS (2022c)	2020	Appendix A.8.1
Dist_school	Distance to closest primary school	CBS (2022c)	2020	Appendix A.8.1
Dist_shops	Average distance to shopping facilities	CBS (2022c)	2020	Appendix A.8.1
Dist_train_station	Distance to closest train station	CBS (2022c)	2020	Appendix A.8.1
Has_city_centre	Train station is available within one kilometer	CBS (2022c)	2020	Appendix A.8.1
Has_hospital	Hospital is available within one kilometer	CBS (2022c)	2020	Appendix A.8.1
Has_public_point	Presence of amusement park, zoo, indoor playground, museum or swimming pool	CBS (2022c)	2020	Appendix A.8.1
Has_school	Primary school is available within one kilometer	CBS (2022c)	2020	Appendix A.8.1
Num_residents	Number of residents	CBS (2022c)	2020	Appendix A.8.1
Perc_15_24_yr	Percentage of residents in age group 15 - 24 years	CBS (2022c)	2020	Appendix A.8.1
Perc_25_44_yr	Percentage of residents in age group 25 - 44 years	CBS (2022c)	2020	Appendix A.8.1
Perc_45_64_yr	Percentage of residents in age group 45 - 64 years	CBS (2022c)	2020	Appendix A.8.1
Perc_65_EO_yr	Percentage of residents of age 65 years and older	CBS (2022c)	2020	Appendix A.8.1
Perc_green_votes	Percentage of votes on green political parties	Open State Foundation (2021)	2021	Appendix A.10
Perc_high_educated	Percentage of high educated residents	CBS (2022c)	2020	Appendix A.8.1
Perc_high_income	Percentage of residents with high income	CBS (2022c)	2020	Appendix A.8.1
Perc_household_child	Percentage households with children	CBS (2022c)	2020	Appendix A.8.1
Perc_low_educated	Percentage of low educated residents	CBS (2022c)	2020	Appendix A.8.1
Perc_low_income	Percentage of residents with low income	CBS (2022c)	2020	Appendix A.8.1
Perc_rental_house	Percentage of rental houses	CBS (2022c)	2020	Appendix A.8.1
Perc_self_employed	Percentage of self employed people	CBS (2022c)	2020	Appendix A.8.1
Population_density	Population density	CBS (2022c)	2020	Appendix A.8.1
Solar_power	Power of the installed solar panels in kW	CBS (2022d)	2020	Appendix A.5

Table 4.2: Percentage NA values

Variable name	NA values (%)
Car_density	3,1
Dist_hospital	0,8
Dist_school	0,8
Dist_train_station	0,8
Perc_green_votes	40,2
Perc_high_educated	2,9
Perc_high_income	4,4
Perc_low_educated	2,9
Perc_low_income	4,4
Perc_rental_houses	3,1
Perc_self_employed	5,9

As can be seen in Table 4.2, eleven of the 25 independent variables had NA values. These NA values will be dealt with by imputing. For neighbourhoods with NA values in *Perc_green_votes* the mean of the connected neighbourhoods has been taken. This is because not every neighbourhood had a voting station in the 2021 elections. It is assumed that the voters went to the closest voting station and, therefore, ended up voting at the voting stations in the surrounding neighbourhoods. Therefore, taking the mean of these surrounding neighbourhoods should be a good indicator of the real percentage. However, still around 32% of the neighbourhoods had NA values after applying this technique.

To replace the remaining NA values, a K nearest neighbour algorithm was used that replaces the NA values by taking the mean of the independent variables of the ten most close neighbourhoods in statistical terms (Torgo, 2014). Therefore, when an NA value occurred in, for example, *Perc_green_votes* the ten most similar neighbourhoods were selected based on a clustering algorithm in the other variables, and the mean of these ten neighbourhoods was used to estimate *Perc_green_votes* for the neighbourhood with the missing value. This resolved all NA values in the dataset.

Furthermore, the independent variables were checked for correlations with each other. For modelling purposes, correlations between independent variables are undesired. Therefore, the correlations between the variables have been checked on the exceedance of 0.8 and -0.8. The correlation matrix can be found in Appendix A.30. Two pairs of variables appeared to have a correlation higher than 0.8 and one pair of variables had a correlation smaller than -0.8. These were the pairs *Perc_low_income* and *Perc_rental_houses* with a correlation of 0.877, the pair of *Avg_household_size* and *Perc_household_children* with a correlation of 0.940, and the pair of *Perc_high_income* and *Perc_low_income* with a correlation of -0.840. It was decided to exclude *Perc_low_income* and *Perc_household_children* from the data set so that the highest number of variables could be used.

4.6.7. Analysis Differences Observed Regions

As stated at the beginning of this chapter, only charging data are available for the MRA-Elektrisch region. However, the purpose of this study is to provide information to the whole of the Netherlands. To make the eventual conclusion relevant for the entire country, the independent variables stated above will be compared between the MRA-Elektrisch region and the rest of the country. This analysis will show whether the MRA-Elektrisch region is a good sample based on the independent variables considered in this section. As we have seen in Section 2 the MRA-Elektrisch region is further in the modal shift in both the adoption of EVs and in the EV charging infrastructure, making the results of this research even more interesting for the rest of the country.

Comparison analysis was performed using statistical tests for all continuous variables. The categorical variables used in this model are all binary, making them not relevant for this kind of analysis. For the continuous variables, the following approach was used. The Kolmogorov-Smirnov test was used first to check whether the two samples were of the same distribution. Subsequently, the samples were tested to see if they had the same mean value. This was done using either the two sample T-test or the Wilcoxon-Mann-Whitney rank sum test depending on the underlying distribution, i.e. if they came from a normal distribution. The full results of these tests and the probability distribution functions of the two samples for every independent variable can be found in Appendix A.5. The results of this analysis for the independent variables in the modelling approach described hereafter can be found in Table 5.4

in Section 5.3. The code used for this analysis can be found in Appendix A.11

4.7. Modelling Approach

Chapter 3 described the categories of mathematical modelling techniques that have been used to understand the adoption of electric vehicles. This study focusses on finding the relationships between the adoption of electric vehicles around public charging points and the characteristics of the neighbourhood. These relations can be found using various techniques as long as these techniques show the statistical significance of these relations and an indication of their importance. The most suitable category of mathematical techniques in Chapter 3 would be that of statistical modelling.

In the statistical modelling subclass, a frequently used technique is regression analysis (Austmann, 2021). This is the statistical method to estimate the relations between a dependent variable and one or more independent variables. As is the goal of this study. In the field of regression analysis, there are many variants. The simplest form of regression is linear regression, where one dependent variable estimates the independent variable by a linear relation. When a combination of multiple independent variables is used, one speaks of multiple linear regression. In mathematical terms, one speaks of the estimation of a dependent variable y by the linear combination of an intercept β_0, dependent variables x_p times their coefficient β_p for all dependent variables p in $1, \ldots, n$ and an error term ε as seen in (4.1)

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i \tag{4.1}$$

As with most statistical techniques, (multiple) linear regression comes with assumptions that must be met for the model to be applicable (Poole & O'Farrell, 1971). The dependent variable is assumed to have a (continuous) linear relation to the independent variables. For the occupancy rate statistic, this assumption can be held due to its continuous distribution, as seen in Figure 4.7. However, for the EV user statistic this is not the case since this the dependent variable is a positive count variable. A better approach for the statistic of EV users would be Poisson regression, which is suitable for count data (Consul & Famoye, 1992). This holds in the case of electric vehicle users, as can be seen in Figure 4.5. However, the EV users statistic contains count data with many zero values of neighbourhoods where no electric vehicle users occurred. The distribution was tested on the presence of excess zeros and the score test, as in Appendix A.25, could not reject that no excess zeros are present. Therefore, the zero-inflated Poisson regression would fit this task better, as there could be excess zeros (Lambert, 1992). Using this type of regression, two submodels are estimated. One model focusses on excess zeros and the other model focusses on the count variable.

Another adjustment to the modelling technique for the EV users statistic must be made. Poisson regression assumes that the mean and variance of the distribution of the independent variable are of the same size (Consul & Famoye, 1992). As we have seen earlier in this section, this is not the case. An overdispersion test was used to test whether the variance is significantly larger than the mean. Which was the case, see Appendix A.26. Therefore, the modelling technique used is that of zero-inflated negative binomial regression (Ridout et al., 2001).

The zero-inflated negative binomial regression can be mathematically explained by the following. This regression model determines the expected number of EV users, $E\,(n_{\text{EV users}} = k)$, based on two processes. A process that simulates the probability of not having adoption of electric vehicles in a neighbourhood times its outcome: $P(\text{No EV adoption}) * 0$. And the other process is the probability of having EV adoption in a neighbourhood multiplied by the expected result: $P(\text{EV adoption}) * E(y = k \mid \text{EV adoption})$. The expected result is the sum of these two processes, as can be seen in (4.2). Since the first part will be zero in all cases, the expected result will always be equal to the count submodel.

$$E\,(n_{\text{EV users}} = k) = P(\text{No EV adoption}) * 0 + P(\text{EV adoption}) * E(y = k \mid \text{EV adoption}) \tag{4.2}$$

Based on the negative binomial probability density function (pdf) as in (4.3), the likelihood function can be formulated as in (4.4) for the distribution, as shown by (UCLA: Statistical Consulting Group, 2021).

$$pdf(y; p, r) = \frac{(y_i + r - 1)!}{y_i!(r-1)!} p_i^r \,(1 - p_i)^{y_i} \tag{4.3}$$

Where pdf is the probability density function of the negative binomial model with p the probability of r successes.

$$L(\mu; y, \alpha) = \prod_{i=1}^{n} \exp\left(y_i \ln\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right) - \frac{1}{\alpha}\ln(1+\alpha\mu_i) + \ln\Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln\Gamma(y_i+1) - \ln\Gamma\left(\frac{1}{\alpha}\right)\right) \tag{4.4}$$

When one transforms this likelihood function into logarithmic likelihood, the exponent in the product of (4.4) falls away and the formulation transforms to the summation as in (4.5)

$$\mathcal{L}(\mu; y, \alpha) = \sum_{i=1}^{n} y_i \ln\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right) - \frac{1}{\alpha}\ln(1+\alpha\mu_i) + \ln\Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln\Gamma(y_i+1) - \ln\Gamma\left(\frac{1}{\alpha}\right) \tag{4.5}$$

Now one can distinguish again between the two cases where $y = 0$ and $y > 0$ which gives the formulation in (4.6). In this formulation p is a function of the independent variables x_i and its coefficients β_i in (4.7) which will be estimated during the regression estimation. The code used for the modelling can be found in Appendices A.8.1 and A.8.1.

$$\mathcal{L} = \begin{cases} \sum_{i=1}^{n} \left[\ln(p_i) + (1-p_i)\left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}}\right] & \text{if } y_i = 0 \\ \sum_{i=1}^{n} \left[\ln(p_i) + \ln\Gamma\left(\frac{1}{\alpha} + y_i\right) - \ln\Gamma(y_i+1) - \ln\Gamma\left(\frac{1}{\alpha}\right) + \left(\frac{1}{\alpha}\right)\ln\left(\frac{1}{1+\alpha\mu_i}\right) + y_i \ln\left(1 - \frac{1}{1+\alpha\mu_i}\right)\right] & \text{if } y_i > 0 \end{cases} \tag{4.6}$$

$$\begin{aligned} p &= \frac{1}{1+e^{-x_i'\beta}} \\ 1-p &= \frac{1}{1+e^{x_i'\beta}} \end{aligned} \tag{4.7}$$

For the occupancy rate, the situation is much simpler. The occupancy has a continuous linear relationship between the dependent and independent variables, which makes it suitable for multiple linear regression as in (4.1). The other assumptions of these modelling techniques are stated and checked in the next section for the occupancy rate statistic.

4.7.1. Assumptions of the Statistical Models

Zero Inflated Negative Binomial Regression

For the zero-inflated negative binomial regression, the following assumptions exist and have been checked where possible:

1. Y values are count observations with excessive zeros:
    - As can be verified in Figure 4.5.
2. Multiplicative effects of independent variables.
3. Linearity in model parameters.
4. The conditional variance of the outcome variable is greater than its conditional mean:
    - This was tested using an overdispersion test and was found to be significant.

Multiple Linear Regression
For the multiple linear regression model, the following assumptions exist and have been checked where possible:

1. Linearity in model parameters:
   - See the figures in Appendix A.7.1. Most of the variables are close to a linear relationship.
2. Endogeneity of regressors:
   - As discussed in Section 4.
3. Normality and homoscedasticity of the error term:
   - See assumption check in Section 5.2.1.
4. No autocorrelation:
   - See assumption check in Section 5.2.1.
5. No multicollinearity:
   - As checked in Section 4.6.6.

4.7.2. Model Fit Measurements

During the selection of the independent variables, a step-down process was used. This means starting with a model that includes all relevant independent variables and excluding them one by one based on the p-value of the T test and the corresponding measurement statistic for the fit of the model. This was repeated until all the variables in the model were significant or had no added value. For the zero-inflated binomial regression model, the pseudo-McFadden R-squared value (McFadden, 2021) was used as a measurement of model fit as in (4.7).

$$pseudo \quad R^2_{\text{McFadden}} = 1 - \frac{\log(L_c)}{\log(L_{\text{null}})} \tag{4.8}$$

Where L_c is the maximum likelihood value of the fitted model and L_{null} the maximum likelihood value of the null model. For the linear regression model, this was the adjusted R squared value as in (4.8) where n is the number of data points and k the number of independent variables and R^2 as in (4.9).

$$R^2_{adjusted} = 1 - \left[\frac{\left(1 - R^2\right)(n-1)}{n-k-1}\right] \tag{4.9}$$

where

$$R^2 = 1 - \frac{\sum\left(y_i - \hat{y}\right)^2}{\sum\left(y_i - \bar{y}\right)^2} \tag{4.10}$$

With the measurements of model fit, the relevant methods of the quantitative part of this study have been described. The next chapter will follow up on this and present the results obtained using these methods.

5 Results

This chapter will present and analyse the results of the study based on the methodology of Chapter 4. Hereby, it will start with discussing the results of the EV user model and the occupancy rate model. Thereafter, the results of these models will be analysed and compared. Next, the results of the comparison analysis of the MRA-Elektrisch region and the rest of the Netherlands will be presented. Using the findings of the previous sections, the results of a further analysis to improve policy will be presented.

5.1. EV Users Model Results

This section will summarise and present the relevant results gathered by a zero-inflated negative binomial model focused on explaining EV users, as stated in Section 4.3. The constructed model uses ten of the independent variables available as stated in Section 4.6. The other variables did not contribute significantly to model construction. Using the ten variables, the model has a pseudo McFadden R-squared value of 0.09. Although this is not high, the model is capable of explaining a significant part of the variance. Furthermore, it was found that the model was significantly better than the empty model that included only the intercept value with a p-value of approximately zero.

A zero-inflated negative binomial model consists of two submodels, as we have seen in Chapter 4. One submodel explains the excess zero values in the dependent variable, whether or not EV users are present in a neighbourhood. The other submodel tries to explain the count measure, in this case the number of EV users in a neighbourhood. The model was tuned on the significance of the independent variables using a step-down approach based. The final model description can be found in Appendix A.27 with the estimates in the logistic representation, the standard errors, and the p-values for all the independent variables. For the count submodel, nine variables were found to contribute significantly. These are the following:

- *Solar_power*: power of solar panel installations in kW.
- *Dist_school*: distance to nearest primary school in kilometers.
- *Perc_rental_house*: percentage of rental houses in a neighborhood.
- *Dist_hospital*: distance to nearest hospital in kilometers.
- *Perc_green_votes*: percentage of green votes.
- *Perc_low_educated*: percentage of low educated people.
- *Dist_shops*: average distance to shopping facilities (department stores, supermarket, pharmacy, and restaurant).
- *Perc_self_employed*: percentage of self employed people.
- *Perc_65_EO_yr*: percentage of the population 65 years or older.

And for the zero submodel, the following variables were found to contribute significantly:

- *Solar_power*: power of solar panel installations in kW.
- *Perc_rental_house*: percentage of rental houses in a neighborhood.
- *AVG_household_size*: average household size.
- *Dist_hospital*: distance to nearest hospital in kilometers.
- *Perc_green_votes*: percentage of green votes.
- *Dist_shops*: average distance to shop facilities (department stores, supermarket, pharmacy, and restaurant).
- *Perc_self_employed*: percentage of self employed people.
- *Perc_45_64_yr*: percentage of the population between 45 and 65 years or older.
- *Perc_65_EO_yr*: percentage of the population 65 years or older.

Using the EV users model, confidence intervals have been created for the coefficients of the variables to make the relations more robust using an unbiased method. Using a bootstrap method with 3000 replications and bias-corrected and accelerated (bca) bootstrap intervals, the 95% confidence intervals were created, as in (Garay et al., 2011). Subsequently, the confidence intervals were transformed using the exponential transformation so that the results could be interpreted, since the normal estimates are for the logistic version as in Equation (4.6). The 95% confidence intervals can be found in Table 5.1.

Table 5.1: EV user model confidence intervals

Variable	Submodel	Estimate	BCA lower bound	BCA upper bound
Intercept	Count	3.927	2.495	6.548
Solar_power	Count	1.001	1.001	1.001
Dist_school	Count	0.783	0.665	0.924
Perc_rental_house	Count	1.007	1.004	1.013
Dist_hospital	Count	0.953	0.936	0.964
Perc_green_votes	Count	1.036	1.019	1.046
Perc_low_educated	Count	0.078	0.023	0.184
Dist_shops	Count	0.902	0.834	1.017
Perc_self_employed	Count	1.011	1.000	1.025
Perc_65_EO_yr	Count	0.975	0.968	0.981
Theta	Count	1.952	1.654	2.561
Intercept	Zero	$7.9 * 10^{-11}$	≈ 0	$1.2 * 10^{121}$
Solar_power	Zero	1.002	≈ 0	1.749
Perc_rental_house	Zero	1.125	≈ 0	$5.6 * 10^{2}$
Avg_household_size	Zero	1.493	≈ 0	$9.7 * 10^{162}$
Dist_hospital	Zero	0.766	≈ 0	$3.3 * 10^{19}$
Perc_green_votes	Zero	0.778	≈ 0	$1.9 * 10^{15}$
Dist_shops	Zero	2.482	≈ 0	$8.6 * 10^{132}$
Perc_self_employed	Zero	1.078	≈ 0	$2.2 * 10^{9}$
Perc_45_64_yr	Zero	1.159	≈ 0	$6.1 * 10^{1}$
Perc_65_EO_yr	Zero	1.093	≈ 0	$6.5 * 10^{4}$

As can be seen in Table 5.1 the estimated intercept of the count model is 3.93. This shows the baseline number of EV users in a neighbourhood. The other values in the estimate column state the factors that the baseline number alters by increasing that independent variable. For example, the expected number of EV users increases by a factor of 1.036 per increase in the percentage of green votes in a neighbourhood, without changing any other variable. The confidence interval for this factor is 1.019-1.046. As can be seen in Figure 5.1 not all variables have the same effect on the expected number of EV users. Some have an increasing effect with factors greater than one, and others have a decreasing effect with factors smaller than one.

The independent variables *Dist_school*, *Dist_hospital*, *Perc_low_educated*, and *Perc_65_EO_yr* have only factors smaller than one in the 95% confidence area. This means that with increasing distance to primary schools and hospitals, the expected number of EV users decreases. Furthermore, as the percentage of low educated people increases and that in the age group of 65 years and older, the expected number of EV users also decreases in the confidence interval. For most of the other variables in the count submodel, it holds that an increase in that variable accounts for an increase in the expected number of EV users as well in the 95% confidence interval. The only exception is *Dist_shops* where the confidence interval reaches from smaller to larger than one. The exact relation between this variable and the EV users is unclear in the confidence interval, although the estimate indicates that the relationship is likely to be negative. In Section 5.3 more on the found relationships will be stated.

For the zero submodel, the interpretation works differently. The estimated intercept gives the baseline odds that a neighbourhood does not have electric vehicle users. In this case, these odds are $7.89 * 10^{-11}$ which gives a similar probability in this case. This indicates that the zero submodel in this case does not add much to the overall model, as the probability that a neighbourhood is not able to have EV users is close to zero. In this case, this makes complete sense due to the selection of neighbourhoods, as will be elaborated in Chapter 6 the Discussion.

Since the contribution of the zero model is small, it is not surprising that the confidence intervals of the coefficients are large. Due to the fact that the zero model can only be estimated on a limited part of the dataset, it is sensitive to changes in the data using a bootstrap method. All confidence intervals have one in their confidence interval, which decreases the reliability of the estimates. As said before, the contribution of the zero model is low to the overall model and indicates that the zero-inflated negative binomial model did not distinguish excess zeros. On this part of the model, no conclusions will be drawn. More on this in Chapter 6.

5.2. Occupancy Rate Model Results

This section will summarise and present the relevant results gathered by the multiple linear regression model that focusses on explaining the occupancy rate, as stated in Section 4.4. The model was created using the stepdown method as a variable selection approach. The model includes 11 of the independent variables available, since those had a significant contribution. The model came to an adjusted R-squared value of 0.365 which shows it can explain part of the variance. Furthermore, the model was found to be significantly better than the empty model that only has an intercept with a p-value of approximately zero. The final model description can be found in Appendix A.29 with the coefficient estimates, the standard error, and the p-values of the independent variables. The following independent variables were included in the model:

- *Num_residents*: number of residents.
- *Solar_power*: power of solar panel installations in kW.
- *Perc_green_votes*: percentage of green votes.
- *Perc_high_income*: percentage of high incomes.
- *Has_school*: elementary school located within one kilometer.
- *Avg_household_size*: average household size.
- *Perc_high_educated*: percentage of high educated people.
- *Population_density*: population density.
- *Perc_self_employed*: percentage of self employed people.
- *Perc_45_64_yr*: percentage of people in the age category 45-64 years.
- *Perc_65_EO_yr*: percentage of the population 65 years or older.

As with the EV users model, 95% confidence intervals have been constructed for the coefficients of the variables. This was done using bootstrapping with 3000 replications and bias-corrected and accelerated bootstrap intervals. The confidence intervals of the coefficients of the variables can be found in Table 5.1

Table 5.2: Occupancy rate model confidence intervals

Variable	Estimate	BCA lower bound	BCA upper bound
Intercept	0.183	0.071	0.301
Num_residents	$1.8 * 10^{-5}$	$1.3 * 10^{-5}$	$2.2 * 10^{-5}$
Solar_power	$-4.7 * 10^{-5}$	$-6.3 * 10^{-5}$	$-3.0 * 10^{-5}$
Perc_high_income	$3.4 * 10^{-3}$	$2.3 * 10^{-3}$	$4.6 * 10^{-3}$
Has_school	0.032	0.013	0.051
Avg_household_size	-0.063	-0.102	-0.-26
Perc_green_votes	$4.1 * 10^{-3}$	$2.7 * 10^{-3}$	$5.4 * 10^{-3}$
Perc_high_educated	0.168	0.070	0.264
Population_density	$1.1 * 10^{-5}$	$8.9 * 10^{-6}$	$1.4 * 10^{-5}$
Perc_self_employed	$1.9 * 10^{-3}$	$8.01 * 10^{-4}$	$3.1 * 10^{-3}$
Perc_45_64_yr	$-2.8 * 10^{-3}$	$-4.1 * 10^{-3}$	$-1.5 * 10^{-3}$
Perc_65_EO_yr	$-1.2 * 10^{-3}$	$-2.0 * 10^{-3}$	$-3.2 * 10^{-4}$

As can be seen in Table 5.2 in the *Estimate* column, not all variables have the same relationship with the occupancy rate of public chargers. The variables *Solar_power*, *Avg_household_size*, *Perc_45_64_yr* and *Perc_65_EO_yr* have a negative relationship with the average occupancy rate of public chargers in neighbourhoods. The other variables have a positive relationship with the average occupancy rate. For all variables, this unilateral relation holds for the 95% confidence interval. Furthermore, the confidence intervals are relatively small, making the relations more robust. In Section 5.3 more on these relationships will be stated.

5.2.1. Check of Assumptions

For the occupancy rate model, several assumptions had to be checked after the construction of the model. This section focusses on these assumptions and checks whether they apply. The first assumption that needs to be checked is the normality and homoskedasticity of the error terms. The residuals of the model have been plotted and can be found in Appendix A.38. This figure shows that the residuals are independent and identically distributed, as assumed. Furthermore, the residuals have been tested on normality. In Appendix A.37 it can be seen that the residuals are assumed to be normally distributed using the Kolmogorov-Smirnov test. Due to the large number of observations, this test was most applicable and, therefore, used in this case. Therefore, it can be concluded that these assumptions hold.

Furthermore, the assumption of no autocorrelation in the model. In Appendix A.34 a plot of the autocorrelation function (ACF) can be found. In Appendix A.35 a plot of the partial autocorrelation function (PACF) can be found. These figures show some lags that cross the significant threshold level. However, no apparent structure was detected. Subsequently, the Durbin-Watson test was used to test for significant first-order autocorrelation in the model. This test concluded that no first-order autocorrelation could be detected, see Appendix 5.2.1. Therefore, the assumption of no autocorrelation was concluded to be valid. The other assumptions of the model were checked beforehand in Section 4 which makes that the model was found applicable.

5.3. Relations between EV adoption and neighbourhoods

The results of the relations between the independent variables and the KPIs for the adoption of electric vehicles have been summarised in Table 5.3. In Table 5.3 the variables that contributed significantly to the models are stated with their corresponding relationship to the EV adoption KPIs. The table has two columns; *EV users* which is the count submodel of the EV user model and *Occupancy rate* which is the occupancy rate model. The table is filled on the basis of the confidence interval of the coefficients of the model. Green means a unilateral positive relation within the 95% confidence interval. The red colour means unilateral negative relations within the 95% confidence interval. The colour grey means that in the confidence interval both positive and negative relations occurred. The exact coefficient values have been let out, since including them would give a distorted view, since these values come from different models. Furthermore, these coefficients are only estimates made by two statistical models. See the disclaimer in Section 5.3.1.

Table 5.3: Summary of the relations

Variable	EV users number	Occupancy rate	In accordance with literature
Num_residents			Yes
Dist_school			Additional variable
Dist_hospital			Additional variable
Population_density			Yes
Avg_household_size			No
Perc_green_votes			Yes, confirmed positive
Has_public_point			Yes
Has_school			Additional variable
Perc_45_65_yr			Yes (disputable in literature)
Perc_65_EO_yr			Yes (disputable in literature)
Perc_self_employed			Yes
Perc_high_income			Yes
Perc_rental_house			No
Perc_high_educated			Yes
Perc_low_educated			Yes
Dist_shops			Disputable
Solar_power			Disputable

As can be seen in Table 5.3 the two models *EV users number* and *Occupancy rate* found relations in both models with some of the variables. This occurred for *Perc_self_employed*, *Perc_green_votes*, *Perc_65_EO_yr* and *Solar_power*. For the first three of these variables, the same relation was found for both the EV adoption KPIs. The higher the percentage of self-employed people in a neighbourhood, the better the EV adoption measured in both the KPI's used in this study. The same holds for the share of people who voted for green political parties, which implies that environmental belief would be a driver of EV adoption. Furthermore, it seems that the age group of 65 and older has a negative influence on the adoption of EVs measured both in the number of EV users and in the occupancy rate of the chargers.

However, the two models find different relations with *Solar_power*. The *EV users number* has a positive relationship, where *Occupancy rate* finds a negative relationship with the power of the solar panels present in a neighbourhood. This difference in the type of relation is remarkable. An explanation could be related to the dimension of private chargers, where private charger owners often also own solar panels as well (Wolterman et al., 2022).

For the other variables, only one of the two models found a relationship. Since the two models measure different KPI's, this does not come as a surprise. Still, similar relations can be distinguished in the models. The *EV users number* has a negative relationship with increasing distance to a primary school and the *Occupancy rate* has a positive relationship with the presence of a primary school in a neighbourhood. This both indicates that the presence of a school has a positive relation on EV adoption as measured in both KPI's.

Combining the relations of the variables *Num_residents* and *Population_density* with *Occupancy rate* it becomes clear that the more people live in a neighbourhood in absolute terms as in relative terms, the higher the adoption of electric vehicles measured in the occupancy rate.

Furthermore, when focussing on variables *Perc_high_educated* and *Perc_low_educated* it becomes clear that the level of education is a characteristic that describes the adoption of electric vehicles. Combining the insights of the two KPI's, it can be stated that a higher education level has a positive influence on the adoption of EVs. As found in the literature study in Chapter 3.

As stated in Table 5.3, most of the relationships found are supported by the literature. The two variables where different relations were observed, as expected from the literature, are the *Avg_household_size* and the *Perc_rental_house*. In the literature, larger households tend to adopt more electric vehicles, as seen in Table 4.1. This study found an opposite effect. An explanation for this contradictory behaviour could be the scope of this study. This study focusses solely on adoption at public charging points and is driven by charging transactions. The studies observed in the literature do not share this scope and are mainly based on survey data. For the *Perc_rental_house* relation contradicts also the findings in the literature. This could also be explained by the scope of this study that focusses solely on the adoption of electric vehicles by public charging points, where Shom et al. (2022) focusses on the adoption of electric vehicles in general.

Moreover, this study confirms the positive relationship between environmental awareness, as expressed in *Perc_green_votes*, and the adoption of electric vehicles. Although this relation was expected to hold, it was not yet confirmed by empirical research (Austmann, 2021). This study finds a positive relationship between both the number of electric vehicle adoptions and the occupancy rate of public chargers with the percentage of green votes in a neighbourhood. Furthermore, this study expands the range of independent variables used to explain electric vehicle adoption, as indicated by the systematic literature review of Austmann (2021) with proximity variables for schools and hospitals. Additionally, two of the relations found have disputable relations, which makes that the results cannot be compared with the literature. These are the variables *Dist_shops* and *Solar_power* that have ambiguous relations with the adoption of electric vehicles, as found in this study.

The effects of all independent variables in the two models have been visualised using plots of the confidence intervals. These statistical representations can be found in Figure 5.1 for *EV users number* and in Figure 5.2 for the *Occupancy rate*.

As can be seen in Figure 5.1 most variables have coefficients close to one, which makes them contribute only minor by increasing the independent variable. It should be noted that the plots are not corrected for variance but show actual confidence intervals. The independent variables that contribute the most per increase in unit are those of *Dist_school*, *Dist_shop* and *Perc_low_educated*. This last variable is remarkable, as it clearly has a different range than the other independent variables. Figure 5.1 implies that the percentage of low-educated people is a significant characteristic, in this case a barrier, of the adoption of EV as measured in this study.

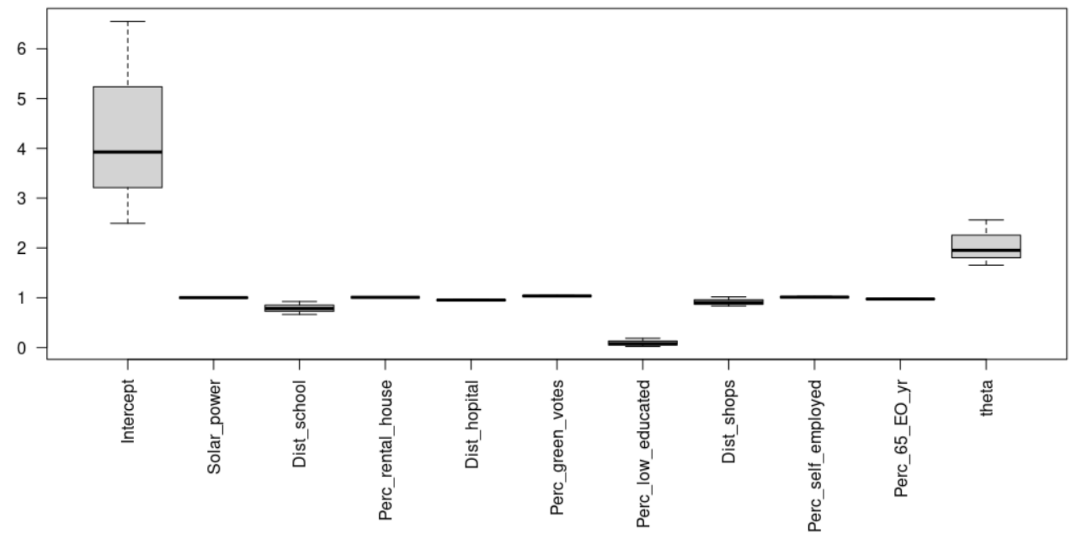

Figure 5.1: Plot of confidence intervals for the EV users model

Moreover, Figure 5.2 shows a similar behaviour for *Perc_high_educated* but in the other direction for the occupancy rate. Apparently, the level of education has a high influence on the adoption of EVs. A further analysis of the confidence intervals shows that the independent variables *Has_school* and *Avg_household_size* differ in the size of their coefficients. Again, the remark is made that the plots are not corrected for the variance, which can lead to a distorted view. For example, it makes sense that *Num_residents* and *Solar_power* have lower coefficients, since these independent variables have a much higher variance.

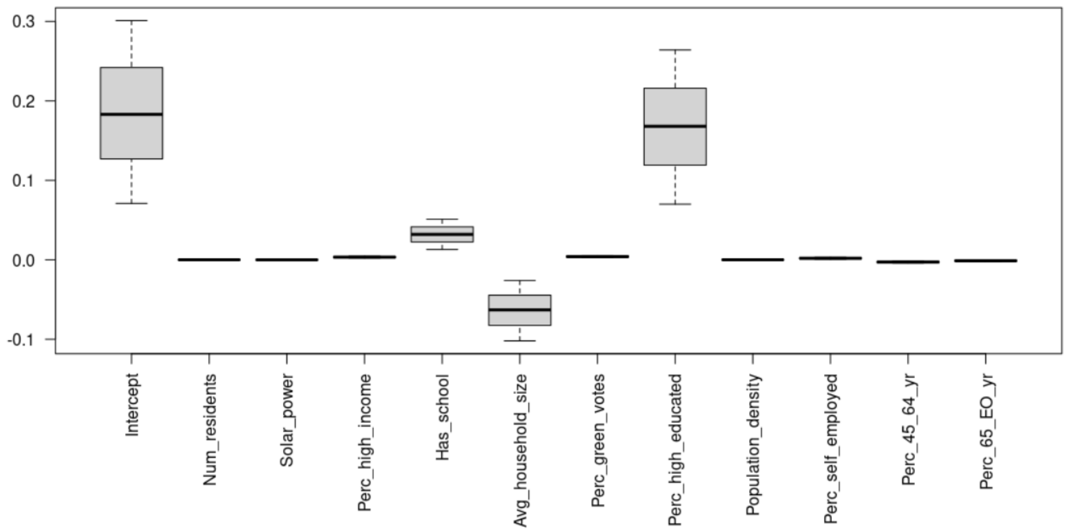

Figure 5.2: Plot of confidence intervals for the occupancy rate model

5.3.1. Disclaimer

The relations described in this section were found using statistical models. Although models can provide good insight and find significant relations, no model is correct. This means that one has to be very careful with conclusions based on such models. In this case, this does not mean that it cannot be said that the variables described in Section 4.6 for which no significant relation was found have no relation to the adoption of electric vehicles. It is not detected by the models used. Furthermore, since the world is far more complex than any model can handle, this means that, although the models found significant relations, this does not mean any causation or correlation in the real world. Moreover, no guarantees are made that the models used are, in this case, the best models for this task. With the use of domain understanding, these modelling techniques were found to be best suited according to the author. The same holds for the composition of the model. However, by definition, any model that describes a complex process in the real world can always be optimised.

5.4. Comparing Regions Analysis

Table 5.4 presents the results of the comparison analysis of the distributions of the MRA-Elektrisch region and the rest of the Netherlands for the independent variables used in this study. As can be seen in Table 5.4 only a portion of the independent variables have comparable distributions and population means in the two regions. However, a major part differs in population mean, distribution, or both between the two regions. Therefore, the two regions cannot be seen to be identical. Furthermore, the differences between these two regions in independent variables could also explain the differences in EV adoption between these two regions, as we have seen in Chapter 2. Further research could focus on this kind of causation. For this research, the difference of both dependent and independent variables means that the results presented in this study must be treated with care when using this study in the context of the Netherlands as a whole.

Table 5.4: Results region analysis

Continuous Variable	Comparable	Different Population Mean	Different Distribution
Num_residents			X
Dist_school		X	
Dist_hospital		X	X
Population_density		X	X
Avg_household_size	X		
Perc_green_votes	X		
Perc_45_64_yr	X		
Perc_65_EO_yr	X		
Perc_self_employed		X	X
Perc_high_income		X	X
Perc_rental_house	X		
Perc_high_educated		X	X
Perc_low_educated		X	X
Dist_shops		X	X
Solar_power		X	X

5.5. Further Analysis for Policy Recommendations

Using the relationships found between neighbourhood characteristics and adoption of EV in previous sections, relevant insights can be created to improve current policy on the implementation of public charging points. Based on the results of the model describing the number of electric vehicles users per neighbourhood, estimates can be made for the potential number of EV users for neighbourhoods in the MRA-Elektrisch region. These estimates are based on the coefficients found in the EV user model for the neighbourhood characteristics and the data set of the neighbourhoods with their characteristics. Using this potential EV adoption statistic and the actual measured adoption of electric vehicles at public charging points, the unused potential of neighbourhoods can be determined. This is assumed to be the difference between the estimate and the actual measured adoption. The unused potential scores for the neighbourhoods can be found in Figure 5.3 for the MRA-Elektrisch region. With these unused potential statistics, relevant insight can be created to improve the rolling out strategy of public charging points. The remainder of this section will perform such analysis.

The potential analysis was extended using the occupancy rate statistic created in Section 4.4. Using this occupancy rate, neighbourhoods can be identified with a high potential score and a high occupancy rate of the chargers present. In other words, these are neighbourhoods that already have a high demand for the current charging infrastructure and potential for additional EV adoption under the current conditions. Therefore, these neighbourhoods can be labelled neighbourhoods that might be prioritised when placing additional public chargers. This will improve the charging experience of current drivers and make the neighbourhoods ready for the near future. Neighbourhoods with a high occupancy rate have been identified as neighbourhoods with occupancy rates above the third quartile value of the occupancy rate distribution of the MRA-Elektrisch region, see Figure 4.7. The neighbourhoods that are suggested to prioritise in the additional placement of public chargers can be found in Figure 5.4.

The potential analysis has also been extended with the ratio of the number of EV users per charger. This ratio is an additional KPI to measure the adoption of EVs (Helmus & Van den Hoed, 2016). Combining this ratio with the potential score measured in another adoption dimension provides additional information on the number of chargers present. With this analysis, neighbourhoods can be identified where the current number of chargers does not meet the potential number of EV users in the future, as the ratio of EV users per charger will be critical in the future. The critical rate of EV users per charger is determined by the third quartile value of the current distribution in the MRA-Elektrisch region. The results of this analysis can be found in Figure 5.5. The code of this policy analysis can be found in Appendix A.8.1.

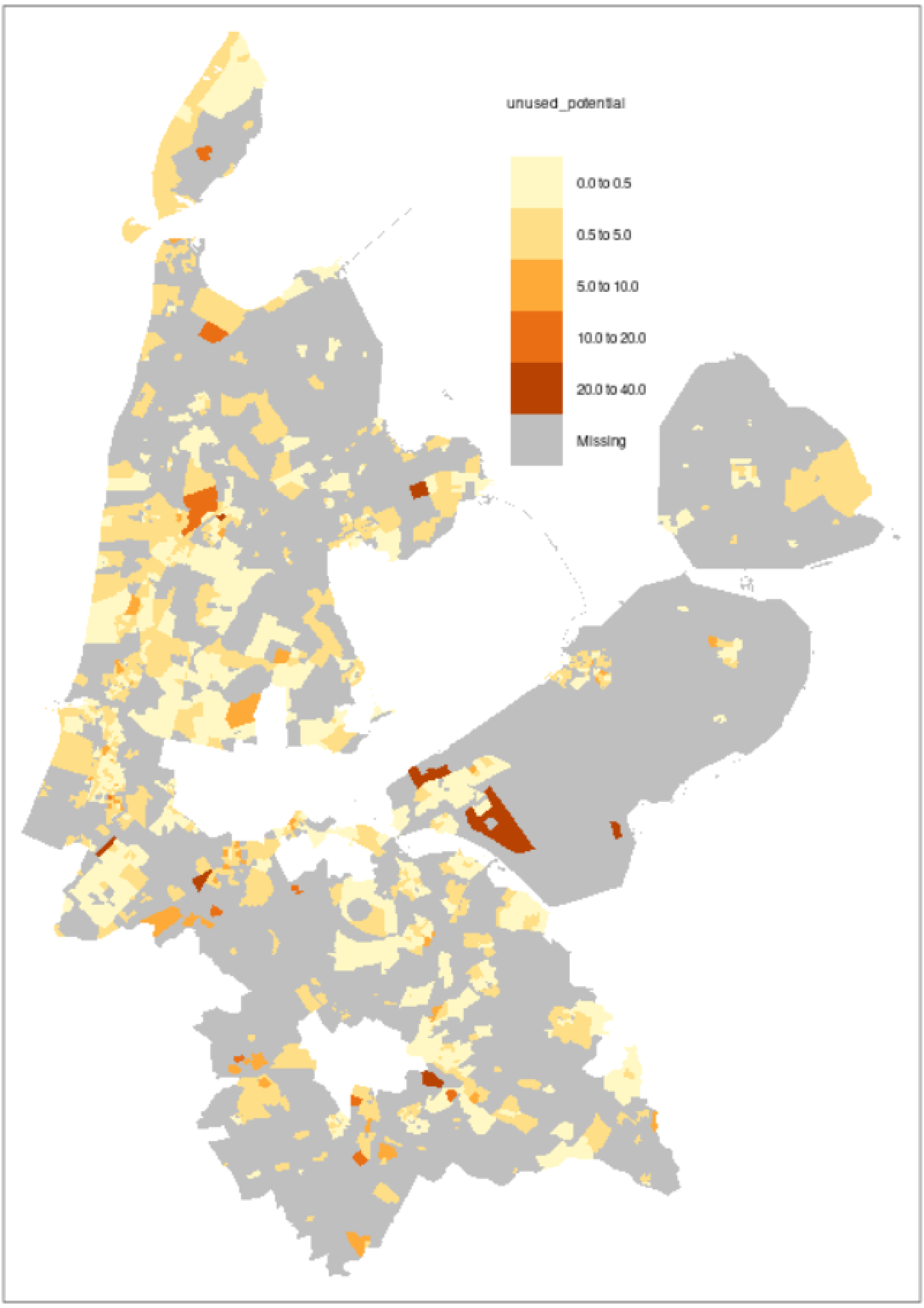

Figure 5.3: Unused potential in EV adoption

Figure 5.3 marks neighbourhoods with high potential adoption of EV users with a more red-like colour. As can be seen, a major part of the neighbourhoods have unused potential for EV adoption, these are the neighbourhoods with a colour different from light yellow. The legend represents the potential number of additional electric vehicle users in each neighbourhood. This additional number is small in most cases, however, in some neighbourhoods the unused potential is high. There are around ten neighbourhoods that show excessive potential with an estimate of ten additional users or more. These neighbourhoods are located throughout the MRA-Elektrisch region without showing some spatial pattern. Although the estimates are based on real data, this does not imply that these potential scores represent reality. However, these potential scores are the best indication available on the basis of these new insights so far.

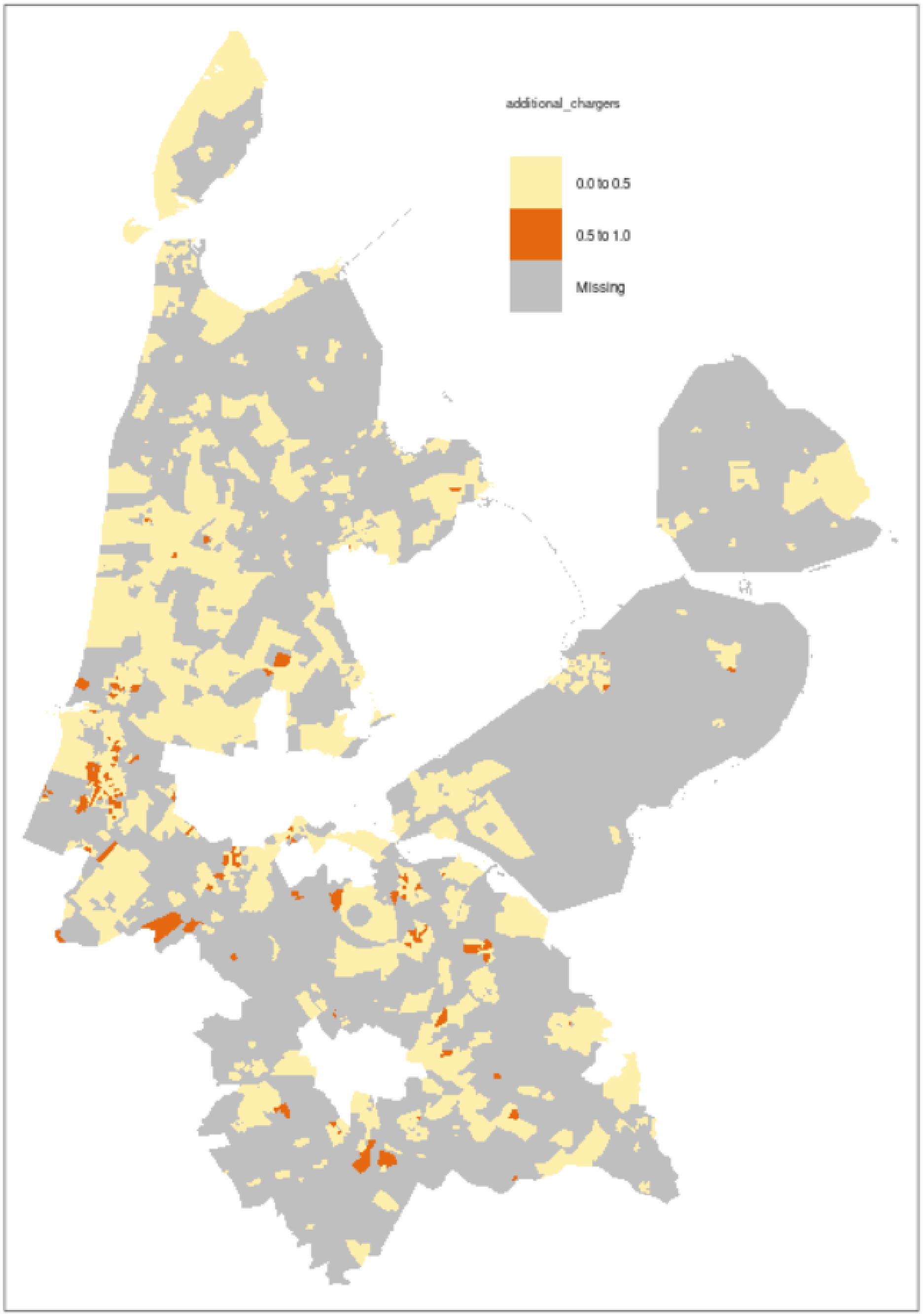

Figure 5.4: Neighbourhoods with potential for extra public charging points in the near future

Figure 5.4 marks the neighbourhoods that are suggested as priority for placing additional chargers in the near future with a red colour. These neighbourhoods are selected based on a high potential for EV adoption and a high occupancy rate of the current charging infrastructure. As can be seen in Figure 5.4, these neighbourhoods are mainly clustered around the cities of Utrecht and Amsterdam, which is in line with the spatial distribution of the occupancy rates in Figure 4.6. In the rest of the MRA-Elektrisch region, neighbourhoods are also suggested as priorities, but not in the same numbers. Furthermore, the number of suggested neighbourhoods is limited, which meet the time span of the analysis. Since the occupancy rate of the suggested neighbourhoods is already high, these neighbourhoods will need additional public charging stations in the near future.

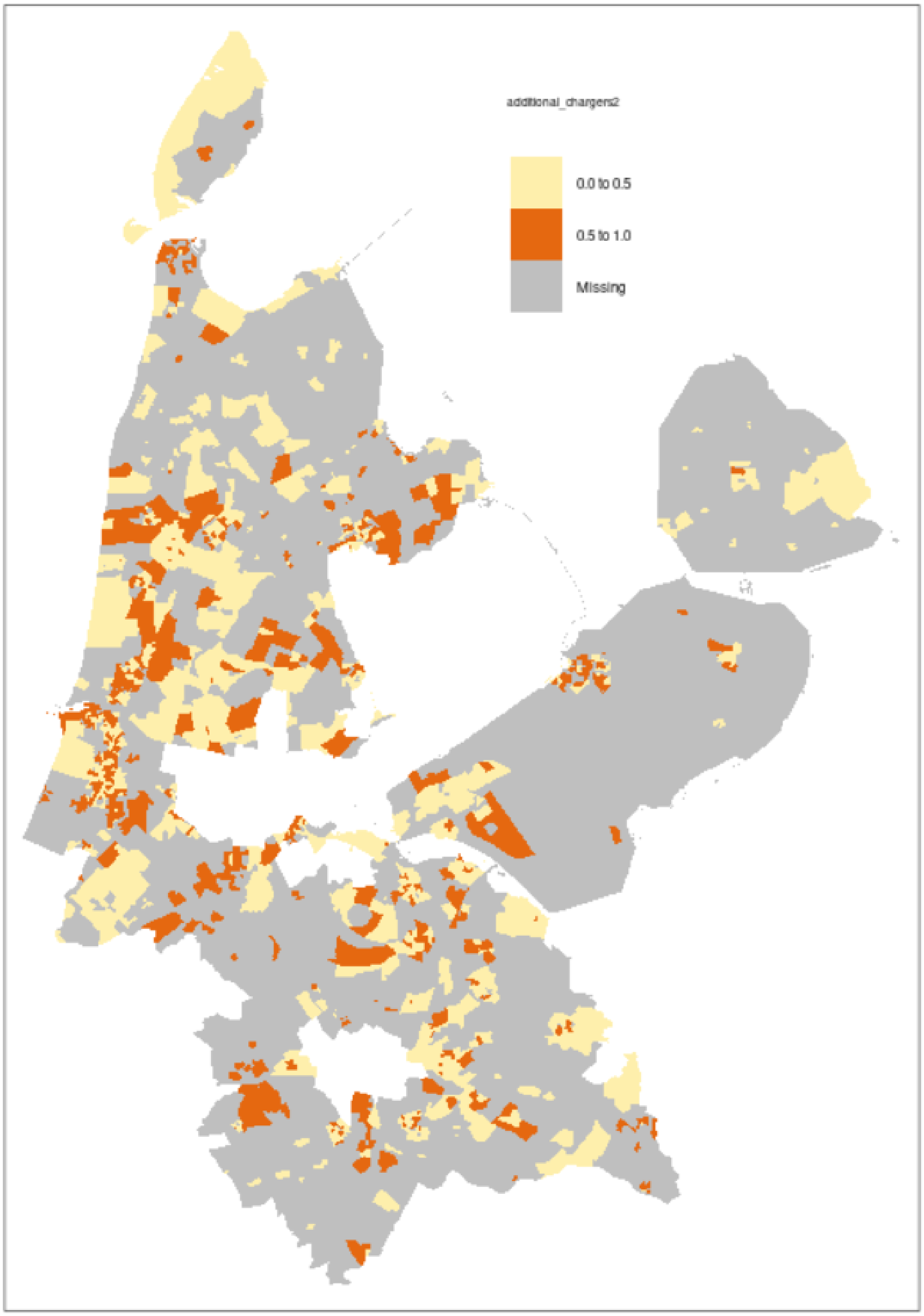

Figure 5.5: Neighbourhoods with potential for additional public charging points in the future

Figure 5.5 marks the neighbourhoods that are suggested to place additional public chargers in the future in red. These neighbourhoods are selected as having a high potential value for EV adoption and a high value for the ratio of potential EV users to available chargers. As can be seen, many neighbourhoods are marked as considerable for adding chargers. Furthermore, these neighbourhoods are spread over the entire MRA-Elektrisch region. Since this analysis focused on a time span further away in time, it makes sense that more neighbourhoods are suggested. Since the Netherlands is in the middle of the modal shift, more chargers will be needed, as described in Chapter 2. Comparing the suggested neighbourhoods of Figure 5.5 with the neighbourhoods suggested to prioritise in the near future, as in Figure 5.4 we see that the two classes do not always overlap, although most neighbourhoods do. In the case of limited resources, one might want to prioritise using both analyses.

6 Discussion

This study focusses on finding the relationships between the adoption of electric vehicles and the characteristics of the neighbourhood measured in the transaction data from public charging points. Therefore, this study uses several methods, including the creation of statistics for the adoption of electric vehicle users and the occupancy rate of public chargers at the neighbourhood level. Using these statistics, a zero-inflated negative binomial model and a multiple linear regression model were built to find the relationships between these statistics and the independent variables. This chapter will be a critical reflection on the methods used and their performance. Furthermore, it will state observations made during the quantitative study. Moreover, it will reflect on the contribution of this study as intended by the scope. Based on this section, the results of this study can be put in context.

6.1. EV Users Statistic

During the processing of data from the *DIM_CHARGEPOINT* and *DIM_LOCATION* tables of the *CHIEF* datamart it was observed that not all public chargers had the correct postal code, *PostalCode*, entry (IDO-LAAD, 2022). During the grouping of public chargers at the neighbourhood level using *ChargePoint_ID*, the values of the charging points per neighbourhood were found that were outliers to the rest of the distribution. Using a publicly accessible data source on the location of public charging points, the neighbourhoods with more than ten charging points were manually checked (ANWB, 2022). Some of them appeared to be correct, however, in most cases the grouping was invalid. Therefore, alternative mappings were considered making use of the longitude and latitude of the chargers. The same problem appeared that brought the conclusion that the data from these locations could be incorrect. Therefore, neighbourhoods that reported more than ten public charging points and were not in line with the second data source have been excluded from the data set. This is important, since the inclusion of them could mean adding incorrect outliers to the models. Checking all neighbourhoods against the secondary data source was not feasible, since no automation could be created for this task, considering more than a thousand neighbourhoods. Therefore, only possible outliers have been treated in this way. It is not clear whether there are other incorrect mappings in the analysis.

Another related point of discussion occurred during the mapping of EV users to neighbourhoods. According to the definition in Section 4.3 and Algorithm 1, one of the constraints is that an RFID identifier must charge more than five times a month at public charging points within a radius of 200 metres to be labelled as an EV user in that neighbourhood. Now let us analyse the following case: Consider an RFID identifier that charges eight times in a single month. These charge sessions occurred on three different chargers, all with a frequency lower than five. At the first charger, the RFID tag charged four times that month. At the second charger, the RFID tag charged three times that month. And at the third charger, the RFID tag charged only once in that month. What makes this case special is that the distance between chargers one and two is more than the 200 metres requested by the algorithm. However, charger three is located within the 200 metre radius of both the other charging points. According to Algorithm 1, the postal code of charging point three is used to map the EV user to a neighbourhood. Now, let us consider that these three charging points lie at the border of two neighbourhoods where charge points one and two lie in neighbourhood A and charge point three lies in neighbourhood

B. In this case, the RFID tag is defined as an EV user in neighbourhood B. However, seven of the eight charge sessions took place in neighbourhood A. Although this behaviour can only occur in special border cases, it shows that Algorithm 1 has its limitations in correctly assigning an RFID tag to a neighbourhood.

Furthermore, EV users are based on unique *RFID_skey* values of RFID tags. However, there is a difference between electric vehicle users and an RFID tag. One EV user can, for example, use multiple RFID tags, and one RFID tag can be used by multiple EV users. This study assumes that every EV user uses one RFID tag and that every RFID tag is used by a single EV user.

Finally, by extending the definition as stated in van Montfort et al. (2016) with the restriction of meeting all requirements at least four times in 2022, adoptions in the last three months of 2022 were excluded from the analysis. The choice was made to increase the quality of the EV users statistics above the increase in the number of EV users. The reason behind this is that the statistic of EV users in this study was measured over a longer period of time to exclude time patterns that can give distorted views. Therefore, some kind of restriction had to be made to extract regular EV users where the choice was given to the reasoning as stated in Section 4.3. On the other hand, possible additional UV users were missed when looking at the increasing charging transactions in the last months of 2022 in the Appendix A.2.

6.2. EV Users Model

During the modelling of EV users using the zero-inflated negative binomial model, several interesting observations were made. One of these was the occurrence of (quasi-)complete separation in the model when all the independent variables were included. Complete separation, also known as perfect separation, is the effect when an independent variable correctly allocates all observations to their corresponding group (Albert & Anderson, 1984). Quasi-complete separation is the case when this correct allocation can be performed using multiple variables. The appearance of (quasi-) complete separation in logistic modelling is undesirable, since logistic regression uses a maximum likelihood estimate for each variable, see Section 4.7, which in the case of complete separation does not exist, and therefore no estimation can be done.

There are several strategies to deal with quasi-complete separation. According to Bruin (2021) one of the applicable strategies for the EV user model without changing the statistical model is to do nothing and keep all the variables in the model. Since the maximum likelihood predictors of the other variables are still valid, the model produces results where only the coefficient of the variable causing the (quasi-) complete separation is incorrect. Another approach is to exclude the variable that causes (quasi-) complete separation, which solves the unreliable estimate for the variable. However, this can have its effects on the estimates of the other predictors. For the EV users model, a combination of both strategies was applied during the composition of the model. Therefore, no conclusions can be drawn about variables that were not included in the model. In the context of the UV user model, only conclusions can be drawn on the significant variables in the model with respect to the explained variance, which was expressed by a pseudo McFadden R-squared value of 0.09. This serves as an additional warning to the disclaimer in Chapter 5.

Another point of discussion is the intercept of the zero submodel of this model. As shown in Chapter 5 the transformed intercept is close to zero, which makes both the odds and the probability of a neighbourhood not capable of having EV users close to zero. The model is tuned to a selected set of neighbourhoods that meet the restriction of having residents, households, houses, and an available public charger in the neighbourhood. Therefore, this limited role of the zero submodel is obvious. One could argue that the zero-inflated modelling technique is excessive for this purpose when such restrictions have been made. However, based on Figure 4.5 it was observed that there might be excess zeros in the distribution. Furthermore, restrictions select neighbourhoods where actual people live, and a public charger is present, which does not directly imply EV adoption. It is reasonable that there are neighbourhoods where public charging points are placed, but there is no EV adoption, e.g., due to a failing strategic placement strategy. However, this was not observed in practise.

Not only did the zero submodel intercept have a minor contribution. The same was true for the coefficients of the variables in this submodel. Looking at the 95% confidence intervals, shown in Figure 5.1, it can be seen that the range of these intervals is large. This comes as no surprise since the zero submodel contributes negligible to the overall model. Therefore, in this study, the zero submodel is perceived as irrelevant and will not be used for any conclusions.

Finally, the overall zero-inflated negative binomial model had a McFadden pseudo R-squared value of 0.09. Although this seems very low, a pseudo R-squared value of 0.2-0.4 is interpreted as extremely well fitted (Domencich & McFadden, 1975). This adds some perspective to these statistics. Therefore, it can be said that the model contributes to explaining some of the variance and is therefore capable of providing insight.

6.3. Occupancy Rate

The occupancy rates of public chargers are estimated by taking the average time a charger was connected in 2022. Although this is a good measurement for the occupancy of public chargers, the purpose of a public charger is to charge and not to be connected. Therefore, the more relevant would be the ratio of time a charger actually charges to that when connected. In the context of this study, the *ChargeTimeHours* column in the CHIEF datamart, see Figure 4.3, describes the charging time of the charging sessions. However, the quality of the data was found to be insufficient. Therefore, the choice was limited and the connection time was used as the occupancy rate.

6.4. Occupancy Rate Model

The occupancy rate model is a version of a multiple linear regression model. This model explained a significant part of the variance with an adjusted R-squared value of 0.3654. Adjusted R-square values must be seen relative to the field for which the model is used. In social sciences, the normal R-squared value of the model would be considered weak moderate (Hair et al., 2011). Since the model was used to gain an understanding of the relationships between neighbourhood characteristics and occupancy rate, it was found to be sufficient for conclusions.

Other observations were made while checking the assumptions of the model. Although the Durbin-Watson test did not give any reason to assume that there would be autocorrelation in the first lag, see Figure A.36, the ACF and PACF reported significant correlations in the first lags. Since the Durbin-Watson test is used to detect autocorrelation in the first lag, which had the largest correlation in the ACF and PACF plots, the conclusion was drawn that this held for the other lags as well, since these correlations were even smaller. Therefore, it was found that the assumption of no autocorrelation was applicable.

Another observation was made while checking the residuals for normality. The corresponding normality test, the Kolmogorov-Smirnov test, indicated that the assumption was valid; the null hypothesis of being normally distributed could not be rejected. However, the p-value was very close to the rejection interval.

6.5. Research Scope

The intention of this study is to provide insights in relations between neighbourhood characteristics and EV adoption in the Netherlands for the public charging point deployment strategy. However, the data available for the quantitative part of this study was limited to the MRA-Elektrisch region, as shown in Figure 4.1. This means that the insights are based on this region. The MRA-Elektrisch region is further in modal shift than the rest of the country, as shown in Figures 2.4 & 2.5. It seems that this makes the MRA-Elektrisch region a less reliable sample for the Netherlands. However, since the MRA-Elektrisch region is further in the adoption of EV, the insights into the relations could be even more relevant for the rest of the country than for the MRA-Elektrisch region itself.

The MRA-Elektrisch region does not only differ in the progress of the modal shift with the rest of the Netherlands. It also has other characteristics measured in the independent variables used in this study. Therefore, the relationships found between neighbourhood characteristics and EV adoption must be used with care when placed in a larger perspective. This study tried to take this into account by testing independent variables on differences in distribution, as can be seen in Table 5.4 in Section 5.3. Although this gives an estimate of whether the distributions are comparable between the MRA-Elektrisch region and the rest of the Netherlands, it does not give insight in whether this affects the found relations in any way.

Moreover, a remark needs to be made about the mapping between the postcodes of the charging points and neighbourhoods. This was done using mapping tables from the year 2020. This means that new neighbourhoods that have been extended in the last years or did not exist in 2020 were not included in the study, as they did not meet the requirements stated in Section 4. Newer mapping tables were available, however, these did not include final versions of the data. This may lead to data loss or to a reduction in the quality of these data. Therefore, the choice was made to go for the most recent final version of these data, as data loss in practise was limited.

Another comment on the scope is that the EV user model used to estimate the potential for electric vehicle adoption appeared to be more biased toward the neighbourhoods of the larger cities. Looking at Figure A.28 the greatest prediction errors can be found for the neighbourhoods of the city of Almere, which is the eight largest city in the Netherlands. No other city of this magnitude was observed in the available data. This raises the concern of the model being over-fit on more rural areas, which decreases the estimation power for neighbourhoods located in the more urban areas.

Furthermore, this study only focusses on electric vehicle adoption measured in public charging transaction. Due to this method, the observed adoption of electric vehicles is only part of the total adoption of electric vehicles in the Netherlands. Including electric vehicle users who use private chargers was not possible with the chosen quantitative approach, since available data on this group is lacking. Therefore, the choice has been made to focus only on the adoption measured in public charging data. When interpreting the results of this study, this scope should be taken into account.

So, conclusions on relations between EV adoption and neighbourhood characteristics in the Netherlands have to be made very carefully. Since, the quantitative study was limited by the available data. Next chapter will conclude on the results considering these limitations and will answer the research question of this study.

7 Conclusions and Recommendations

This thesis focused on identifying the relationships between the adoption of electric vehicles at public charging points and characteristics of neighbourhoods. This was achieved by extending a definition of electric vehicle adoption in the literature and creating two regression models that focus on explaining the number of electric vehicle users per neighbourhood and the occupancy rate of public chargers at the neighbourhood level. Both statistics were created from transaction data from public chargers in the MRA-Elektrisch region. This chapter concludes this study by answering the research question:

> *"Which relations are there between neighbourhood characteristics and the electric vehicle adoption at public charging points in the Netherlands?"*

These conclusions will be twofold. First, conclusions will be drawn on the relations between EV adoption at public charging points and neighbourhood characteristics based on the results of a zero inflated negative binomial model, used for the analysis of EV users, and a multiple linear regression model, used for the analysis of occupancy rates. Subsequently, policy recommendations for municipalities in the Netherlands will be made based on the results of the further analysis for the improvement of policy, incorporating the results of the relations between the adoption of EV and the neighbourhoods. This chapter ends with suggestions for further research.

7.1. Relations between EV Adoption and Neighbourhoods

Several significant relationships have been found between neighbourhood characteristics and EV adoption measurements. Of the 25 variables found during the literature study, 17 were found to have a significant relationship with the number of EV users or the occupancy rate. Four of these variables were found to have a relation with both these EV adoption KPI's. In all the cases except one, the constructed 95% confidence interval of the coefficients of the variables showed an unambiguous relation. This means that in all cases except one, the direction of the relationship between the neighbourhood characteristic and the adoption of electric vehicles could be determined. The remainder of this section will describe these relations.

One of the relations observed is that between the number of citizens in a neighbourhood, *Num_residents*, and the average occupancy rate of public chargers in that neighbourhood. This was found to be a positive relation, where an increase in number of citizens is assumed to increase the average occupancy rate. The same holds for the population density, *Population_density*, which also has a positive relationship with the occupancy rate. Therefore, both the absolute and relative numbers of citizens have a positive influence on the occupancy rate of the chargers found in the MRA-Elektrisch region.

Another relation that was distinguished is that of the distance to the closest primary school and the number of adoptions of electric vehicles. The farther away this school is, the lower the expected number of electric vehicle users. Similar behaviour was found for the occupancy rate; when a school is present in a neighbourhood, the occupancy rate is expected to be higher. Both relationships indicate that the presence of a nearby school estimates an increase in the adoption of electric vehicles. This indicates a relationship between some kind of school-related characteristic and the adoption of EV.

Similar relations were found for variables that indicate the distance to other facilities. Among them, the relationship between the distance to the shopping facilities, *Dist_shops* and the number of electric vehicle users, although the confidence interval did not give an unambiguous direction. The distance to the closest hospital, *Dist_hospital*, showed the same negative relation as the distance to schools. Furthermore, when a public point (an amusement park, zoo, indoor playground, museum, or swimming pool) is present in a neighbourhood, *Has_public_point*, the expected number of EV users is also higher. It seems that the adoption of EV is greater in the neighbourhoods with more facilities present.

Another relation that was observed is that of the percentage of green votes in the last national elections, *Perc_green_votes*, and the adoption of electric vehicles. Both the number of EV adoptions and the occupancy rate have a significant positive relationship with this variable. The findings of these relationships are in line with literature (Austmann, 2021).

One more observation is the relationship between the total power of all solar panels in a neighbourhood, *Solar_power*, and EV adoption. The higher the total power, the higher the expected number of EV users. The contradiction is that the higher the total power, the lower the occupancy rate. This is remarkable. An explanation could be the presence of private chargers.

Moreover, there appears to be a negative relation between the percentage of people in the age group 45-64, *Perc_45_64_yr*, and 65+, *Perc_65_EO_yr*, and the occupancy rate of public chargers. Furthermore, this negative relation holds as well for the percentage of the population 65 + years of age and the expected number of EV users. The adoption of electric vehicles on public chargers appears to have a negative correlation with elderly people. An explanation could be that this group makes more use of private chargers, as indicated by Wolterman et al. (2022). And that the group 65+ years of age makes less use of (electric) vehicles.

Furthermore, the percentage of people with high income, *Perc_high_income* has a positive correlation with the occupancy rate. The same relation holds for the percentage of people who have higher education *Perc_high_educated*. Furthermore, there is a negative correlation between the percentage of low-educated people, *Perc_low_educated*, and the number of EV users. It seems that the level of education is a driver of the adoption of electric vehicles. Furthermore, it was found that this level of education has a greater impact than most of the other variables stated before.

Another relation found is between the percentage of self-employed people and the measurements of the adoption of EVs. Combining this with the previous relationship found, we see a trend of increased adoption of electric vehicles in neighbourhoods with a better financial welfare state (education, income and occupation).

The other relationships found are between the average size of the household, *Avg_household_size* and the occupancy rate, which was found to have a negative correlation. And the percentage of rental houses in a neighbourhood, *Perc_rental_house*, and the number of EV users. Which had a positive correlation. A possible explanation is that owners of larger bought houses tend to use private chargers more often.

Combining the relations, one could create a profile of a neighbourhood with high adoption of EVs via public charging points. This neighbourhood has more rental houses, smaller households, and more urban characteristics such as a large number of facilities. The residents in these neighbourhoods are younger in age, well-educated and have high incomes. In addition, residents are more likely to be self-employed and have green views. This resident profile comes close to that of a so-called starter. Although the above conclusions tend to support this profile, the relations only suggest that there is a correlation between the variables and EV adoption measurements. This does not have to imply that this is the correct interpretation of the underlying characteristics and, moreover, that these relations can be combined in this way. Furthermore, both the zero-inflated negative binomial model and the multiple linear regression model used for the establishing of these relations had goodness-of-fit measurements which indicated that the models did not fit extremely well. In addition, the results are based on a quantitative study that assumed that every EV user uses a single RFID tag to charge and that every RFID tag is used by a single EV user. And that the transaction data used for the quantitative study was of sufficient quality. Furthermore, the statistical models used in the quantitative study are assumed to be applicable in this context, as could be partially verified.

Overall, it can be concluded that the majority of variables found in the scientific literature are significant explainers for EV adoption in the MRA-Elektrisch region in the Netherlands. Moreover, most of the relations found are in agreement with the literature. The exceptions to this were found in the relationship between the average household size, where this study found a negative relationship instead of a positive relationship with the adoption of electric vehicles. And the relations between the percentage of rental houses, where this study found a positive rather than a negative relationship. Furthermore, this study confirmed a positive relationship between environmental awareness and EV adoption, which was expected but not yet empirically proved in the literature.

So, this research contributes to the scientific literature as a case study of the observed independent variables, indicating which and how these variables influence the EV adoption. In this way, this study contributes to the understanding of the overall modal shift. In addition, for most of the factors described in the scientific literature, a mapping has been constructed with publicly accessible data in the Netherlands.

7.2. Policy Recommendations

The findings in this study are translated into recommendations for municipal rollout strategies for the placement of public charging points. These recommendations are based on the relationships found between EV adoption and neighbourhood characteristics in the MRA-Elektrisch region, and the analysis focussing on the identification of neighbourhoods that should be prioritised in this region. In the policy recommendations, a distinction will be made between general policy recommendations relevant for all municipalities in the Netherlands and municipalities located in the MRA-Elektrisch region.

General Policy Recommendations

- A general recommendation for the implementation of public charging points is the recognition of the target audience. The profile found of EV users using public charging points clearly distinguishes itself from the profile of EV users who mostly use private chargers, as sketched by (Wolterman et al., 2022). Although profiling is a dangerous method, especially for policymaking, it is important to understand that the demand for public chargers is not homogeneous. The demand for public chargers is higher in neighbourhoods with certain characteristics (more rental houses, smaller households, more facilities; residents who are well-educated, with high incomes, are more likely to be self-employed and to have green views). Therefore, the current policy of placing public charging points on demand is a good method to meet this heterogeneous demand.

- The placement strategy by demand can be improved by adding a potential analysis that indicates the potential of a requested location. When this potential is high, the municipality might consider placing a public charger with higher capacity. In this way, the charging infrastructure is made more robust.

- Furthermore, for strategic placement of public charging points, it is also recommended to perform this potential analysis and check whether the suggested location has sufficient demand. The relationship found between the occupancy rate of public chargers and the presence of public points in neighbourhoods, *Has_public_point*, indicates that the occupancy rate in locations generally chosen by a strategic placement is higher. This shows that this strategic placement strategy seems to work well.

- To improve the prediction of EV adoption in the Netherlands, it is suggested that Buurtprognoses by NAL (2022a) update their predictions based on the findings of this study. This can be done by adding the relationships between neighbourhood characteristics and current adoption of electric vehicles, as found in this study. Additionally, an analysis can be performed to determine whether and how these relations will change in the future.

- In addition, it is recommended that close cooperation with grid providers is continued due to the capacity shortage of the grid in many parts of the country. These grid providers have the power to deny a connection request when the electricity grid does not have sufficient capacity. Therefore, municipalities should actively support the upgrade of the electricity grid.

Municipalities Located in MRA-Elektrisch

- Based on the findings of this study, it is suggested to place more public charging points in neighbourhoods with a higher number of citizens. For the MRA-Elektrisch region, the higher the number of citizens and the density of the population in a neighbourhood, the higher the occupancy rate of public chargers. This indicates that public charging points are not yet evenly distributed across the population. Combining this finding with the current municipal rollout strategy, it is suggested that requests for public charges in neighbourhoods with a higher number of citizens are prioritised. Furthermore, it is suggested that municipalities consider placing charging points with higher capacity in such neighbourhoods. By enhancing the capacity in the initial placement, resources are saved in the future.

- Another recommendation is the prioritisation of the neighbourhoods indicated in Figure 5.4 for the placement of additional chargers. Based on further analysis for policy improvement, these neighbourhoods are marked as locations with unused potential for EV adoption combined with high demand for the current charging infrastructure. Therefore, it is assumed that in the near future the charging infrastructure in these neighbourhoods will not be sufficient. To address this, these neighbourhoods should be prioritised in the placement of new public charging stations in the near future.

- Additionally, it is suggested that the municipalities of MRA-Elektrisch review their current public charger plans using Figure 5.5. The selected neighbourhoods in this figure show the potential for additional charging infrastructure under current conditions. This has been estimated based on the ratio of potential EV users per charging point, where the selected neighbourhoods have ratios higher than the third quartile of the current distribution. Using Figure 5.5 neighbourhoods could be made more future-proof by increasing the capacity of the planned public chargers in these neighbourhoods. Adding additional chargers could also be considered when no neighbourhood is prioritised in the municipality by the previous recommendations.

7.3. Suggestions for Further Research

Based on the methodology and results of this study, suggestions will be made for further research. The purpose of this study was to provide insights in relations between electric vehicle adoption and neighbourhood characteristics using quantitative research. However, by adjusting the scope of this research, more interesting insights can be created, which are beneficial for the improvement of the Dutch EV policy.

The first suggestion is to include the adoption of electric vehicles through private chargers in research on electric vehicle adoption. In the Netherlands, the majority of EV adoptions occur via private chargers, and therefore this group should not be forgotten in further research for the improvement of Dutch EV policy. For this inclusion, it is necessary to connect users using private chargers to neighbourhoods within the applicable privacy legislation. This mapping could be created using data on registrations of electric cars on neighbourhood level combined with the statistic of electric vehicle adoptions using public chargers of this study. Data on electric car registrations are already available on higher aggregate levels in the Netherlands.

This study used a zero-inflated negative binomial model to distinguish between neighbourhoods with and without possible adoption of electric vehicles. The zero submodel, for the detection of a neighbourhood without adoption of electric vehicles, did not provide insights due to restrictions on the selected neighbourhoods. For further research, it is suggested to give up such limitations and see if a model can be created to identify neighbourhoods where EV adoption does not occur through public charging points.

Another suggestion for further research is the inclusion of the placement strategy per neighbourhood as an independent variable and test whether the strategy used has an effect on the adoption of EV in a neighbourhood. In this way, the different rollout policies can be evaluated and compared. This research will benefit the overall placement in the Netherlands, since data-driven conclusions can be drawn on which strategy performs best in the adoption of EVs.

Another suggestion is the extension of the observed region. Although a significant part of the Netherlands has been analysed in this study, this region is not very compatible as a sample for the Netherlands. The MRA-Elektrisch region differs both in EV adoption and neighbourhood characteristics from the rest of the country. Furthermore, the only major city in the observed region appears to have a large prediction error in the EV user model. Therefore, it is suggested to perform a similar analysis on a more heterogeneous sample of the Netherlands.

The last suggestion given is to use the relationships found in this study and use them in a more dynamic modelling technique, such as system dynamics, to explore the options of the Dutch government to stimulate the adoption of electric vehicles in neighbourhoods where it is behind on average. Using a more dynamic modelling technique, these options can be evaluated over time.

Bibliography

Agreement, P. (2015). Paris agreement. *Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change (21st Session, 2015: Paris). Retrived December*, *4*, 2017.

Albert, A., & Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, *71*(1), 1–10.

ANWB. (2022). Hier staan de laadpalen [Accessed: 2022-12-10]. https://www.anwb.nl/auto/elektrisch-rijden/waar-staan-de-oplaadpunten

Austmann, L. M. (2021). Drivers of the electric vehicle market: A systematic literature review of empirical studies. *Finance Research Letters*, *41*, 101846.

Axsen, J., Goldberg, S., & Bailey, J. (2016). How might potential future plug-in electric vehicle buyers differ from current “pioneer” owners? *Transportation Research Part D: Transport and Environment*, *47*, 357–370.

Ben-Akiva, M. E., Lerman, S. R., Lerman, S. R., et al. (1985). *Discrete choice analysis: Theory and application to travel demand* (Vol. 9). MIT press.

Berkeley, N., Jarvis, D., & Jones, A. (2018). Analysing the take up of battery electric vehicles: An investigation of barriers amongst drivers in the uk. *Transportation Research Part D: Transport and Environment*, *63*, 466–481.

Bruin, J. (2021). *Faq what is complete or quasi-complete separation in logistic/probit regression and how do we deal with them?* https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faqwhat-is-complete-or-quasi-complete-separation-in-logisticprobit-regression-and-how-do-we-deal-with-them/

CBS. (2020). Buurt, wijk en gemeente 2020 voor postcode huisnummer [Accessed: 2022-12-15]. https://www.cbs.nl/nl-nl/maatwerk/2020/39/buurt-wijk-en-gemeente-2020-voor-postcode-huisnummer

CBS. (2022a). Buurt [Accessed: 2022-12-15]. https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/buurt

CBS. (2022b). Verdeling personenauto’s naar provincie, 2020-2021 [Accessed: 2022-11-01]. https://www.cbs.nl/nl-nl/maatwerk/2022/20/verdeling-personenauto-s-naar-provincie-2020-2021

CBS. (2022c). Wijk- en buurtkaart 2020 [Accessed: 2022-12-15]. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2020

CBS. (2022d). Zonnestroom; vermogen en vermogensklasse, bedrijven en woningen, regio [Accessed: 2022-12-15]. https://opendata.cbs.nl/%5C#/CBS/nl/dataset/85005NED/table

Chen, T. D., Wang, Y., & Kockelman, K. M. (2015). Where are the electric vehicles? a spatial model for vehicle-choice count data. *Journal of Transport Geography*, *43*, 181–188.

Chen, Z., Liu, W., & Yin, Y. (2017). Deployment of stationary and dynamic charging infrastructure for electric vehicles along traffic corridors. *Transportation Research Part C: Emerging Technologies*, *77*, 185–206.

Coffman, M., Bernstein, P., & Wee, S. (2015). Factors affecting ev adoption: A literature review and ev forecast for hawaii. *Electric Vehicle Transportation Center*, (April 2015), 1–36.

Consul, P., & Famoye, F. (1992). Generalized poisson regression model. *Communications in Statistics-Theory and Methods*, *21*(1), 89–109.

Crooks, A. T., & Heppenstall, A. J. (2012). Introduction to agent-based modelling. In *Agent-based models of geographical systems* (pp. 85–105). Springer.

Domencich, T. A., & McFadden, D. (1975). *Urban travel demand-a behavioral analysis* (tech. rep.).

Dong, Y. (2018). Electric vehicle charging station quantity forecasting model. *AIP Conference Proceedings*, *1971*(1), 040020.

Egnér, F., & Trosvik, L. (2018). Electric vehicle adoption in sweden and the impact of local policy instruments. *Energy policy*, *121*, 584–596.

European Commission. (2020). Delivering the european green deal [Accessed: 2022-11-27]. https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/delivering-european-green-deal_en

European Commission. (2021). Fit for 55 [Accessed: 2022-11-27]. https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition

European Commission. (2022). Recharging systems [Accessed: 2022-11-28]. https://alternative-fuels-observatory.ec.europa.eu/general-information/recharging-systems

European Environment Agency. (2022a). Decarbonising road transport — the role of vehicles, fuels and transport demand [Accessed: 2022-11-27]. https://www.eea.europa.eu/publications/transport-and-environment-report-2021/download

European Environment Agency. (2022b). New registrations of electric vehicles in europe [Accessed: 2022-11-01]. https://www.eea.europa.eu/ims/new-registrations-of-electric-vehicles

EV-Volumes. (2022). Global ev sales for 2022 h1 [Accessed: 2022-11-01]. https://www.ev-volumes.com/country/total-world-plug-in-vehicle-volumes/

Funke, S. Á., Sprei, F., Gnann, T., & Plötz, P. (2019). How much charging infrastructure do electric vehicles need? a review of the evidence and international comparison. *Transportation research part D: transport and environment*, *77*, 224–242.

Garay, A. M., Hashimoto, E. M., Ortega, E. M., & Lachos, V. H. (2011). On estimation and influence diagnostics for zero-inflated negative binomial regression models. *Computational Statistics & Data Analysis*, *55*(3), 1304–1318.

Gemeente Amsterdam. (2022). Ik moet tijdelijk mijn huis uit vanwege renovatie, verbouwing of sloop. moet ik dit doorgeven aan de gemeente? [Accessed: 2022-12-19]. https://www.amsterdam.nl/veelgevraagd/?caseid=%5C%7B85206F10-BD07-466E-A492-91AEDB0E513%5C5%5C%7D

Gupta, R. S., Tyagi, A., & Anand, S. (2021). Optimal allocation of electric vehicles charging infrastructure, policies and future trends. *Journal of Energy Storage*, *43*, 103291.

Haidar, B., & Rojas, M. T. A. (2022). The relationship between public charging infrastructure deployment and other socio-economic factors and electric vehicle adoption in france. *Research in Transportation Economics*, 101208.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). Pls-sem: Indeed a silver bullet. *Journal of Marketing theory and Practice*, *19*(2), 139–152.

Hardman, S. (2019). Understanding the impact of reoccurring and non-financial incentives on plug-in electric vehicle adoption–a review. *Transportation Research Part A: Policy and Practice*, *119*, 1–14.

He, Z., Zhou, Y., Chen, X., Wang, J., Shen, W., Wang, M., & Li, W. (2022). Examining the spatial mode in the early market for electric vehicles adoption: Evidence from 41 cities in china. *Transportation Letters*, *14*(6), 640–650.

Helmus, J., & Van den Hoed, R. (2016). Key performance indicators of charging infrastructure. *World Electric Vehicle Journal*, *8*(4), 733–741.

Hermans, L., Kwakkel, J., Thissen, W., Koppenjan, J., & Bots, P. (2010). Policy analysis of multi-actor systems. *The hague: LEMMA*.

Hoekstra, A., & Refa, N. (2017). Characteristics of dutch ev drivers.

IDO-LAAD. (2022). Chief.dm for researchers [R4. CHIEF.DM for researchers for the new DataMart structure.].

International Energy Agency. (2022). Trends in charging infrastructure [Accessed: 2022-11-27]. https://www.iea.org/reports/global-ev-outlook-2022/trends-in-charging-infrastructure

Javid, R. J., & Nejat, A. (2017). A comprehensive model of regional electric vehicle adoption and penetration. *Transport Policy*, *54*, 30–42.

Kaufmann, R., Newberry, D., Xin, C., Gopal, S., et al. (2021). Feedbacks among electric vehicle adoption, charging, and the cost and installation of rooftop solar photovoltaics. *Nature Energy*, *6*(2), 143–149.

Kiesklimaat. (2021). Hoe stemt men in de 2e kamer tot aan de verkiezingen in 2021? [Accessed: 2022-12-20]. https://kiesklimaat.nl/view_klimaatlabel_2021

Köhler, J., Turnheim, B., & Hodson, M. (2020). Low carbon transitions pathways in mobility: Applying the mlp in a combined case study and simulation bridging analysis of passenger transport in the netherlands. *Technological Forecasting and Social Change*, *151*, 119314.

Kumar, R. R., & Alok, K. (2020). Adoption of electric vehicle: A literature review and prospects for sustainability. *Journal of Cleaner Production*, *253*, 119911.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1–14.

Li, S., & Zhou, Y. (2015). Dynamics of technology adoption and critical mass: The case of us electric vehicle market. *Available at SSRN 2672065*.

Liao, F., Molin, E., & van Wee, B. (2017). Consumer preferences for electric vehicles: A literature review. *Transport Reviews*, *37*(3), 252–275.

Liu, X., Roberts, M. C., & Sioshansi, R. (2017). Spatial effects on hybrid electric vehicle adoption. *Transportation Research Part D: Transport and Environment*, *52*, 85–97.

Maybury, L., Corcoran, P., & Cipcigan, L. (2022). Mathematical modelling of electric vehicle adoption: A systematic literature review. *Transportation Research Part D: Transport and Environment*, *107*, 103278.

McFadden, D. (2021). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In *Behavioural travel modelling* (pp. 279–318). Routledge.

Melliger, M. A., van Vliet, O. P., & Liimatainen, H. (2018). Anxiety vs reality–sufficiency of battery electric vehicle range in switzerland and finland. *Transportation Research Part D: Transport and Environment*, *65*, 101–115.

Mersky, A. C., Sprei, F., Samaras, C., & Qian, Z. S. (2016). Effectiveness of incentives on electric vehicle adoption in norway. *Transportation Research Part D: Transport and Environment*, *46*, 56–68.

Morton, C., Lovelace, R., & Anable, J. (2017). Exploring the effect of local transport policies on the adoption of low emission vehicles: Evidence from the london congestion charge and hybrid electric vehicles. *Transport Policy*, *60*, 34–46.

Mukherjee, S. C., & Ryan, L. (2020). Factors influencing early battery electric vehicle adoption in ireland. *Renewable and Sustainable Energy Reviews*, *118*, 109504.

Mulholland, E., Tattini, J., Ramea, K., Yang, C., & Gallachóir, B. P. Ó. (2018). The cost of electrifying private transport–evidence from an empirical consumer choice model of ireland and denmark. *Transportation Research Part D: Transport and Environment*, *62*, 584–603.

NAL. (2019). Nationale agenda laadinfrastructuur [Accessed: 2022-11-27]. https://www.klimaatakkoord.nl/binaries/klimaatakkoord/documenten/publicaties/2019/06/28/achtergrondnotitie-mobiliteit-laadinfrastructuur/Nationale+Agenda+Laadinfrastructuur.pdf

NAL. (2022a). Buurtprognoses [Accessed: 2022-12-06]. https://www.agendalaadinfrastructuur.nl/ondersteuning+gemeenten/buurtprognoses/default.aspx

NAL. (2022b). Monitoring landelijk [Accessed: 2022-10-20]. https://www.agendalaadinfrastructuur.nl/monitoring+2021/monitoring+landelijk/default.aspx

Narassimhan, E., & Johnson, C. (2018). The role of demand-side incentives and charging infrastructure on plug-in electric vehicle adoption: Analysis of us states. *Environmental Research Letters*, *13*(7), 074032.

Nayum, A., Klöckner, C. A., & Mehmetoglu, M. (2016). Comparison of socio-psychological characteristics of conventional and battery electric car buyers. *Travel Behaviour and Society*, *3*, 8–20.

Netbeheernederland. (2023). Capaciteitskaart afname elektriciteitsnet [Accessed: 2023-01-04]. https://capaciteitskaart.netbeheernederland.nl/

Netherlands Enterprise Agency. (2019). Electric vehicle charging [Accessed: 2022-11-28]. https://nklnederland.nl/wp-content/uploads/2021/12/Electric_Vehicle_Charging_-_Definitions_and_Explanation_-_january_2019.pdf

Netherlands Enterprise Agency. (2022). Electric vehicles statistics in the netherlands [Accessed: 2022-11-01]. https://www.rvo.nl/onderwerpen/elektrisch-rijden/cijfers

Nian, V., Hari, M., & Yuan, J. (2019). A new business model for encouraging the adoption of electric vehicles in the absence of policy support. *Applied energy*, *235*, 1106–1117.

Open State Foundation. (2021). Verkiezingsuitslagen tweede kamerverkiezingen 2021 [Accessed: 2022-12-19]. https://data.openstate.eu/dataset/verkiezingsuitslagen-tweede-kamerverkiezingen-2021

Parlement.com. (2022). Opkomst bij tweede kamerverkiezingen [Accessed: 2022-12-19]. https://www.parlement.com/id/vh8lnhrp8wsz/opkomst_bij_tweede_kamerverkiezingen

Planbureau voor de Leefomgeving. (2021). Nederland fit for 55 [Accessed: 2022-11-27]. https://www.pbl.nl/sites/default/files/downloads/pbl-2021-nederland-fit-for-55_4697.pdf

Plötz, P., Schneider, U., Globisch, J., & Dütschke, E. (2014). Who will buy electric vehicles? identifying early adopters in germany. *Transportation Research Part A: Policy and Practice*, *67*, 96–109.

Poole, M. A., & O'Farrell, P. N. (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 145–158.

Priessner, A., Sposato, R., & Hampl, N. (2018). Predictors of electric vehicle adoption: An analysis of potential electric vehicle drivers in austria. *Energy policy*, *122*, 701–714.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Richardson, G. P. (2011). Reflections on the foundations of system dynamics. *System dynamics review*, *27*(3), 219–243.

Ridout, M., Hinde, J., & Demétrio, C. G. (2001). A score test for testing a zero-inflated poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*, *57*(1), 219–223.

Rijksoverheid. (2019a). Klimaatakkoord [Accessed: 2022-11-27]. https://www.klimaatakkoord.nl/documenten/publicaties/2019/06/28/klimaatakkoord

Rijksoverheid. (2019b). Klimaatwet [Accessed: 2022-11-27]. https://wetten.overheid.nl/BWBR0042394/2022-03-02

Rijksoverheid. (2021). Klimaatnota 2021 [Accessed: 2022-11-27]. https://open.overheid.nl/repository/ronl-c6578318-bfa3-47b5-9c8d-85d72f817646/1/pdf/Klimaatnota%5C%202021.pdf

Rijksoverheid. (2022a). Locaties milieuzones [Accessed: 2022-12-15]. https://www.milieuzones.nl/locaties-milieuzones

Rijksoverheid. (2022b). Overheid stimuleert milieuvriendelijker rijden [Accessed: 2022-04-26]. https://www.rijksoverheid.nl/onderwerpen/auto/overheid-stimuleert-milieuvriendelijker-rijden

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432–448). Routledge.

Secinaro, S., Calandra, D., Lanzalonga, F., & Ferraris, A. (2022). Electric vehicles' consumer behaviours: Mapping the field and providing a research agenda. *Journal of Business Research*, *150*, 399–416.

Shankar, A., & Kumari, P. (2019). Exploring the enablers and inhibitors of electric vehicle adoption intention from sellers' perspective in india: A view of the dual-factor model. *International Journal of Nonprofit and Voluntary Sector Marketing*, *24*(4), e1662.

Shom, S., James, K., & Alahmad, M. (2022). Understanding the correlation of demographic features with bev uptake at the local level in the united states. *Sustainability*, *14*(9), 5016.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology*, *70*, 747–770.

Sierzchula, W., Bakker, S., Maat, K., & Van Wee, B. (2014). The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy policy*, *68*, 183–194.

Smith, B., Olaru, D., Jabeen, F., & Greaves, S. (2017). Electric vehicles adoption: Environmental enthusiast bias in discrete choice models. *Transportation Research Part D: Transport and Environment*, *51*, 290–303.

Sovacool, B. K., Kester, J., Noel, L., & de Rubens, G. Z. (2018). The demographics of decarbonizing transport: The influence of gender, education, occupation, age, and household size on electric mobility preferences in the nordic region. *Global Environmental Change*, *52*, 86–100.

Stockkamp, C., Schäfer, J., Millemann, J. A., & Heidenreich, S. (2021). Identifying factors associated with consumers' adoption of e-mobility—a systematic literature review. *Sustainability*, *13*(19), 10975.

Straka, M., De Falco, P., Ferruzzi, G., Proto, D., Van Der Poel, G., Khormali, S., & Buzna, L. (2020). Predicting popularity of electric vehicle charging infrastructure in urban context. *IEEE Access*, *8*, 11315–11327.

Tal, G., Xing, Y., Wang, Y., & Sun, S. (2018). Motivations and barriers associated with the adoption of battery electric vehicles in beijing: A multinomial logit model approach.

Tian, X., Su, H., Wang, F., Zhang, K., & Zheng, Q. (2019). A electric vehicle charging station optimization model based on fully electrified forecasting method. *Engineering Letters*, *27*(4).

Torgo, L. (2014). An infra-structure for performance estimation and experimental comparison of predictive models in r. *arXiv preprint arXiv:1412.0436*.

UCLA: Statistical Consulting Group. (2021). Zero-inflated negative binomial regression | r data analysis examples [Accessed: 2023-02-20]. https://stats.oarc.ucla.edu/r/dae/zinb/

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

van Montfort, K., et al. (2016). Voorspellen van benodigde infrastructuur van publieke laadpalen voor elektrische auto's. *Tijdschrift vervoerswetenschap*, *52*.

Veeger, M. (2021). Voor iedere auto een laadpaal? dat kan ons stroomnet helemaal niet aan [Accessed: 2023-01-04]. https://www.rtlnieuws.nl/economie/bedrijven/artikel/5253093/elektrisch-netwerk-laadpaal-auto-inftrastructuur-liander

Vergis, S., & Chen, B. (2015). Comparison of plug-in electric vehicle adoption in the united states: A state by state approach. *Research in Transportation Economics*, *52*, 56–64.

Vermeij, R., & Veger, J. (2019). Prognose laadinfrastructuur [Accessed: 2022-11-27]. https://www.agendalaadinfrastructuur.nl/ondersteuning+gemeenten/documenten+en+links/documenten+in+bibliotheek/handlerdownloadfiles.ashx?idnv=1773488

Vidyattama, Y., Tanton, R., Sinclair, D., Schirmer, J., & McCaffrie, B. (2022). Estimating the impact of electric vehicle policy in the australian capital territory. *Available at SSRN 4209060*.

Vilchez, J. J. G., & Jochem, P. (2019). Simulating vehicle fleet composition: A review of system dynamics models. *Renewable and Sustainable Energy Reviews*, *115*, 109367.

Voegtlin, C., Scherer, A. G., Stahl, G. K., & Hawn, O. (2022). Grand societal challenges and responsible innovation. *Journal of Management Studies*, *59*(1), 1–28.

Wee, S., Coffman, M., & Allen, S. (2020). Ev driver characteristics: Evidence from hawaii. *Transport Policy*, *87*, 33–40.

Wesseling, J. H., Bidmon, C., & Bohnsack, R. (2020). Business model design spaces in socio-technical transitions: The case of electric driving in the netherlands. *Technological Forecasting and Social Change*, *154*, 119950.

Wolterman, B., Duurkoop, T., Zweistra, M., Hiep, E., & van Biezen, M. (2022). *Nationaal laadonderzoek 2022*. ElaadNL, de Vereniging Elektrische Rijders, Rijksdienst voor Ondernemend Nederland and Rijksuniversiteit Groningen.

Yang, Y., & Tan, Z. (2019). Investigating the influence of consumer behavior and governmental policy on the diffusion of electric vehicles in beijing, china. *Sustainability*, *11*(24), 6967.

Yao, J., Xiong, S., & Ma, X. (2020). Comparative analysis of national policies for electric vehicle uptake using econometric models. *Energies*, *13*(14), 3604.

Zambrano-Gutiérrez, J. C., Nicholson-Crotty, S., Carley, S., & Siddiki, S. (2018). The role of public policy in technology diffusion: The case of plug-in electric vehicles. *Environmental science & technology*, *52*(19), 10914–10922.

A
Appendix

A.1. Public Charging Points Municipality Level

Figure A.1 shows the number of public charging points per 1000 citizens on municipal level. It is a more detailed version of 2.2 showing the same statistic per province. The map is not complete since not all municipalities provide this data. However, it can still be seen that it is in line with 2.2. Furthermore, the relative number of charging points seems to differ per municipality.

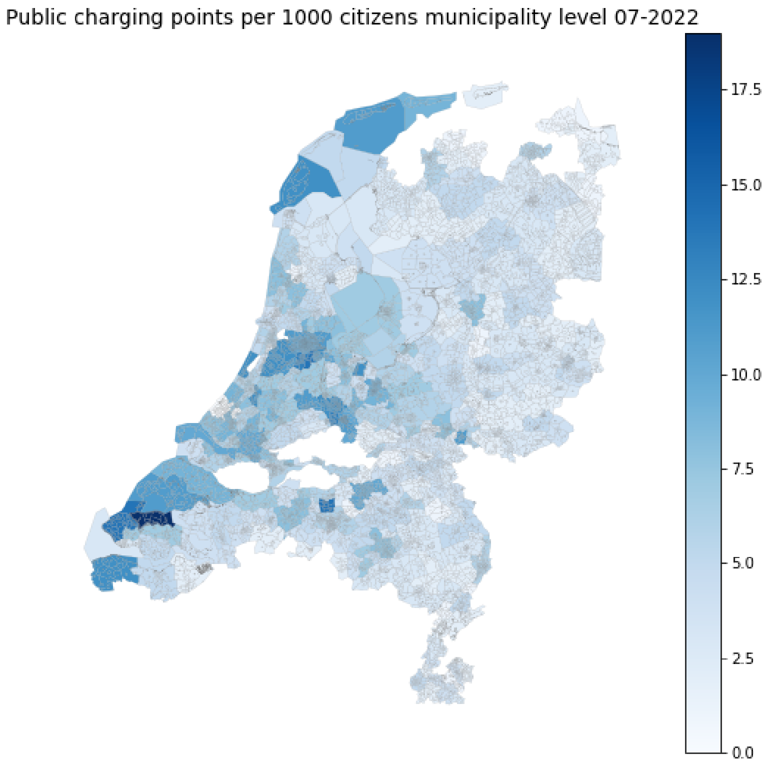

Figure A.1: Public charging points per thousand citizens, municipal level

A.2. Analysis Policy on Municipality Level

This section shows the results of the analysis of the policy of a selection of municipalities in the MRA-Elektrisch region. This includes whether they make use of preprepared preference locations and whether citizens can request the placement of public charging points.

Table A.1: Overview of the policy in selection of the MRA-Elektrisch region

Municipality	Province	Preverence Locations	On Request
Alkmaar	Noord-Holland	No	Yes
Almere	Flevoland	Yes	Yes
Amersfoort	Utrecht	Yes	Yes
De Bilt	Utrecht	No	Yes
Den-Helder	Noord-Holland	Yes	Yes
Enkhuizen	Noord-Holland	Yes	Yes
Haarlem	Noord-Holland	No	Yes
Haarlemmermeer	Noord-Hollands	Yes	No
Hilversum	Noord-Holland	No	Yes
Lelystad	Flevoland	Yes	Yes
Nieuwegein	Utrecht	No	Yes
Noordoostpolder	Flevoland	No	Yes
Stichtse Vecht	Utrecht	Yes	Yes
Zaanstad	Noord-Holland	No	Yes
Zeist	Utrecht	Yes	Yes

A.3. CHIEF.DM FACT_CHARGESESSION Transactions

This section describes the data from the *FACT_CHARGESESSION* table used for the modelling in this study. In total, 3016012 transactions were available for this analysis in 2022. 2818475 of these were found to meet the restrictions described in Section 4.3. And 2842703 transactions were found to be applicable to the occupancy rate statistic as in Section 4.4. An overview of these numbers over the year can be found in Figure A.2 and Table A.2.

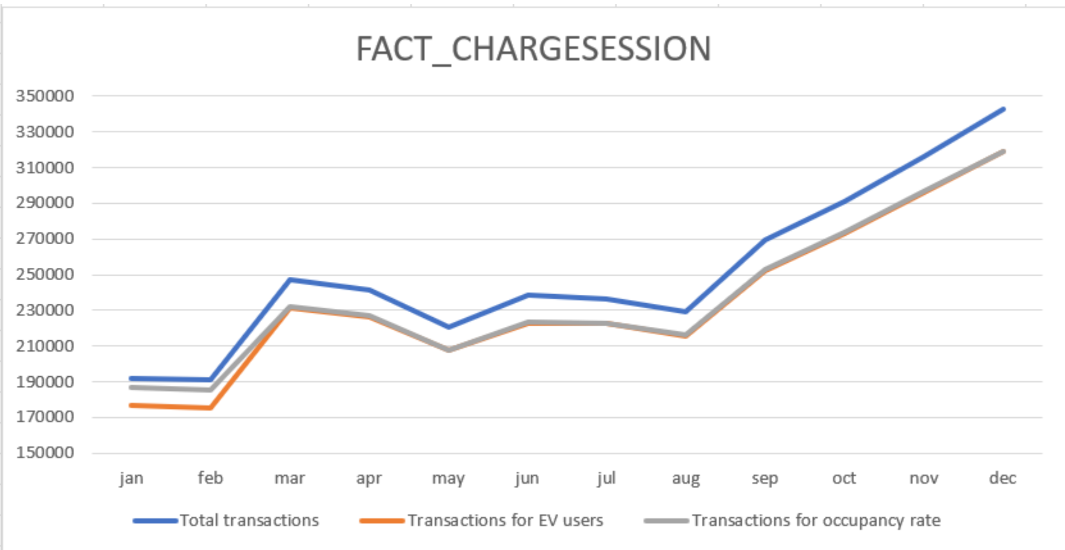

Figure A.2: Plot of transactions per month

Table A.2: Overview transactions per month

Month	Total transactions	Transactions for EV users	Transactions for occupancy rate
Jan	191749	176610	186686
Feb	191015	175178	185082
Mar	247207	231595	232040
Apr	241266	226307	226805
May	220645	207432	207854
Jun	238243	222687	223138
Jul	236643	222496	222911
Aug	229407	215816	216189
Sep	269567	252450	252824
Oct	291390	273126	273518
Nov	316076	296060	296491
Dec	342804	318718	319165

A.4. Deleted Neighbourhoods

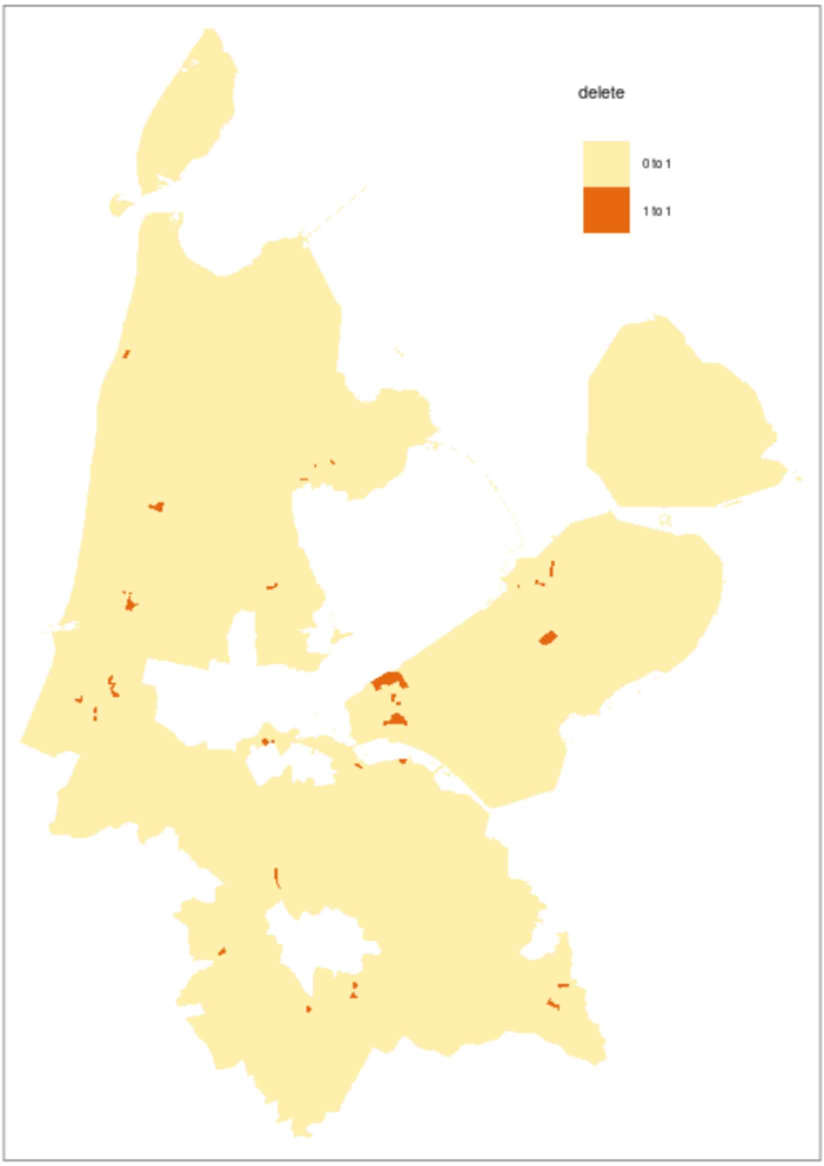

Figure A.3: Deleted neighbourhoods due to high percentage of NA values

A.5. Comparison of Samples of the Independent Variables

This section contains the results of the comparison analyses between the MRA-Elektrisch region and the rest of the Netherlands. The figures below show the distributions of the independent variables considered in this study. The legend of each figure states the p-values of the statistical test as described in Section 4.6.7 that indicates whether the samples come from the same distribution and whether the samples have the same mean value.

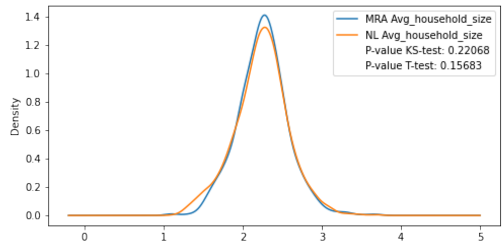

Figure A.4: Avg_household_size

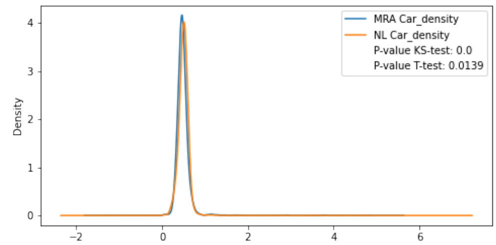

Figure A.5: Car_density

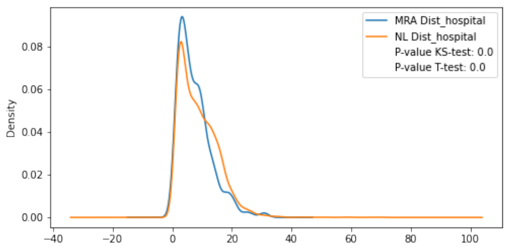

Figure A.6: Dist_hospital

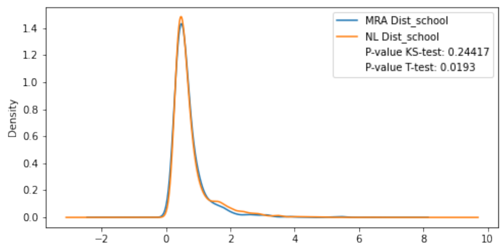

Figure A.7: Dist_school

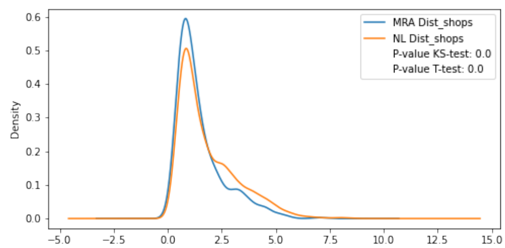

Figure A.8: Dist_shops

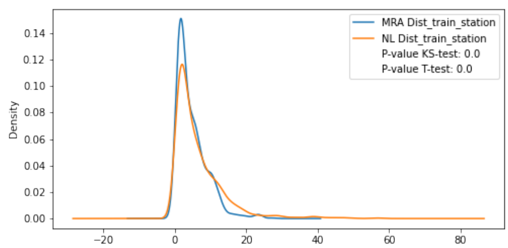

Figure A.9: Dist_train_station

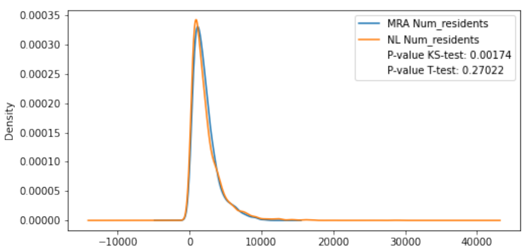

Figure A.10: Num_residents

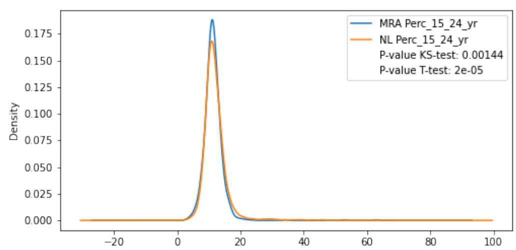

Figure A.11: Perc_15_24_yr

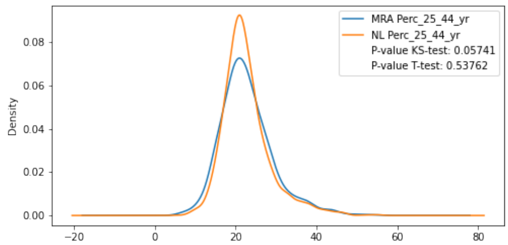

Figure A.12: Perc_25_44_yr

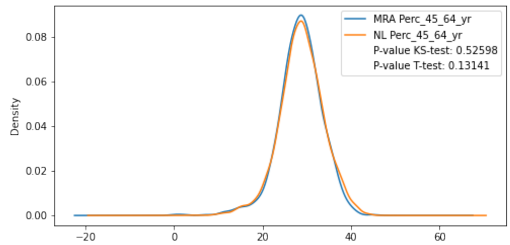

Figure A.13: Perc_45_64_yr

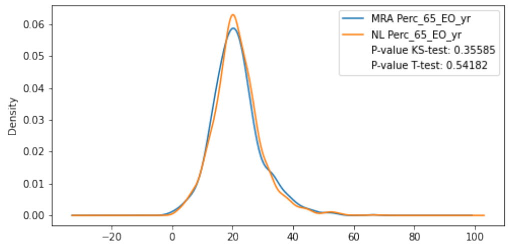

Figure A.14: Perc_65_EO_yr

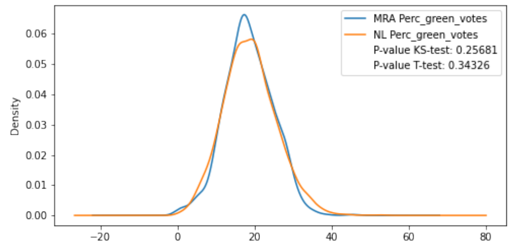

Figure A.15: Perc_green_votes

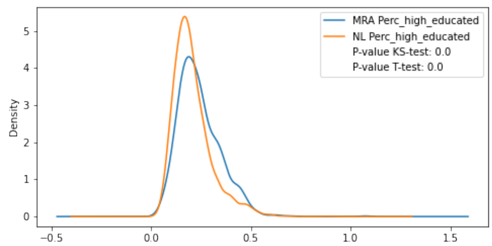

Figure A.16: Perc_high_educated

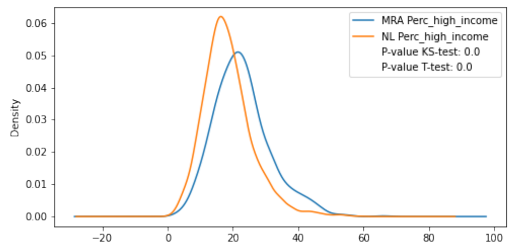

Figure A.17: Perc_high_income

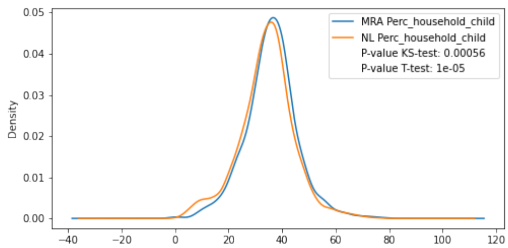

Figure A.18: Perc_household_child

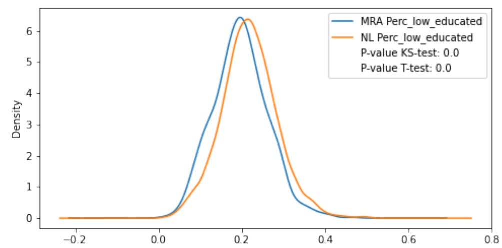

Figure A.19: Perc_low_educated

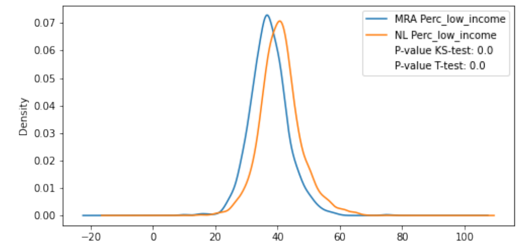

Figure A.20: Perc_low_income

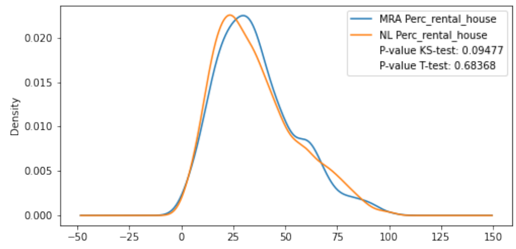

Figure A.21: Perc_rental_house

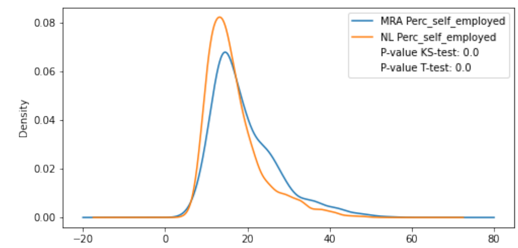

Figure A.22: Perc_self_employed

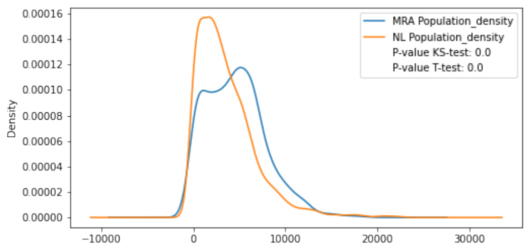

Figure A.23: Population_density

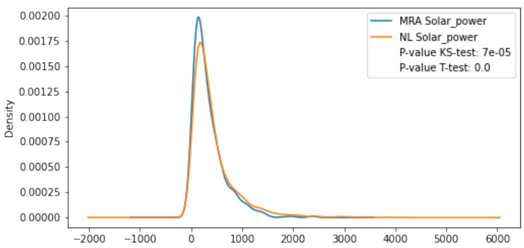

Figure A.24: Solar_power

A.6. EV Users Model

This section contains the additional results of the EV user model, as in Section 4.3. First, the results of the statistical test of a zero inflated distribution can be found in Figure A.25 showing that hypotheses of the distribution of EV users not being zero inflated cannot be rejected. Subsequently, Figure A.26 shows that the same distribution is overdispersed, which means that the variance is greater than the mean. This indicates that not Poisson regression, but negative binomial regression, should be used. The complete model description can be found in A.27. Finally, Figure A.28 shows the prediction errors on the number of EV users per neighbourhood of the EV users model.

```
Score test for zero inflation

        Chi-square = 235.79444
        df = 1
        pvalue: < 2.22e-16
```

Figure A.25: Score test for zero inflation

```
        Overdispersion test

data:  fmp
z = 9.6956, p-value < 2.2e-16
alternative hypothesis: true dispersion is greater than 1
sample estimates:
dispersion
   2.17504
```

Figure A.26: Test for overdispersion

```
Pearson residuals:
    Min      1Q  Median      3Q     Max
-1.2561 -0.7498 -0.2415  0.4727  7.2848

Count model coefficients (negbin with log link):
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)       1.368e+00  2.235e-01   6.121 9.27e-10 ***
Solar_power       9.755e-04  5.684e-05  17.160  < 2e-16 ***
Dist_school       2.441e-01  6.723e-02  -3.631 0.000283 ***
Perc_rental_house 7.205e-03  1.830e-03   3.936 8.28e-05 ***
Dist_hospital     4.781e-02  5.982e-03  -7.993 1.32e-15 ***
Perc_green_votes  3.511e-02  6.227e-03   5.639 1.71e-08 ***
Perc_low_educated 2.553e+00  4.799e-01  -5.321 1.03e-07 ***
Dist_shops        1.030e-01  3.924e-02  -2.625 0.008668 **
Perc_self_employed 1.080e-02  5.104e-03   2.116 0.034358 *
Perc_65_EO_yr     2.561e-02  3.194e-03  -8.019 1.06e-15 ***
Log(theta)        6.689e-01  7.591e-02   8.813  < 2e-16 ***

Zero-inflation modeel coefficients (binomial with logit link):
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)       2.326e+01  7.514e+00  -3.096 0.001962 **
Solar_power       1.502e-03  5.717e-04   2.628 0.008597 **
Perc_rental_house 1.180e-01  3.357e-02   3.515 0.000439 ***
Avg_household_size 5.006e+00  1.774e+00   2.821 0.004781 **
Dist_hospital     2.660e-01  1.294e-01  -2.056 0.039801 *
Perc_green_votes  2.505e-01  7.320e-02  -3.422 0.000622 ***
Dist_shops        9.091e-01  3.890e-01   2.337 0.019457 *
Perc_self_employed 7.552e-02  3.633e-02   2.079 0.037662 *
Perc_45_64_yr     1.475e-01  6.188e-02   2.383 0.017168 *
Perc_65_EO_yr     8.905e-02  3.488e-02   2.553 0.010674 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 1.9522
Number of iterations in BFGS optimization: 62
Log-likelihood: -3011 on 21 Df
```

Figure A.27: EV users model

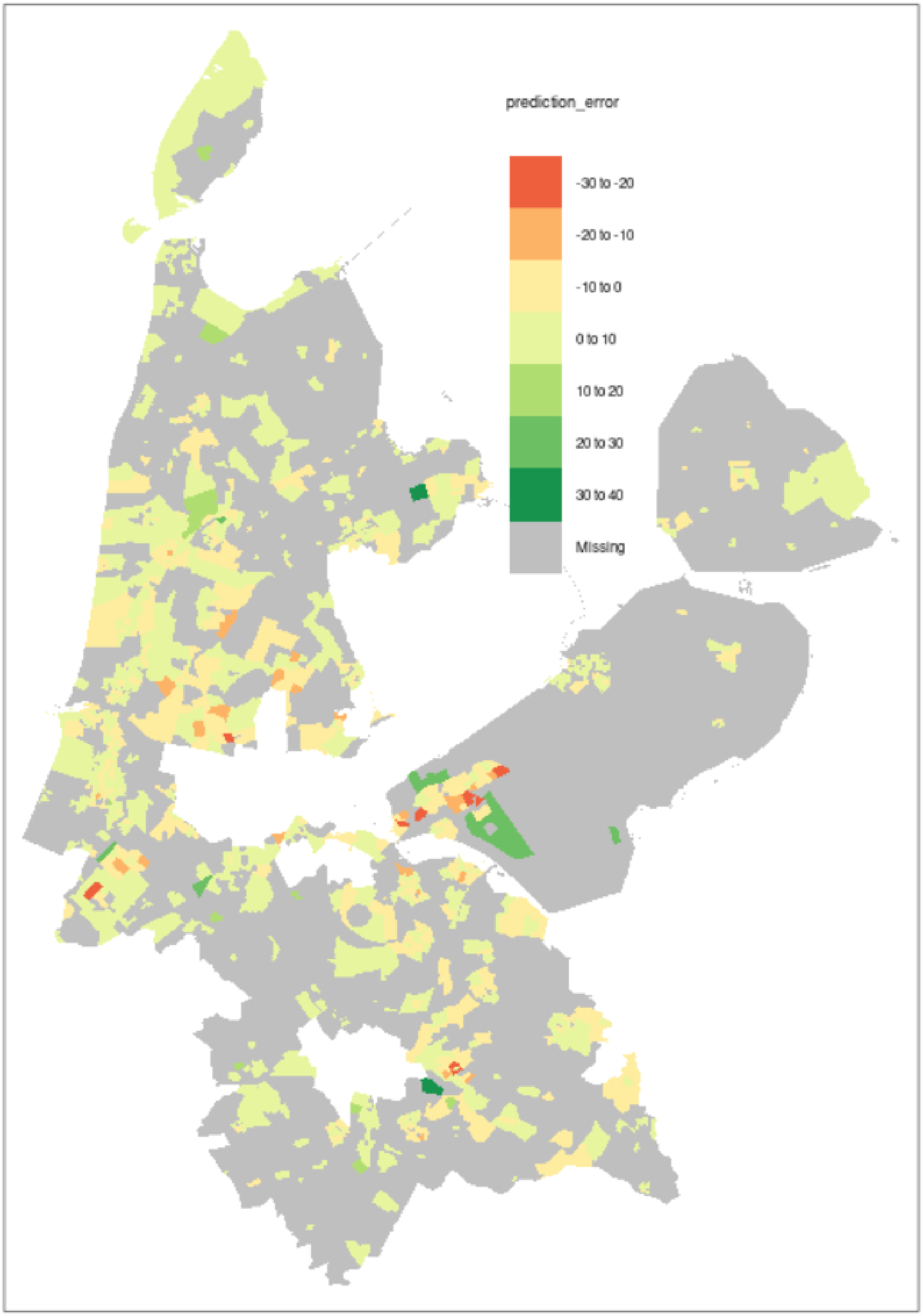

Figure A.28: Prediction error EV users model

A.7. Occupancy Rate Model

This section contains the figures and tests used to verify the assumptions of the occupancy rate model as described in Chapter 4. First, the complete description of the model can be found in Figure A.29. Next, the correlation matrix of the model variables will be shown, see Figure A.30. Subsequently, the pair plots of the dependent and independent variable are shown in Figures A.31, A.32 & A.33.

Subsequently, the ACF plot, Figure A.34 and the PACF plot, Figure A.35, are shown. Followed by the results of the statistical test for autocorrelation of the first lag, Figure A.36. Finally, the test for normality of the residuals, Figure A.37, and a plot of these residuals, Figure A.38, will be shown.

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.32727 -0.06274 -0.00624  0.06114  0.43684 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)        1.834e-01  4.932e-02   3.718 0.000209 ***
Num_residents      1.798e-05  2.818e-06   6.379 2.49e-10 ***
Solar_power       -4.699e-05  1.024e-05  -4.591 4.86e-06 ***
Perc_high_income   3.443e-03  5.587e-04   6.163 9.55e-10 ***
Has_school         3.198e-02  8.605e-03   3.717 0.000211 ***
Avg_household_size -6.263e-02  1.759e-02  -3.561 0.000384 ***
Perc_green_votes   4.068e-03  6.690e-04   6.081 1.58e-09 ***
Perc_high_educated  1.685e-01  4.476e-02   3.764 0.000175 ***
Population_density  1.112e-05  1.172e-06   9.491  < 2e-16 ***
Perc_self_employed  1.943e-03  5.455e-04   3.562 0.000382 ***
Perc_45_64_yr     -2.791e-03  5.905e-04  -4.727 2.53e-06 ***
Perc_65_EO_yr     -1.183e-03  3.771e-04  -3.138 0.001742 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.09948 on 1280 degrees of freedom
Multiple R-squared:  0.3708,	Adjusted R-squared:  0.3654 
F-statistic: 68.58 on 11 and 1280 DF,  p-value: < 2.2e-16
```

Figure A.29: Occupancy rate model

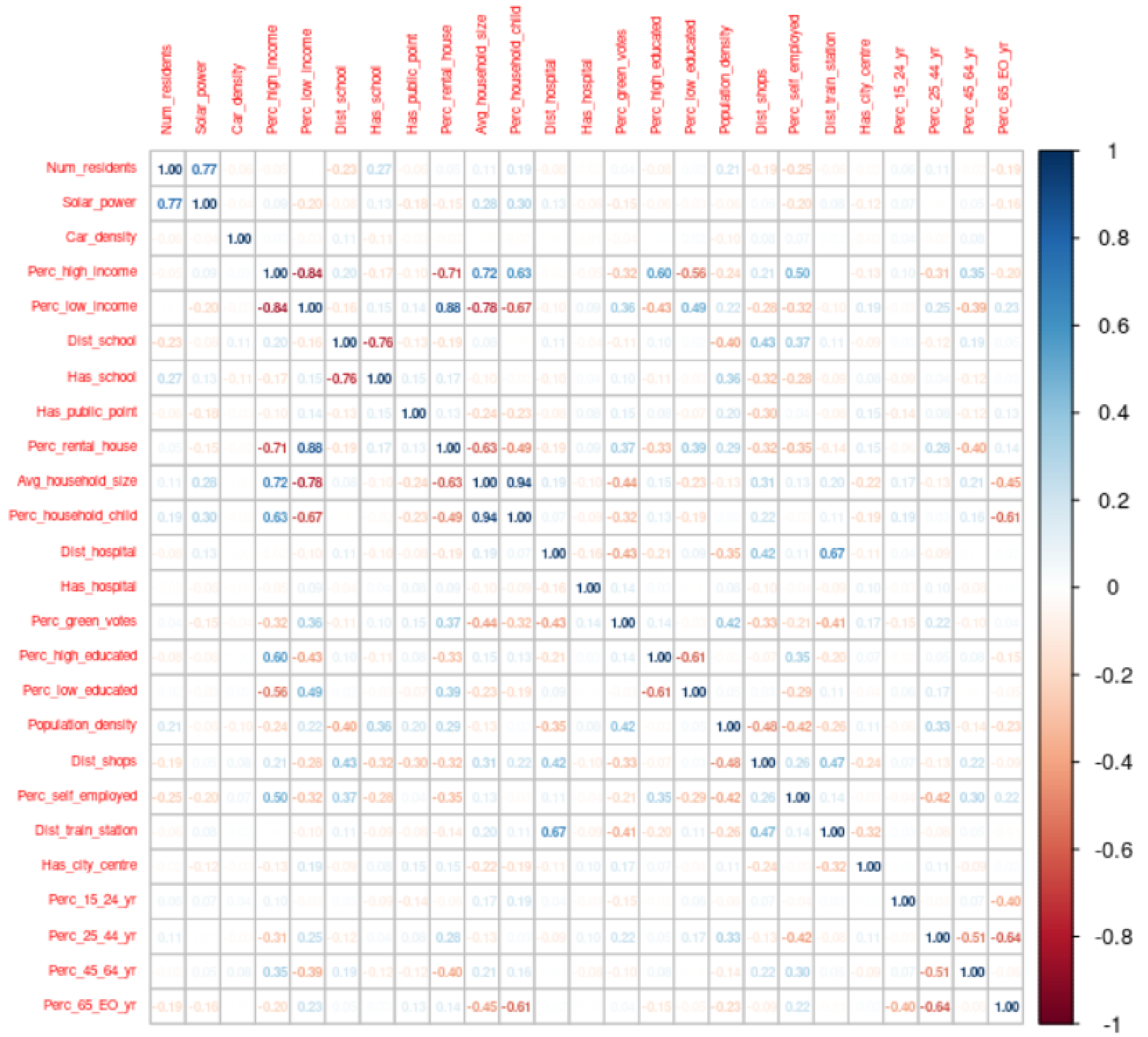

Figure A.30: Correlation matrix

A.7.1. Pairplots

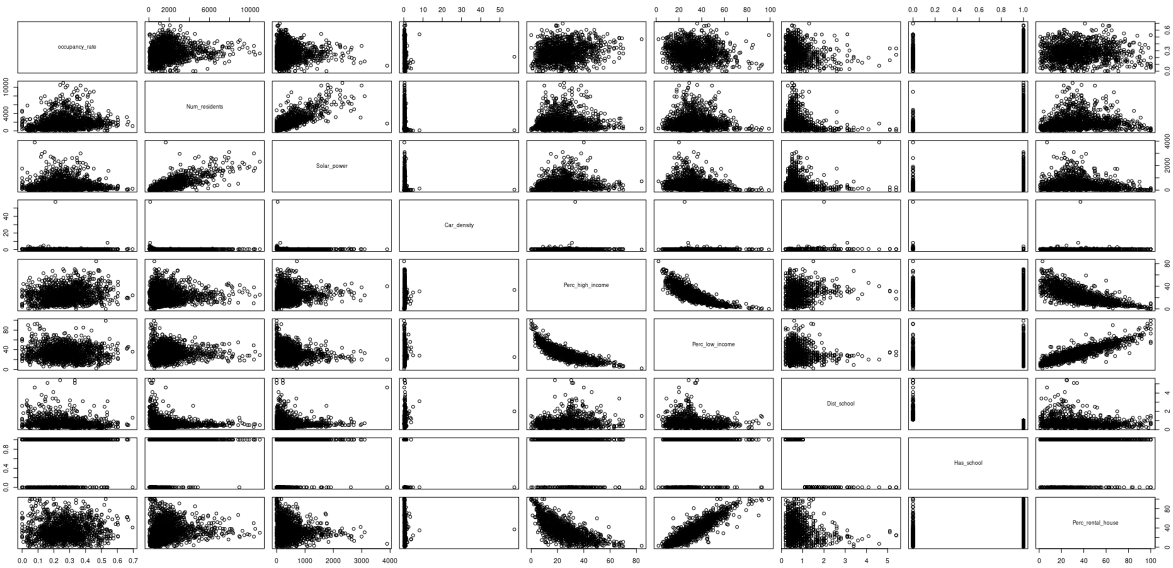

Figure A.31: Pair plots, part 1

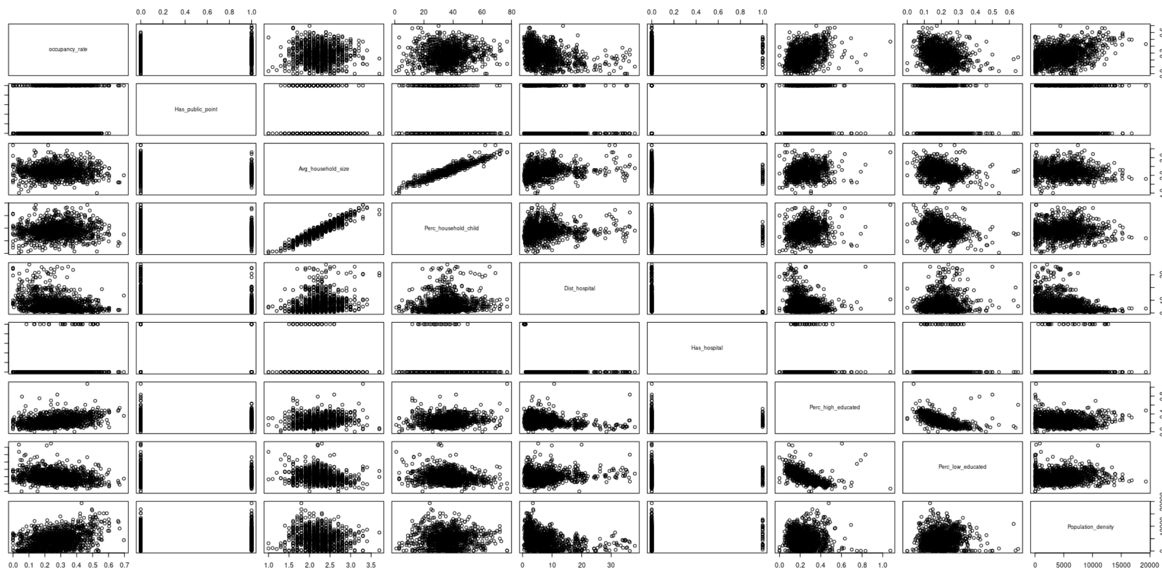

Figure A.32: Pair plots, part 2

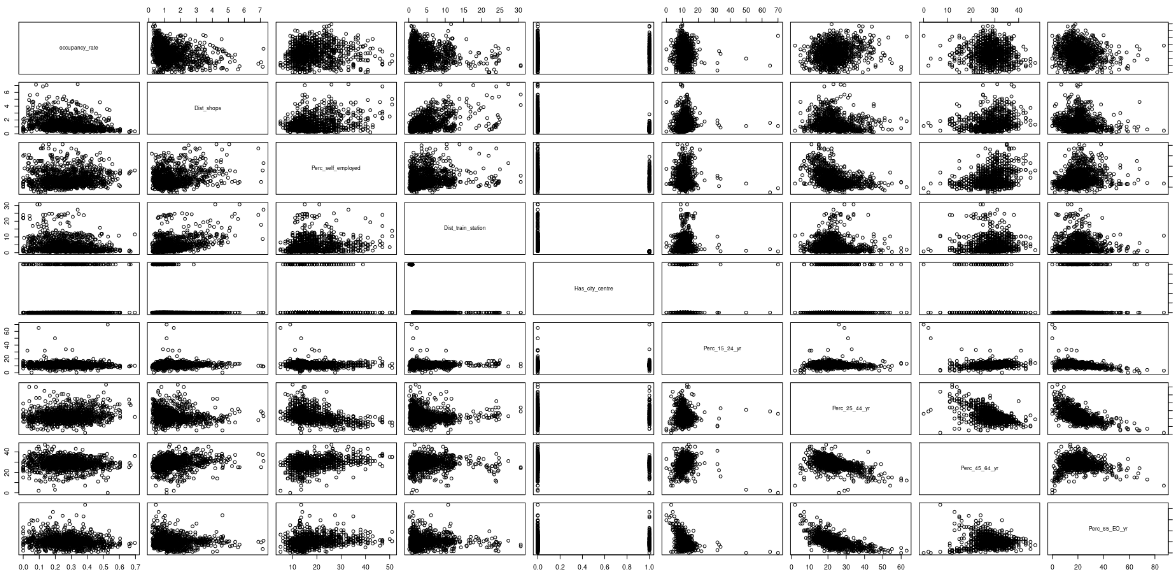

Figure A.33: Pair plots, part 3

A.7.2. Auto Correlation

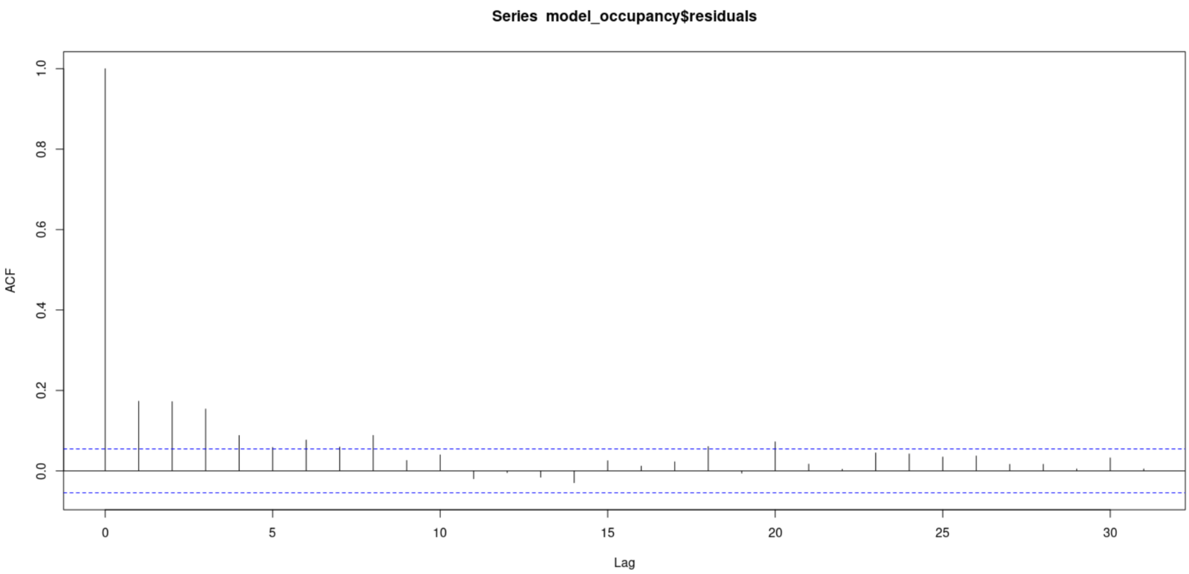

Figure A.34: ACF plot

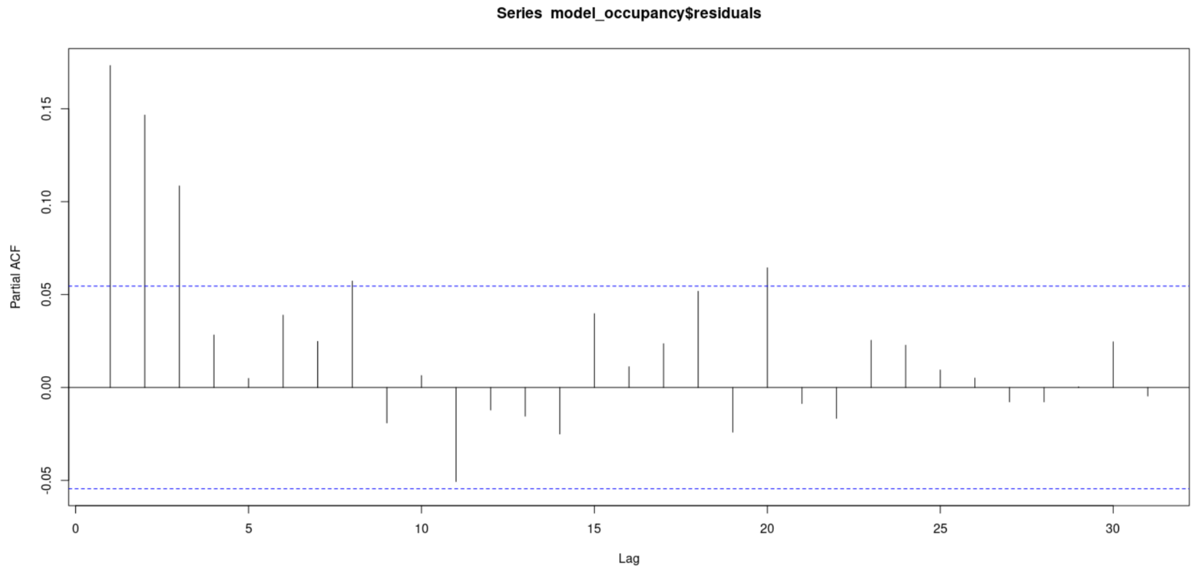

Figure A.35: PACF plot

```
        Durbin-Watson test

data:  model_occupancy
DW = 1.6508, p-value = 0.2398
alternative hypothesis: true autocorrelation is greater than 0
```

Figure A.36: Auto-correlation test

A.7.3. Residuals

Test	Statistic	pvalue
Shapiro-Wilk	0.9934	0.0000
Kolmogorov-Smirnov	0.0375	0.0529
Cramer-von Mises	351.104	0.0000
Anderson-Darling	2.0344	0.0000

Figure A.37: Test Normality Residuals

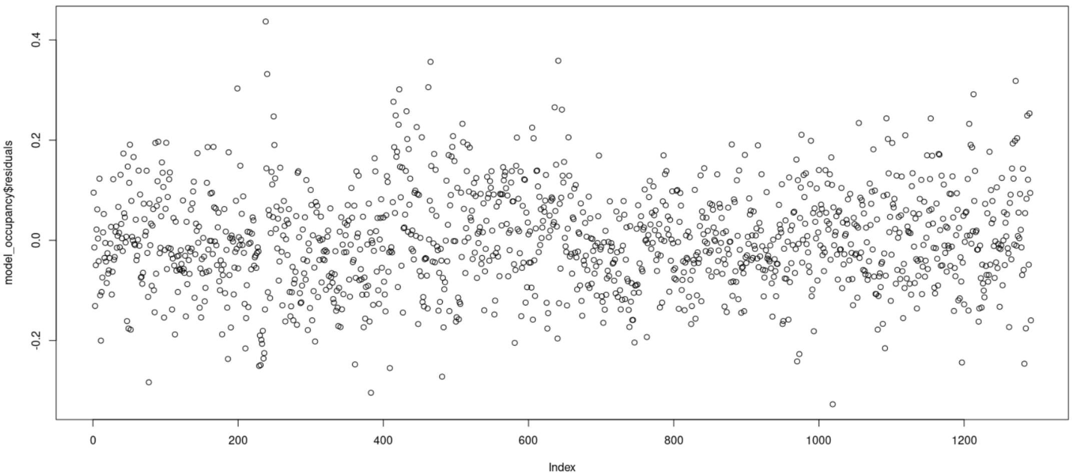

Figure A.38: Residuals plot

A.8. Code

In this section, a selection of the code used for the quantitative part of the study is listed. The code is organised by programming language. The comments in the code sections explain how the code works and how it was used. The selection of code can also be found on: https://github.com/dkoopman/master_thesis_ev_adoption.

A.8.1. R code

```r
#This script returns the following:
# neighborhoodsMRA: geodf with the polygons of all neighborhoods that can be considered
#   in the modelling of ev adoption in the mra region. the valid column presents whether the
#   neighborhood meets the requirements to be included in the modelling.
# relevant_chargers: all the relevant charging points that can be taken into consideration

#import dependencies and set correct settings
options(java.parameters=’-Xmx2g’)
library(magrittr)
library(lubridate)
library(dplyr)
library(sp)
library(sf)
library(tidyverse)
library(DBI)
library(rJava)
library(RJDBC)
library(stringr)
library(readr)
library(tmap)

#First run buurt_level_data script before this one

#Filter data based on DIM_ChargePoint table
chargingPoints <- DIM_CHARGEPOINT(username, password)
chargingPoints <- chargingPoints[chargingPoints$IsFastCharger == ’0’,] #however all charging points have value 0, is this correct?
chargingPoints <- chargingPoints[chargingPoints$LastActiveDateTime != chargingPoints$FirstActiveDateTime,] #so there is an interval
chargingPoints <- chargingPoints[chargingPoints$LastActiveDateTime >= chargingPoints$FirstActiveDateTime,] #interval should be positive
chargingPoints <- chargingPoints[chargingPoints$LastActiveDateTime >= ”2022-01-01”,] #interval should be positive

#decide the interval the charging point is ’open’ as good as that can be done
chargingPoints <- chargingPoints %>%
  mutate(OperationInterval = difftime(LastActiveDateTime,FirstActiveDateTime, units = ’weeks’))
chargingPoints <- chargingPoints[chargingPoints$OperationInterval >= 4,]
chargingPoints <- chargingPoints[chargingPoints$status == 1,]
chargingPoints <- chargingPoints[chargingPoints$NumberOfSockets >= 1,]

#remove rows without ChargePoint_skey,
chargingPoints <- chargingPoints %>% drop_na(ChargePoint_skey)
chargingPoints <- chargingPoints %>% drop_na(OperationInterval)

#merge with location table
LOOKUP_LOCATION_CHARGEPOINT_USER <- LOOKUP_LOCATION_CHARGEPOINT_USER(username,password)
chargingPoints <- left_join(chargingPoints, LOOKUP_LOCATION_CHARGEPOINT_USER)
DIM_LOCATION <- DIM_LOCATION(username, password)
chargingPoints <- left_join(chargingPoints, DIM_LOCATION)
chargingPoints <- chargingPoints %>% drop_na(PostalCode)

#Code to see the chargers per postcode
#test <- chargingPoints %>% group_by(PostalCode) %>%
#   summarise(total_count=n(),
#             .groups = ’drop’)
#rm(test)

#drop the following postcodes since they have more than 10 occurences and are manually checked and found invalid
```

```
chargingPoints <- chargingPoints[!chargingPoints$PostalCode == "3812PH",] #would have 235 chargers, are not in the area
chargingPoints <- chargingPoints[!chargingPoints$PostalCode == "3433PG",] #would have 210 chargers, are in the area but not on that postcode
chargingPoints <- chargingPoints[!chargingPoints$PostalCode == "3709JK",] #would have 42 chargers, are in the area but not on that postcode

#remove the duplicates
chargingPointsDuplicates <- chargingPoints[duplicated(chargingPoints[ , c("ChargePoint_skey")]), ]
chargingPointsNoDuplicates <- chargingPoints[!duplicated(chargingPoints[ , c("ChargePoint_skey")]), ]
relevant_chargers <- chargingPointsNoDuplicates$ChargePoint_skey
rm(chargingPointsDuplicates)

#count the number of chargers per postcode area
chargersPerPostcode <- chargingPointsNoDuplicates %>%
  group_by(PostalCode) %>%
  summarise(charging_points = n_distinct(ChargePoint_ID))

#load mapping tables
KoppelPostcodeBuurt <- read.csv(file = 'Postcode/brt2020.csv', sep = ';')
KoppelPostcode <- read.csv(file = 'Postcode/pc6-gwb2020.csv', sep = ';')

#join mapping tbales to shapefile
KoppelPostcode <- left_join(KoppelPostcode, KoppelPostcodeBuurt, by = c('Buurt2020' = 'buurtcode2020'))
KoppelPostcode <- KoppelPostcode[,c("PC6","Buurt2020","buurtnaam2020","GM2020", "GM_NAAM")]

#join the neighborhood data
chargersPerNeighborhood <- left_join(chargersPerPostcode, KoppelPostcode, by = c('PostalCode' = 'PC6'))
chargersPerNeighborhood <- chargersPerNeighborhood %>%
  group_by(Buurt2020) %>%
  summarise(charging_points = sum(charging_points))
sum(chargersPerNeighborhood$charging_points)
chargersPerNeighborhood <- chargersPerNeighborhood %>% drop_na(Buurt2020)

#read in neighborhood data
my_spdf_buurt <- read_sf(
  dsn= paste0(getwd(),"/Buurtdata/WijkBuurtkaart_2020_v3/buurt_2020_v3.shp"),
)
my_spdf_buurt <- my_spdf_buurt[,c("BU_CODE", "AANT_INW", "WONINGEN", "AANTAL_HH")]
my_spdf_buurt <- my_spdf_buurt %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
my_spdf_buurt$BU_CODE = as.numeric(as.character(my_spdf_buurt$BU_CODE))

#join with the chargersPerNeighborhood
chargersPerNeighborhood <- left_join(my_spdf_buurt, chargersPerNeighborhood, by = c('BU_CODE' = 'Buurt2020'))

#only select neighborhoods with citizens and homes
chargersPerNeighborhood$AANT_INW[chargersPerNeighborhood$AANT_INW <= 0] <- NA
chargersPerNeighborhood$WONINGEN[chargersPerNeighborhood$WONINGEN <= 0] <- NA
chargersPerNeighborhood$AANTAL_HH[chargersPerNeighborhood$AANTAL_HH <= 0] <- NA

#create column whether neighborhood can be taken into the modelling
chargersPerNeighborhood$Valid <- 1
chargersPerNeighborhood[is.na(chargersPerNeighborhood$AANT_INW), "Valid"] <- 0
chargersPerNeighborhood[is.na(chargersPerNeighborhood$WONINGEN), "Valid"] <- 0
chargersPerNeighborhood[is.na(chargersPerNeighborhood$AANTAL_HH), "Valid"] <- 0

#determing the available municipalities
unique_cities <- as.vector(LOOKUP_LOCATION_CHARGEPOINT_USER$City)
unique_cities <- unique(unique_cities)
unique_cities <- replace(unique_cities, unique_cities=="Bergen NH", "Bergen (NH.)")
unique_cities <- replace(unique_cities, unique_cities=="Ouder Amstel", "Ouder-Amstel")
unique_cities <- replace(unique_cities, unique_cities=="Stichtse vecht", "Stichtse Vecht")
unique_cities <- replace(unique_cities, unique_cities=="Ijsselstein", "IJsselstein")
unique_cities <- replace(unique_cities, unique_cities=="Edam Volendam", "Edam-Volendam")
```

```
unique_cities <- append(unique_cities, "Velsen") #for IJmuiden
unique_cities <- append(unique_cities, "Heerhugowaard") #for Dijk en Waard
unique_cities <- append(unique_cities, "Beemster") #for Middenbeemster
unique_cities <- append(unique_cities, "Langedijk") #for Middenbeemster
unique_cities <- append(unique_cities, "Edam-Volendam") #for Middenbeemster
unique_cities <- append(unique_cities, "Woudenberg") #for Middenbeemster

#join mapping tbales to shapefile
chargersPerNeighborhood <- left_join(chargersPerNeighborhood, KoppelPostcodeBuurt, by = c('BU
    _CODE' = 'buurtcode2020'))

#select only available municipalities
chargersPerNeighborhood$MRA_AREA <- 0
chargersPerNeighborhood$MRA_AREA[chargersPerNeighborhood$GM_NAAM %in% unique_cities] <- 1
chargersPerNeighborhoodMRA <- chargersPerNeighborhood[chargersPerNeighborhood$MRA_AREA == 1,]

#make plot of the chargingpoint in the MRA region
tm_shape(chargersPerNeighborhoodMRA) +
  tm_fill(col = 'charging_points') +
  tm_layout(legend.title.size = 0.5,
            legend.text.size = 0.3,
            legend.position = c("right","top"),
            legend.bg.color = "white",
            legend.bg.alpha = 0.8)

tm_shape(chargersPerNeighborhood) +
  tm_fill(col = 'MRA_AREA', legend.show = FALSE)

#create neighborhoodsMRA dataframe
neighborhoodsMRA <- chargersPerNeighborhoodMRA# %>% drop_na(charging_points)
neighborhoodsMRA[is.na(neighborhoodsMRA$charging_points), "Valid"] <- 0
neighborhoodsMRA <- neighborhoodsMRA[,c("BU_CODE", "AANT_INW", "charging_points", "Valid", "
    geometry")]

#remove variables
rm(unique_cities)
rm(KoppelPostcodeBuurt)
rm(chargingPoints)
rm(my_spdf_buurt)
rm(LOOKUP_LOCATION_CHARGEPOINT_USER)
rm(chargingPointsNoDuplicates)
```

Listing A.1: Code for the creating dataframe with suitable neighbourhoods in the MRA-Elektrisch region.

```
#import dependencies and set correct settings
options(java.parameters='-Xmx2g')
library(magrittr)
library(lubridate)
library(dplyr)
library(sp)
library(tidyverse)
library(DBI)
library(rJava)
library(RJDBC)

#function to transform the data
clear_chargingsession_rfid <- function(chargingData){
  #clear rows with NA in critical columns
  chargingData <- chargingData[!is.na(chargingData$RFID_skey),]
  chargingData <- chargingData[!is.na(chargingData$UseType),]
  chargingData <- chargingData[!is.na(chargingData$kWh),]
  chargingData <- chargingData[!is.na(chargingData$IsValid),]
  chargingData <- chargingData[!is.na(chargingData$StartConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$EndConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$ChargePoint_skey),]

  #clear rows with values that are not correct
  chargingData <- chargingData[chargingData$UseType == 'regulier',]
  chargingData <- chargingData[chargingData$RFID_skey >= 0,]
  chargingData <- chargingData[chargingData$kWh > 0,]
  chargingData <- chargingData[chargingData$IsValid == 1,]
```

```
#relevant_chargers as in Relevant_charging_points script
chargingData <- chargingData[chargingData$ChargePoint_skey %in% relevant_chargers,]

#now select only rfid’s that have less than 50 charging sessions per month
chargeSessionsPerMonth <- chargingData
chargeSessionsPerMonth <- chargeSessionsPerMonth %>%
  mutate(
    start_month = month(chargeSessionsPerMonth$StartConnectionDateTime),
    year = year(chargeSessionsPerMonth$StartConnectionDateTime)
  )

monthlySessions <-  chargeSessionsPerMonth %>% group_by(start_month, year, RFID_skey) %>%
  summarise(sessionsMonth=n(),
            .groups = ’drop’)
delete_overusage_users <- monthlySessions[monthlySessions$sessionsMonth >= 50,]$RFID_skey
chargingData <- chargingData[!chargingData$RFID_skey %in% delete_overusage_users, ]

#now get the starting time of the charge sessions
chargingData <- chargingData %>%
  mutate(
    start_hour = hour(chargingData$StartConnectionDateTime)
  )

#select only transactions that started between 16pm and 4 am
chargers_evening <- chargingData[chargingData$start_hour >= 16,]
chargers_night <- chargingData[chargingData$start_hour < 4,]
chargingData <- rbind(chargers_evening, chargers_night)

#add location data to charging data
DIM_LOCATION <- DIM_LOCATION(username, password)
charingPointsData <- left_join(chargingData, DIM_LOCATION)

#delete sessions where no postcode is known
charingPointsData <- charingPointsData[!is.na(charingPointsData$PostalCode),]
charingPointsData <- charingPointsData[!is.na(charingPointsData$Location_skey),]

#now select only the transactions of people that charge more than five times a month as EV user
#this has to be done with geo analysis since nearby points are also valid
possible_ev_users <- chargingData$RFID_skey
possible_ev_users <- unique(possible_ev_users)

#create dataframe to save ev users
Postcodes <- unique(chargersPerPostcode$PostalCode)
Homechargers <- rep(0,length(Postcodes))
df <- data.frame(Postcodes, Homechargers)
df <- na.omit(df)

#create progresbar
pb <- txtProgressBar(min=0, max=length(possible_ev_users), style=3, width=50, char="=")

for(i in seq_along(possible_ev_users)){
  #select charging sessions for that rfid
  selection <- charingPointsData[charingPointsData$RFID_skey == possible_ev_users[[i]],]
  distinct_locations <- length(unique(selection$Location_skey))
  distinct_sessions <- length(unique(selection$ChargeSession_skey))

  if(distinct_sessions > 5){
    #group by charger and get occurrences and sort descending
    chargers <- selection %>%
      group_by(Location_skey) %>%
      summarise(occurences=n(),
                .groups = ’drop’) %>%
      arrange(desc(occurences))

    if(chargers[1,’occurences’] > 5){
      #add the $Location_skey since it is a tibble
      skey <- chargers[1,”Location_skey”]$Location_skey
      location_charger = selection[selection$Location_skey == skey,]
      postal_code <- location_charger[1,”PostalCode”]
      value <- df[df$Postcodes == postal_code,”Homechargers”]
```

```r
        if(length(value) != 0 && value == 0){
          df[df$Postcodes == postal_code,"Homechargers"] <- c(possible_ev_users[[i]])
        }else{
          #add to list
          value <- append(value,possible_ev_users[[i]])
          df[df$Postcodes == postal_code,"Homechargers"] <- toString(value)
        }
      }else{
        #create a coordinate matrix
        coordinates <- subset(selection, select =  c("Latitude", "Longitude"))
        coordinates <- matrix(unlist(coordinates), ncol = 2, byrow = FALSE)

        #count the number of points that are within 200 meters of each point
        close_points <- apply(spDists(coordinates, longlat=TRUE), 2,
                              function(x) paste(length(which(x < 1.2)), collapse=', '))

        #get the maximum charging sessions within 200 meter radius
        max_charging_sessions = max(close_points)

        #if the maximum occurred charging sessions in a 200 meter radius exceeds 5 than count the observation
        if(max_charging_sessions > 5){
          #take the postcode of the charger with the highest number of chargin session for the user
          max_charger = which.max(close_points)
          max_occurence_charger <- selection$Location_skey[[max_charger]]

          location_charger = selection[selection$Location_skey == max_occurence_charger,]
          postal_code <- location_charger[1,"PostalCode"]
          value <- df[df$Postcodes == postal_code,"Homechargers"]
          if(length(value) != 0 && value == 0){
            df[df$Postcodes == postal_code,"Homechargers"] <- c(possible_ev_users[[i]])
          }else{
            #add new rfid user to list
            value <- append(value,possible_ev_users[[i]])
            df[df$Postcodes == postal_code,"Homechargers"] <- toString(value)
          }
          setTxtProgressBar(pb, i)
        }
      }
    }else{
      setTxtProgressBar(pb, i)
    }
  }
  return(df)
}

#select month to run
months <- seq(from = 1, to = 12, by = 1)
years <- seq(from = 22, to = 22, by = 1)

#code to create ev users per postcode per month and write to csv
for(year in years){
  for(month in months){
    if(length(month)<1){
      month <- paste("0",month,sep="")
    }
    date_string <- paste("20",year,"-",month,"-01",sep = "")
    start_date <- as.Date(date_string, "%Y-%m-%d")
    days_month <- days_in_month(start_date)
    date_string <- paste("20",year,"-",month,"-",days_month,sep = "")
    end_date <- as.Date(date_string, "%Y-%m-%d")
    start_date <- paste ("'", start_date, "'", sep = "")
    end_date <- paste ("'", end_date, "'", sep = "")
    print(end_date)

    data <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = start_date, endDateView = end_date)
    data_transformed <- clear_chargingsession_rfid(data)
    filename <- paste("Data_monthly_home_chargers_rfid//", month,"20", year,"_rfid.csv",sep = "")
```

```
    write.csv(data_transformed,filename, row.names = FALSE)
    print(paste('Processed file ',filename))
  }
}
```

Listing A.2: Code for the processing of charging transactions to EV users on monthly basis.

```
#occupancy rate
options(java.parameters='-Xmx2g')
library(magrittr)
library(lubridate)
library(dplyr)
library(sp)
library(tidyverse)
library(DBI)
library(rJava)
library(RJDBC)

#read in mapping table and chargepoint data
KoppelPostcode <- read.csv(file = 'Postcode/pc6-gwb2020.csv', sep = ';')
DIM_CHARGEPOINT <- DIM_CHARGEPOINT(username, password)
DIM_CHARGEPOINT <- DIM_CHARGEPOINT[, c('ChargePoint_skey', 'NumberOfSockets')]

clear_chargingsession_occupancy <- function(chargingData){
  #clear rows with NA in critical columns
  chargingData <- chargingData[!is.na(chargingData$RFID_skey),]
  chargingData <- chargingData[!is.na(chargingData$kWh),]
  chargingData <- chargingData[!is.na(chargingData$IsValid),]
  chargingData <- chargingData[!is.na(chargingData$StartConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$EndConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$ChargePoint_skey),]
  chargingData <- chargingData[!is.na(chargingData$ConnectionTimeHours),]

  #clear rows with values that are not correct
  chargingData <- chargingData[chargingData$RFID_skey >= 0,]
  chargingData <- chargingData[chargingData$kWh > 0,]
  chargingData <- chargingData[chargingData$IsValid == 1,]
  chargingData <- chargingData[chargingData$ChargePoint_skey %in% relevant_chargers,]# relevant_chargers as in Relevant_charging_points script

  # determine sum of connectiontime hours per chargingpoint
  chargingData <- chargingData[,c("ChargePoint_skey", "kWh", "ConnectionTimeHours", "Location_skey")]
  chargingData <- chargingData %>%
    group_by(ChargePoint_skey) %>%
    summarise(Sum_hours = sum(ConnectionTimeHours), Location_skey = min(Location_skey) )

  #divide by the number of connectors at the chargingpoint
  chargingData <- left_join(chargingData, DIM_CHARGEPOINT)
  chargingData$Sum_hours <- chargingData$Sum_hours / chargingData$NumberOfSockets

  #get the location data and map by postcode to neighborhood
  DIM_LOCATION <- DIM_LOCATION(username, password)
  chargingData <- left_join(chargingData, DIM_LOCATION)
  chargingData <- chargingData[,c("ChargePoint_skey", "Sum_hours", "Location_skey", "PostalCode")]
  chargingData <- left_join(chargingData, KoppelPostcode, by = c("PostalCode" = "PC6"))

  #take the average of the chargingstation per neighborhood
  chargingData <- chargingData %>%
    group_by(Buurt2020) %>%
    summarise(avg_hours = mean(Sum_hours))
  return(chargingData)
}

months <- seq(from = 1, to = 12, by = 1)
years <- seq(from = 22, to = 22, by = 1)

#code to process all transaction data per month and save csv file with occupancy hours
for(year in years){
  for(month in months){
```

```
    if(length(month)<1){
      month <- paste(”0”,month,sep=””)
    }
    date_string <- paste(”20”,year,”-”,month,”-01”,sep = ””)
    start_date <- as.Date(date_string, ”%Y-%m-%d”)
    days_month <- days_in_month(start_date)
    date_string <- paste(”20”,year,”-”,month,”-”,days_month,sep = ””)
    end_date <- as.Date(date_string, ”%Y-%m-%d”)
    start_date <- paste (”’”, start_date, ”’”, sep = ””)
    end_date <- paste (”’”, end_date, ”’”, sep = ””)
    print(end_date)

    data <- FACT_CHARGESESSION(username,password,country = ”’NLD’” ,
                               startDateView = start_date, endDateView = end_date)
    data_transformed <- clear_chargingsession_occupancy(data)
    filename <- paste(”Data_monthly_home_chargers_occupancy//”, month,”20”, year,”.csv”,sep = ””)
    write.csv(data_transformed,filename, row.names = FALSE)
    print(paste(’Processed file ’,filename))
  }
}
```

Listing A.3: Code for the processing of charging transactions in average usage per month.

```
#This script process the EV adoption kpi’s as produced by their corresponding scripts to
#and aggregates them for the whole year of 2022. Thereafter, it adds them to the
#neighborhoodsMRA dataframe so that they can be used for modelling and plotting.

#import dependencies
library(dplyr)
library(stringr)
library(readr)

################################################################################
#occupancy rate

#combine the data of all months for the occupancy rates, based on the csv files written
#by the script occupancy_rate_kpi.R
l <- list()
l2 <- data.frame()
months <- seq(from = 1, to = 12, by = 1)
for(month in months){
  df <- read_csv(paste(”Data_monthly_home_chargers_occupancy/”, month, ”2022.csv”, sep = ””))
  l2 <- rbind(l2, df)
}

#take the mean values over the months
occupancy_rates <- l2 %>%
  group_by(Buurt2020) %>%
  summarise(occupancy_rate=mean(avg_hours),
            .groups = ’drop’)

#devide by the average hours in a month
occupancy_rates$occupancy_rate <- occupancy_rates$occupancy_rate/730.48

#add to neighborhoodsMRA df, as made in relevant_charging_points.R
neighborhoodsMRA <- left_join(neighborhoodsMRA, occupancy_rates, by = c(”BU_CODE” = ”Buurt2020”))
neighborhoodsMRA[is.na(neighborhoodsMRA$occupancy_rate) & neighborhoodsMRA$Valid == 1, ”occupancy_rate”] <- 0
neighborhoodsMRA[neighborhoodsMRA$Valid == 0, ”occupancy_rate”] <- NA
rm(occupancy_rates)

#create statistics and histogram
occupancy_rate <- neighborhoodsMRA$occupancy_rate
hist(occupancy_rate)
mean(neighborhoodsMRA$occupancy_rate, na.rm = TRUE)
sd(neighborhoodsMRA$occupancy_rate, na.rm = TRUE)

#make plot of MRA region with occupancy rates
tm_shape(neighborhoodsMRA) +
```

```
  tm_fill(col = 'occupancy_rate', breaks = c(0,0.2,0.4,0.6,0.8,1)) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.45,
            legend.position = c("right","top"),
            legend.bg.color = "white",
            legend.bg.alpha = 0)


################################################################################
#ev users

#combine the data of all months for the ev users, based on the csv files written
#by the script ev_users.R
l <- list()
l2 <- data.frame()
months <- seq(from = 1, to = 12, by = 1)
for(month in months){
  df <- read_csv(paste("Data_monthly_home_chargers_rfid/", month, "2022_rfid.csv", sep = ""))
  avector <- as.vector(df['Homechargers'])
  l <- append(l,avector)
  l2 <- rbind(l2, df)
}
l2 <- separate_rows(l2, Homechargers, sep = "\\, ")

#seperate the values in one row to multiple rows
vec <- Reduce(c,l)
vec <- unlist(strsplit(vec, "\\, "))
v1 <- table(vec)
rfids <- as.integer(names(v1)[v1 >= 4])
#delete first element since this is the 0 occurrences
rfids <- rfids[-1]

#create new dataframe with ev users per postocde
ev_users_df <- data.frame()
for(rfid in rfids){
  selection <- l2[l2$Homechargers == rfid,]
  unique_values <- length(unique(selection$Postcodes))
  if(unique_values == 1){
    ev_users_df <- rbind(ev_users_df, selection[1,])
  }else{
    ev_users_df <- rbind(ev_users_df, selection)
  }
}

#add neighborhood data to the ev_users_df
ev_users_df <- left_join(ev_users_df, KoppelPostcode, by = c('Postcodes'= 'PC6'))
drop <- c('buurtnaam2020','GM2020', 'GM_NAAM')
ev_users_df = ev_users_df[,!(names(ev_users_df) %in% drop)]

#group per neighborhood and count the distinct rfid's
home_chargers_per_buurt <- ev_users_df%>%
  group_by(Buurt2020) %>%
  summarise(ev_users=n_distinct(Homechargers),
            .groups = 'drop')
home_chargers_per_buurt$ev_users[is.na(home_chargers_per_buurt$ev_users)] <- 0

#add the data to the neighborhoodsMRA dataframe and replace na values by 0
neighborhoodsMRA <- left_join(neighborhoodsMRA, home_chargers_per_buurt, by = c("BU_CODE" = "
    Buurt2020"))
neighborhoodsMRA[is.na(neighborhoodsMRA$ev_users) & neighborhoodsMRA$Valid == 1, "ev_users"]
    <- 0
rm(home_chargers_per_buurt)

#give the statistics and make a histogram of the distribution
ev_users <- neighborhoodsMRA$ev_users
hist(ev_users, breaks = seq(from=0, to=42, by=1))
max(ev_users, na.rm = TRUE)
mean(ev_users, na.rm = TRUE)
sd(ev_users, na.rm = TRUE)
rm(ev_users)
```

```
#make a plot of the number of ev users for the MRA region
tm_shape(neighborhoodsMRA) +
  tm_fill(col = 'ev_users', breaks = c(0,1,5,10,15,20,45)) +
  tm_layout(legend.title.size = 1.1,
            legend.text.size = 0.5,
            legend.width = 1,
            legend.position = c("right","top"),
            legend.bg.color = "white",
            legend.bg.alpha = 0)

#remove variables
rm(l)
rm(l2)
rm(avector)
rm(selection)
```

Listing A.4: Code for the processing and aggregating of the dependent variables into the modelling dataframe.

```
#prepare solar panel data
#read in data from csv
df_solar_panels <- read.csv("Upload_data_sets/Zonnestroom__wijken_en_buurten__2019_14122022_103834.csv", sep =";")
#consider only neighborhoods
df_solar_panels <- df_solar_panels[grep("BU", df_solar_panels$Regioaanduiding.Codering..code.), ]
#select relevant columns and change names
df_solar_panels <- df_solar_panels[,c("Regioaanduiding.Codering..code.", "Opgesteld.vermogen.van.zonnepanelen..kW.")]
names(df_solar_panels)[names(df_solar_panels) == 'Regioaanduiding.Codering..code.'] <- 'BU_CODE'
names(df_solar_panels)[names(df_solar_panels) == 'Opgesteld.vermogen.van.zonnepanelen..kW.'] <- 'Solar_power_kw'
#set all NaN values to 0
df_solar_panels[is.na(df_solar_panels)] <- 0
#strip the first parth of the Buurt code so it matches the other data
df_solar_panels <- df_solar_panels %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
df_solar_panels$BU_CODE = as.numeric(as.character(df_solar_panels$BU_CODE))
```

Listing A.5: Code for the processing of solar panel data.

```
#This file processes and prepares the independent variables and writes them to a csv file

#import dependencies
library("writexl")

#load the buurt shape file
my_spdf_buurt <- read_sf(
  dsn= paste0(getwd(),"/Buurtdata/WijkBuurtkaart_2020_v3/buurt_2020_v3.shp"),
)

#green_votes
df_green_votes <- read.csv("Upload_data_sets/greenVotes.csv")
df_green_votes <- na.omit(df_green_votes)
df_green_votes <- df_green_votes %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
df_green_votes <- df_green_votes[,c('BU_CODE', 'Green_votes_perc')]
names(df_green_votes)[names(df_green_votes) == 'Green_votes_perc'] <- 'Perc_green_votes'


#city_centre
df_city_centre <- my_spdf_buurt[,c("BU_CODE","AF_TREINST")]
df_city_centre[df_city_centre$AF_TREINST == -99999999,2] <- NA
df_city_centre$Has_city_centre <- as.numeric(cut(df_city_centre$AF_TREINST, 65))
df_city_centre[is.na(df_city_centre$Has_city_centre),'Has_city_centre'] <- 0
df_city_centre[df_city_centre$Has_city_centre > 1,'Has_city_centre'] <- 0
df_city_centre <- df_city_centre %>% st_drop_geometry()
df_city_centre <- df_city_centre %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
```

```
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
names(df_city_centre)[names(df_city_centre) == ’AF_TREINST’] <- ’Dist_train_station’

#income
df_income <- my_spdf_buurt[,c(”BU_CODE”,”P_HOOGINKH”, ”P_LAAGINKH”)]
df_income[df_income$P_HOOGINKH == -99999999,2] <- NA
df_income[df_income$P_LAAGINKH == -99999999,3] <- NA
df_income <- df_income %>%
  mutate_at(”BU_CODE”, str_replace, ”BU”, ””) %>%
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
df_income <- df_income %>% st_drop_geometry()
names(df_income)[names(df_income) == ’P_HOOGINKH’] <- ’Perc_high_income’
names(df_income)[names(df_income) == ’P_LAAGINKH’] <- ’Perc_low_income’

#education
df_education <- my_spdf_buurt[,c(”BU_CODE”, ”A_OPL_HG”, ”A_OPL_MD”, ”A_OPL_LG”, ”AANT_INW”)]
df_education[df_education$A_OPL_HG == -99999999,2] <- NA
df_education[df_education$A_OPL_MD == -99999999,3] <- NA
df_education[df_education$A_OPL_LG == -99999999,4] <- NA
df_education[df_education$AANT_INW == -99999999,5] <- NA
df_education$P_OPL_HG <- df_education$A_OPL_HG/df_education$AANT_INW
df_education$P_OPL_MD <- df_education$A_OPL_MD/df_education$AANT_INW
df_education$P_OPL_LG <- df_education$A_OPL_LG/df_education$AANT_INW
df_education <- df_education[,c(’BU_CODE’,’P_OPL_HG’, ’P_OPL_LG’)]
df_education <- df_education %>%
  mutate_at(”BU_CODE”, str_replace, ”BU”, ””) %>%
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
df_education <- df_education %>% st_drop_geometry()
names(df_education)[names(df_education) == ’P_OPL_HG’] <- ’Perc_high_educated’
names(df_education)[names(df_education) == ’P_OPL_LG’] <- ’Perc_low_educated’


#age
df_age <- my_spdf_buurt[,c(”BU_CODE”, ”P_00_14_JR”,”P_15_24_JR”,”P_25_44_JR”, ”P_45_64_JR”, ”P_65_EO_JR”)]
df_age[df_age$P_00_14_JR == -99999999,2] <- NA
df_age[df_age$P_15_24_JR == -99999999,3] <- NA
df_age[df_age$P_25_44_JR == -99999999,4] <- NA
df_age[df_age$P_45_64_JR == -99999999,5] <- NA
df_age[df_age$P_65_EO_JR == -99999999,6] <- NA
df_age <- df_age[,c(’BU_CODE’, ’P_15_24_JR’, ’P_25_44_JR’, ’P_45_64_JR’, ’P_65_EO_JR’)]
df_age <- df_age %>% st_drop_geometry()
df_age <- df_age %>%
  mutate_at(”BU_CODE”, str_replace, ”BU”, ””) %>%
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
names(df_age)[names(df_age) == ’P_15_24_JR’] <- ’Perc_15_24_yr’
names(df_age)[names(df_age) == ’P_25_44_JR’] <- ’Perc_25_44_yr’
names(df_age)[names(df_age) == ’P_45_64_JR’] <- ’Perc_45_64_yr’
names(df_age)[names(df_age) == ’P_65_EO_JR’] <- ’Perc_65_EO_yr’

#self employed
df_self_employed <- my_spdf_buurt[,c(”BU_CODE”, ”P_ARB_ZS”)]
df_self_employed[df_self_employed$P_ARB_ZS == -99999999,2] <- NA
df_self_employed <- df_self_employed %>% st_drop_geometry()
df_self_employed <- df_self_employed %>%
  mutate_at(”BU_CODE”, str_replace, ”BU”, ””) %>%
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
names(df_self_employed)[names(df_self_employed) == ’P_ARB_ZS’] <- ’Perc_self_employed’

#housing type
df_house_type <- my_spdf_buurt[,c(”BU_CODE”, ”P_MGEZW”)]
df_house_type[df_house_type$P_MGEZW == -99999999,3] <- NA
df_house_type <- df_house_type[,c(”BU_CODE”,”P_MGEZW”)]
df_house_type <- df_house_type %>% st_drop_geometry()
df_house_type <- df_house_type %>%
  mutate_at(”BU_CODE”, str_replace, ”BU”, ””) %>%
  mutate_at(”BU_CODE”, str_remove, ”^0+”)
names(df_house_type)[names(df_house_type) == ’P_MGEZW’] <- ’Perc_multi_house’

#housing ownership
df_house_own <- my_spdf_buurt[,c(”BU_CODE”, ”P_HUURWON”)]
```

```
df_house_own[df_house_own$P_HUURWON == -99999999,2] <- NA
df_house_own <- df_house_own[,c("BU_CODE","P_HUURWON")]
df_house_own <- df_house_own %>% st_drop_geometry()
df_house_own <- df_house_own %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
names(df_house_own)[names(df_house_own) == 'P_HUURWON'] <- 'Perc_rental_house'

#density
df_density <- my_spdf_buurt[,c("BU_CODE", "BEV_DICHTH")]
df_density[df_density$BEV_DICHTH == -99999999,2] <- NA
df_density <- df_density %>% st_drop_geometry()
df_density <- df_density %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
names(df_density)[names(df_density) == 'BEV_DICHTH'] <- 'Population_density'


#household_size
df_household_size <- my_spdf_buurt[,c("BU_CODE", "GEM_HH_GR", "P_HH_M_K")]
df_household_size[df_household_size$GEM_HH_GR == -99999999,2] <- NA
df_household_size[df_household_size$P_HH_M_K == -99999999,3] <- NA
df_household_size <- df_household_size %>% st_drop_geometry()
df_household_size <- df_household_size %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
names(df_household_size)[names(df_household_size) == 'GEM_HH_GR'] <- 'Avg_household_size'
names(df_household_size)[names(df_household_size) == 'P_HH_M_K'] <- 'Perc_household_child'

#shops
df_shops <- my_spdf_buurt[,c("BU_CODE", "AF_APOTH", "AF_SUPERM", "AF_WARENH", "AF_RESTAU")]
df_shops[df_shops$AF_APOTH == -99999999,2] <- NA
df_shops[df_shops$AF_SUPERM == -99999999,3] <- NA
df_shops[df_shops$AF_WARENH == -99999999,4] <- NA
df_shops[df_shops$AF_RESTAU == -99999999,5] <- NA
df_shops$shop_distance <- (df_shops$AF_APOTH + df_shops$AF_SUPERM + df_shops$AF_WARENH + df_
    shops$AF_RESTAU)/4
df_shops <- df_shops %>% st_drop_geometry()
df_shops <- df_shops %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
df_shops <- df_shops[,c('BU_CODE', 'shop_distance')]
names(df_shops)[names(df_shops) == 'shop_distance'] <- 'Dist_shops'

#public points
df_public_point <- my_spdf_buurt[,c("BU_CODE", "AF_ATTRAC", "AF_MUSEUM", "AF_ZWEMB")]
df_public_point[df_public_point$AF_ATTRAC == -99999999,2] <- NA
df_public_point[df_public_point$AF_MUSEUM == -99999999,3] <- NA
df_public_point[df_public_point$AF_ZWEMB == -99999999,4] <- NA
df_public_point <- transform(df_public_point, public_distance = pmin(AF_ATTRAC, AF_MUSEUM, AF
    _ZWEMB))
df_public_point$has_public_point <- as.numeric(cut(df_public_point$public_distance, 27))
df_public_point[is.na(df_public_point$has_public_point),'has_public_point'] <- 0
df_public_point[df_public_point$has_public_point >1, 'has_public_point'] <- 0
df_public_point <- df_public_point %>% st_drop_geometry()
df_public_point <- df_public_point[,c('BU_CODE', 'has_public_point')]
df_public_point <- df_public_point %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
names(df_public_point)[names(df_public_point) == 'has_public_point'] <- 'Has_public_point'

#hospitals
df_hospital <- my_spdf_buurt[,c("BU_CODE", "AF_ZIEK_E")]
df_hospital[df_hospital$AF_ZIEK_E == -99999999,2] <- NA
df_hospital$has_hospital <- as.numeric(cut(df_hospital$AF_ZIEK_E, 80))
df_hospital[is.na(df_hospital$AF_ZIEK_E),'has_hospital'] <- 0
df_hospital[df_hospital$has_hospital >1, 'has_hospital'] <- 0
df_hospital <- df_hospital %>% st_drop_geometry()
df_hospital <- df_hospital %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
```

```r
names(df_hospital)[names(df_hospital) == 'has_hospital'] <- 'Has_hospital'
names(df_hospital)[names(df_hospital) == 'AF_ZIEK_E'] <- 'Dist_hospital'

#schools
df_school <- my_spdf_buurt[,c("BU_CODE", "AF_ONDBAS")]
df_school[df_school$AF_ONDBAS == -99999999,2] <- NA
df_school$has_school <- as.numeric(cut(df_school$AF_ONDBAS, 10))
df_school[is.na(df_school$AF_ONDBAS),'has_school'] <- 0
df_school[df_school$has_school >1, 'has_school'] <- 0
df_school <- df_school %>% st_drop_geometry()
df_school <- df_school %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
names(df_school)[names(df_school) == 'AF_ONDBAS'] <- 'Dist_school'
names(df_school)[names(df_school) == 'has_school'] <- 'Has_school'

#cars density
df_cars <- read.csv("Upload_data_sets/df_cars.csv")
df_cars <- df_cars[,c('BU_CODE', 'G_PAU_KM')]
temp <- my_spdf_buurt[,c("BU_CODE", "AANT_INW")]
df_cars <- left_join(df_cars, temp)
df_cars <- df_cars %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
df_cars$car_density <- df_cars$G_PAU_KM/df_cars$AANT_INW
df_cars <- df_cars[,c('BU_CODE','car_density')]
names(df_cars)[names(df_cars) == 'car_density'] <- 'Car_density'

#prepare solar panel data
#read in data from csv
df_solar_panels <- read.csv("Upload_data_sets/Zonnestroom.csv", sep =";")
#consider only neighborhoods
df_solar_panels <- df_solar_panels[grep("BU", df_solar_panels$WijkenEnBuurten), ]
#select relevant columns and change names
df_solar_panels <- df_solar_panels[,c("WijkenEnBuurten", "OpgesteldVermogenVanZonnepanelen_6"
    )]
names(df_solar_panels)[names(df_solar_panels) == 'WijkenEnBuurten'] <- 'BU_CODE'
names(df_solar_panels)[names(df_solar_panels) == 'OpgesteldVermogenVanZonnepanelen_6'] <- '
    Solar_power_kw'
#set all NaN values to 0
df_solar_panels[is.na(df_solar_panels)] <- 0
#strip the first parth of the Buurt code so it matches the other data
df_solar_panels <- df_solar_panels %>%
  mutate_at("BU_CODE", str_replace, "BU", "") %>%
  mutate_at("BU_CODE", str_remove, "^0+")
df_solar_panels$BU_CODE = as.character(df_solar_panels$BU_CODE)
names(df_solar_panels)[names(df_solar_panels) == 'Solar_power_kw'] <- 'Solar_power'

#construct dataframe
basis_df <- neighborhoodsMRA[, c('BU_CODE')]
basis_df$BU_CODE = as.character(basis_df$BU_CODE)

#add all the dataframes together
df_independent <- purrr::reduce(list(basis_df, df_solar_panels, df_cars, df_income, df_school
    , df_public_point,
                                     df_house_own, df_house_type, df_household_size, df_
    hospital,
                                     df_green_votes, df_education, df_density, df_shops,
                                     df_self_employed, df_city_centre, df_age), dplyr::left_
    join, by = 'BU_CODE')

#drop the geometry and write to file
df_independent <- df_independent %>% st_drop_geometry()
write_xlsx(df_independent,'independent_variables.xlsx')
```

Listing A.6: Code for the processing and preparation of the independent variables and to write them to a csv file.

```r
#import dependencies
library(performanceEstimation)
library(corrplot)
library(readxl)
```

```r
library(AER)
library(pscl)
library(mpath)
library("olsrr")
library(boot)

#construct modelling dataframe
independent_variables <- read_excel("independent_variables.xlsx")
neighborhoodsMRA$BU_CODE = as.character(neighborhoodsMRA$BU_CODE)
total_df <- left_join(neighborhoodsMRA, independent_variables, by = 'BU_CODE')

#make plot of which neighborhoods to delete, is done later in the code for total_df
total_df_delete_plot <- total_df
total_df_delete_plot$delete <- 0
total_df_delete_plot[(rowSums(is.na(total_df_delete_plot)) > ncol(total_df_delete_plot)*.25)
    &
                       (total_df_delete_plot$Valid == 1),'delete'] <- 1

tm_shape(total_df_delete_plot) +
  tm_fill(col = 'delete', breaks = c(0,1,1)) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.5,
            legend.position = c("right","top"),
            legend.bg.color = "white",
            legend.bg.alpha = 0)


#select only valid neighborhoods and delete columns
total_df <- total_df[total_df$Valid == 1,]
total_df <- total_df[,!names(total_df) %in% c("Valid")]

#cast columns
total_df$car_density <- as.numeric(total_df$car_density)
total_df$charging_points <- as.numeric(total_df$charging_points)

#delete rows with more than 25% na values
delete_rows <- total_df[rowSums(is.na(total_df)) > ncol(total_df)*.25,]
total_df <- total_df[!rowSums(is.na(total_df)) > ncol(total_df)*.25,]

#get new percentage of NA values
colMeans(is.na(total_df))

#for the green votes impute by the mean of the neighborhoods that touch a neighborhood with
    na value
index <- st_touches(total_df, total_df)
total_df <- total_df %>%
  mutate(Green_votes_perc = ifelse(is.na(Green_votes_perc),
                                   apply(index, 1, function(i){mean(.$Green_votes_perc[i])}),
                                   Green_votes_perc))

#drop the geometry column
total_df <- total_df %>% st_drop_geometry()

#impute the other values using knn with 10 neighbors
total_df$BU_CODE <- as.numeric(total_df$BU_CODE)
total_df <- knnImp(total_df, k = 10)
colMeans(is.na(total_df))

#create a copy for later use and delete BU_CODE column
final_df <- total_df
total_df <- total_df[,!names(total_df) %in% c("BU_CODE")]

############################################################################################
#ev users

#create own dataframe
home_chargers_model_df <- total_df[,!names(total_df) %in% c("occupancy_rate")]

#test on over dispersion under simple poisson regression
fmp <- glm(ev_users ~ Solar_power_kw , data = home_chargers_model_df, family=poisson)
dispersiontest(fmp)
```

```
#small p value so over dispersion -> therefore negative binomial model

#print all the columns in the dataframe
colnames(home_chargers_model_df)

#the following model was selected after trial and error using step down method
model_ev_users <- zeroinfl(ev_users ~ Solar_power+Dist_school+Perc_rental_house+Dist_hospital+
                             Perc_green_votes+Perc_low_educated+Dist_shops+Perc_self_employed+
                             Perc_65_EO_yr | Solar_power+Perc_rental_house+Avg_household_size+
                             Dist_hospital+Perc_green_votes+Dist_shops+Perc_self_employed+
                             Perc_45_64_yr+Perc_65_EO_yr, data = home_chargers_model_df, dist = "negbin")

#create mcfadden pseudo r squared value
pR2(model_ev_users)
summary(model_ev_users)

#check whether it is statistical better than zero model
InterceptModel <- update(model_ev_users, . ~ 1)
logLik(InterceptModel)
pchisq(2 * (logLik(model_ev_users) - logLik(InterceptModel)), df = 3, lower.tail=FALSE)

#create start values for bootstrap method
dput(coef(model_ev_users, "count"))
dput(coef(model_ev_users, "zero"))

#perform bootstrap experiments
f <- function(data, i) {
  require(pscl)
  m <- zeroinfl(ev_users ~ Solar_power_kw+AF_ONDBAS+P_HUURWON+AF_ZIEK_E+
                  Green_votes_perc+P_OPL_LG+shop_distance+P_ARB_ZS+P_65_EO_JR | Solar_power_kw+
                  P_HUURWON +GEM_HH_GR+AF_ZIEK_E+Green_votes_perc+shop_distance+P_ARB_ZS+P_45_64_JR+P_65_EO_JR, data = data[i, ], dist="negbin",
                start = list(count = c(1.368, 0.001, -0.244, 0.007, -0.048, 0.035, -2.553, -0.103, 0.011, -0.26), zero = c(-23.262, 0.002, 0.118, 5.006, -0.266, -0.250, 0.910, 0.076, 0.147, 0.090)))
  print(as.vector(t(do.call(rbind, coef(summary(m)))[, 1:2])))
}
set.seed(10)
res <- boot(home_chargers_model_df, f, R = 3000, parallel = "snow", ncpus = 1)

#create bca confidence intervals for coefficients
parms <- t(sapply(c(1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ,35, 37, 39, 41), function(i) {
  out <- boot.ci(res, index = c(i, i + 1), type = c("perc", "bca"))
  with(out, c(Est = t0, pLL = percent[4], pUL = percent[5],
              bcaLL = bca[4], bcaUL = bca[5]))
}))

#construct matrix
var_names <- names(coef(model_ev_users))
var_names <- append(var_names, "count_theta", 10)
row.names(parms) <- var_names
parms

#now do the same for the transformed coefficients; h = exp
expparms <- t(sapply(c(1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ,35, 37, 39, 41), function(i) {
  out <- boot.ci(res, index = c(i, i + 1), type = c("bca"), h = exp)
  with(out, c(Est = t0, bcaLL = bca[4], bcaUL = bca[5]))
}))

## add row names
row.names(expparms) <- var_names
expparms

#create boxplot of the confidence intervals
```

```
par(mar=c(11,8,1,1))
boxplot <- boxplot(t(head(expparms,11)), las =2, outline = TRUE)

##########################################################################################
#occupancy rate model

#construct modelling dataframe
occupancy_model_df <- total_df
names(occupancy_model_df)[names(occupancy_model_df) == ’AANT_INW’] <- ’Num_residents’
occupancy_model_df <- total_df[,!names(total_df) %in% c(”home_chargers”,”NumberChargingPoints”)]
colnames(occupancy_model_df)
occupancy_model_df$occupancy_rate<- occupancy_model_df$avg_hours/730.48
occupancy_model_df <- occupancy_model_df[,!names(occupancy_model_df) %in% c(”avg_hours”)]

#linearity check
pairs(~occupancy_rate + Num_residents + Solar_power  + Car_density + Perc_high_income + Perc_low_income + Dist_school
    + Has_school + Perc_rental_house, data = occupancy_model_df)

pairs(~occupancy_rate +Has_public_point + Avg_household_size + Perc_household_child + Dist_hospital +
        Has_hospital + Perc_high_educated + Perc_low_educated + Population_density, data = occupancy_model_df)

pairs(~occupancy_rate  + Dist_shops + Perc_self_employed + Dist_train_station + Has_city_centre + Perc_15_24_yr +
        Perc_25_44_yr + Perc_45_64_yr + Perc_65_EO_yr, data = occupancy_model_df)

#correlations check
cor_mat <- cor(occupancy_model_df)
cor_mat[cor_mat > 0.8]
cor_mat[cor_mat < -0.8]
corrplot(cor_mat, method=”number”,tl.cex=0.5, number.cex=0.5)
occupancy_model_df <- occupancy_model_df[,!names(occupancy_model_df) %in% c(”P_HH_M_K”, ”P_LAAG_INK_H”)]

#with the step down method the following model was constructed
model_occupancy <- lm(occupancy_rate ~ Num_residents+Solar_power+
                        Perc_high_income+Has_school+
                        Avg_household_size+Perc_green_votes + Perc_high_educated+Population_density+
                        Perc_self_employed+Perc_45_64_yr+Perc_65_EO_yr, data = occupancy_model_df)
summary(model_occupancy)

#check no autoregression
acf(model_occupancy$residuals)
pacf(model_occupancy$residuals)

#test first lag autocorrelation
dwtest(model_occupancy)
# -> there is no significant autocorrelation in the model

#check normality of error terms
plot(model_occupancy)
ols_test_normality(model_occupancy)

#check residuals
mean(model_occupancy$residuals)
plot(model_occupancy$residuals)

#bootstrap experiments
dput(coef(model_occupancy))
res_occupancy <- Boot(model_occupancy, R=3000)

#create confidence intervals
conf_occupancy <- confint(res_occupancy)
conf_occupancy <- cbind(coef(model_occupancy), conf_occupancy)
colnames(conf_occupancy) <- c(”Est”, ”bcaLL”, ”bcaUL”)
conf_occupancy
```

```

#create boxplot of the confidence intervals
par(mar=c(8,5,1,1))
boxplot(t(conf_occupancy), las =2)
```

Listing A.7: Code for the creation of the two statistical models.

```
#policy analysis

#make use of the ev users model and the finaldf created in model.R and determine potential scores
final_df$potential <- predict(model_ev_users, final_df,
                              type = c(”count”))
selection <- final_df[,c(’BU_CODE’, ’potential’)]
selection$BU_CODE <- as.character(selection$BU_CODE)

#add potential to neighborhoodsMRA
neighborhoodsMRA <- left_join(neighborhoodsMRA, selection, by = ’BU_CODE’)
neighborhoodsMRA$unused_potential <- neighborhoodsMRA$potential - neighborhoodsMRA$ev_users
neighborhoodsMRA$prediction_error <- neighborhoodsMRA$potential - neighborhoodsMRA$ev_users

#change negative potential to zero
neighborhoodsMRA <- neighborhoodsMRA %>% mutate(unused_potential = if_else(unused_potential < 0.5, 0, unused_potential))

#create plots
tm_shape(neighborhoodsMRA) +
  tm_fill(col = ’prediction_error’) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.5,
            legend.position = c(”right”,”top”),
            legend.bg.color = ”white”,
            legend.bg.alpha = 0)


tm_shape(neighborhoodsMRA) +
  tm_fill(col = ’unused_potential’, breaks = c(0,0.5,5,10,20,40)) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.5,
            legend.position = c(”right”,”top”),
            legend.bg.color = ”white”,
            legend.bg.alpha = 0)

#determine third quartile value
quantile(neighborhoodsMRA$occupancy_rate, na.rm= TRUE)

#mark neighborhoods as additional charger based on this value
neighborhoodsMRA <-neighborhoodsMRA %>% mutate(high_demand = if_else(occupancy_rate >= 0.3493349, 1, 0))
neighborhoodsMRA$additional_chargers <- neighborhoodsMRA$high_demand * neighborhoodsMRA$unused_potential
neighborhoodsMRA <-neighborhoodsMRA %>% mutate(additional_chargers = if_else(additional_chargers >  0, 1, 0))

#create plot
tm_shape(neighborhoodsMRA) +
  tm_fill(col = ’additional_chargers’, breaks = c(0,0.5,1)) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.5,
            legend.position = c(”right”,”top”),
            legend.bg.color = ”white”,
            legend.bg.alpha = 0)

#do the same for the ev user/charger analysis
neighborhoodsMRA$usersPerCharger <- neighborhoodsMRA$ev_users/neighborhoodsMRA$charging_points
quantile(neighborhoodsMRA$usersPerCharger, na.rm = TRUE)
neighborhoodsMRA$potentialPerCharger <- neighborhoodsMRA$potential/neighborhoodsMRA$charging_points
neighborhoodsMRA <-neighborhoodsMRA %>% mutate(additional_chargers2 = if_else(potentialPerCharger >=  1.333333, 1, 0))
```

```
#create plot
tm_shape(neighborhoodsMRA) +
  tm_fill(col = 'additional_chargers2', breaks = c(0,0.5,1)) +
  tm_layout(legend.title.size = 0.8,
            legend.text.size = 0.5,
            legend.position = c("right","top"),
            legend.bg.color = "white",
            legend.bg.alpha = 0)
```

Listing A.8: Code for policy analysis.

```
#analysis for appendix on fact_chargersessions

correct_records <- function(chargingData){
  #clear rows with NA in critical columns
  chargingData <- chargingData[!is.na(chargingData$RFID_skey),]
  chargingData <- chargingData[!is.na(chargingData$UseType),]
  chargingData <- chargingData[!is.na(chargingData$kWh),]
  chargingData <- chargingData[!is.na(chargingData$IsValid),]
  chargingData <- chargingData[!is.na(chargingData$StartConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$EndConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$ChargePoint_skey),]

  #clear rows with values that are not correct
  chargingData <- chargingData[chargingData$UseType == 'regulier',]#only regulier usetype
  chargingData <- chargingData[chargingData$RFID_skey >= 0,]
  chargingData <- chargingData[chargingData$kWh > 0,]
  chargingData <- chargingData[chargingData$IsValid == 1,]
  chargingData <- chargingData[chargingData$ChargePoint_skey %in% relevant_chargers,]#
    relevant_chargers as in Relevant_charging_points script

  DIM_LOCATION <- DIM_LOCATION(username, password)

  #add location data to chargindata
  charingPointsData <- left_join(chargingData, DIM_LOCATION)
  #delete sessions where no postalcode is known
  charingPointsData <- charingPointsData[!is.na(charingPointsData$PostalCode),]
  charingPointsData <- charingPointsData[!is.na(charingPointsData$Location_skey),]
  return(charingPointsData)
}

correct_records_occupancy <- function(chargingData){
  #clear rows with NA in critical columns
  chargingData <- chargingData[!is.na(chargingData$RFID_skey),]
  chargingData <- chargingData[!is.na(chargingData$kWh),]
  chargingData <- chargingData[!is.na(chargingData$IsValid),]
  chargingData <- chargingData[!is.na(chargingData$StartConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$EndConnectionDateTime),]
  chargingData <- chargingData[!is.na(chargingData$ChargePoint_skey),]

  #clear rows with values that are not correct
  chargingData <- chargingData[chargingData$RFID_skey >= 0,]
  chargingData <- chargingData[chargingData$kWh > 0,]
  chargingData <- chargingData[chargingData$IsValid == 1,]
  chargingData <- chargingData[chargingData$ChargePoint_skey %in% relevant_chargers,]#
    relevant_chargers as in Relevant_charging_points script

  DIM_LOCATION <- DIM_LOCATION(username, password)

  #add location data to chargindata
  charingPointsData <- left_join(chargingData, DIM_LOCATION)
  #delete sessions where no postalcode is known
  charingPointsData <- charingPointsData[!is.na(charingPointsData$PostalCode),]
  charingPointsData <- charingPointsData[!is.na(charingPointsData$Location_skey),]
  return(charingPointsData)
}


#januari
data_jan <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
```

```
                                startDateView = "'2022-01-01'", endDateView = "'2022-01-31'")
nrow(data_jan)
#191749
correct_data_jan <- correct_records(data_jan)
nrow(correct_data_jan)
#176610
correct_data_jan_occupancy <- correct_records_occupancy(data_jan)
nrow(correct_data_jan_occupancy)
#186686

#februari
data_feb <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                                startDateView = "'2022-02-01'", endDateView = "'2022-02-28'")
nrow(data_feb)
#191015
correct_data_feb <- correct_records(data_feb)
nrow(correct_data_feb)
#175178
correct_data_feb_occupancy <- correct_records_occupancy(data_feb)
nrow(correct_data_feb_occupancy)
#185082


#march
data_mar <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                                startDateView = "'2022-03-01'", endDateView = "'2022-03-31'")
nrow(data_mar)
#247207
correct_data_mar <- correct_records(data_mar)
nrow(correct_data_mar)
#231595
correct_data_mar_occupancy <- correct_records_occupancy(data_mar)
nrow(correct_data_mar_occupancy)
#232040



#april
data_apr <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                                startDateView = "'2022-04-01'", endDateView = "'2022-04-30'")
nrow(data_apr)
#241266
correct_data_apr <- correct_records(data_apr)
nrow(correct_data_apr)
#226307
correct_data_apr_occupancy <- correct_records_occupancy(data_apr)
nrow(correct_data_apr_occupancy)
#226805

#may
data_may <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                                startDateView = "'2022-05-01'", endDateView = "'2022-05-31'")
nrow(data_may)
#220645
correct_data_may <- correct_records(data_may)
nrow(correct_data_may)
#207432
correct_data_may_occupancy <- correct_records_occupancy(data_may)
nrow(correct_data_may_occupancy)
#207854

#june
data_jun <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                                startDateView = "'2022-06-01'", endDateView = "'2022-06-30'")
nrow(data_jun)
#238243
correct_data_jun <- correct_records(data_jun)
nrow(correct_data_jun)
#222687
correct_data_jun_occupancy <- correct_records_occupancy(data_jun)
nrow(correct_data_jun_occupancy)
```

```
#223138

#july
data_jul <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-07-01'", endDateView = "'2022-07-31'")
nrow(data_jul)
#236643
correct_data_jul <- correct_records(data_jul)
nrow(correct_data_jul)
#222496
correct_data_jul_occupancy <- correct_records_occupancy(data_jul)
nrow(correct_data_jul_occupancy)
#222911

#august
data_aug <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-08-01'", endDateView = "'2022-08-31'")
nrow(data_aug)
#229407
correct_data_aug <- correct_records(data_aug)
nrow(correct_data_aug)
#215816
correct_data_aug_occupancy <- correct_records_occupancy(data_aug)
nrow(correct_data_aug_occupancy)
#216189

#september
data_sep <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-09-01'", endDateView = "'2022-09-30'")
nrow(data_sep)
#269567
correct_data_sep <- correct_records(data_sep)
nrow(correct_data_sep)
#252450
correct_data_sep_occupancy <- correct_records_occupancy(data_sep)
nrow(correct_data_sep_occupancy)
#252824

#october
data_oct <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-10-01'", endDateView = "'2022-10-31'")
nrow(data_oct)
#291390
correct_data_oct <- correct_records(data_oct)
nrow(correct_data_oct)
#273126
correct_data_oct_occupancy <- correct_records_occupancy(data_oct)
nrow(correct_data_oct_occupancy)
#273518

#november
data_nov <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-11-01'", endDateView = "'2022-11-30'")
nrow(data_nov)
#316076
correct_data_nov <- correct_records(data_nov)
nrow(correct_data_nov)
#296060
correct_data_nov_occupancy <- correct_records_occupancy(data_nov)
nrow(correct_data_nov_occupancy)
#296491

#december
data_dec <- FACT_CHARGESESSION(username,password,country = "'NLD'" ,
                               startDateView = "'2022-12-01'", endDateView = "'2022-12-31'")
nrow(data_dec)
#342804
correct_data_dec <- correct_records(data_dec)
nrow(correct_data_dec)
#318718
correct_data_dec_occupancy <- correct_records_occupancy(data_dec)
```

```
nrow(correct_data_dec_occupancy)
#319165
```

Listing A.9: Code for describing the number of used transactions per month.

A.8.2. Python Code

```
import pandas as pd
import json

#read in data from geo.json file
with open('tk2021.geo.json') as data_file:
    data = json.load(data_file)
df = pd.json_normalize(data, 'features')

#make selection and drop na values
df = df.groupby('properties.Adres').sum(numeric_only=True)
df = df[['properties.Geldige stemmen', 'properties.SP (Socialistische Partij)', 'properties.
    GROENLINKS', 'properties.Partij voor de Dieren', 'properties.Partij van de Arbeid (P.v.d.
    A.)']]
df = df.dropna()
df = df.apply(pd.to_numeric)
df.reset_index(inplace = True)

#load mathcing from postcode to neighborhood
match_postcode = pd.read_csv('pc6-gwb2020.csv', sep = ';')
match_postcode = match_postcode[['PC6', 'Buurt2020']]
match_postcode['Buurt2020'] = match_postcode['Buurt2020'].astype(str)
match_postcode['Buurt2020'] = match_postcode['Buurt2020'].str.zfill(8)
match_postcode['Buurt2020'] = "BU" + match_postcode['Buurt2020']

#group by neighborhood and determine percentage green votes
df = df.merge(match_postcode, left_on='properties.Adres', right_on='PC6')
df = df.groupby('Buurt2020').sum(numeric_only=True)
df['Green_votes'] = df['properties.Partij voor de Dieren'] + df['properties.GROENLINKS'] + df
    ['properties.SP (Socialistische Partij)'] + df['properties.Partij van de Arbeid (P.v.d.A
    .)']
df['Green_votes_perc'] = (df['Green_votes']/df['properties.Geldige stemmen'])*100
df.reset_index(inplace=True)
df = df[['Buurt2020', 'Green_votes_perc']]
df.rename({'Buurt2020': 'BU_CODE'}, axis=1, inplace=True)
df.to_csv('greenVotes.csv')
```

Listing A.10: Code for the generating of green vote percentages.

```
#import dependencies
import pandas as pd
import geopandas as gpd
from scipy.stats import ks_2samp
import scipy.stats as stats
import matplotlib.pyplot as plt
import numpy as np

#read in the data of the shapefile
imd_shp = 'WijkBuurtkaart_2020_v3\\buurt_2020_v3.shp'
imd = gpd.read_file(imd_shp)

#process car_density variable
df_cars = pd.read_csv('df_cars.csv')
imd  = imd.merge(df_cars, on='BU_CODE', how='left')
imd['Car_density'] = imd['G_PAU_KM']/imd['AANT_INW']
imd.replace([np.inf, -np.inf], np.nan, inplace=True)

#process Green_votes_perc
df_green_votes = pd.read_csv('greenVotes.csv')
df_green_votes = df_green_votes[['BU_CODE', 'Green_votes_perc']]
imd  = imd.merge(df_green_votes, on='BU_CODE', how='left')

#process Solar_power_kw
df_solar_panels = pd.read_csv('solarpanels.csv')
df_solar_panels = df_solar_panels[['BU_CODE', 'Solar_power_kw']]
```

```
imd  = imd.merge(df_solar_panels, on='BU_CODE', how='left')

#reconstruct shop_distance
imd['shop_distance'] = (imd['AF_APOTH'] + imd['AF_SUPERM'] + imd['AF_WARENH'] + imd['AF_RESTAU'])/4
imd['P_OPL_HG'] = imd['A_OPL_HG']/ imd['AANT_INW']
imd['P_OPL_LG'] = imd['A_OPL_LG']/ imd['AANT_INW']

df = pd.read_excel('Gemeenten alfabetisch 2022.xlsx')
df = df[['GemeentecodeGM', 'Provincienaam']]
df_total = pd.merge(imd, df, left_on='GM_CODE', right_on='GemeentecodeGM')
print(df_total.shape)

#select the MRA region
df_selection = df_total[df_total['Provincienaam'].isin(['Flevoland', 'Noord-Holland', 'Utrecht'])]
df_selection = df_selection[~df_selection['GM_NAAM'].isin(['Amsterdam', 'Utrecht'])]

#select the rest of the Netherlands
df_selection2 = df_total[~df_total['Provincienaam'].isin(['Flevoland', 'Noord-Holland', 'Utrecht'])]

#select columns to check
columns_to_check = ['AF_TREINST',
'AF_ZIEK_E',
'AF_ONDBAS',
'P_LAAGINKP',
'P_HOOGINKP',
'P_OPL_HG',
'P_OPL_LG',
'P_15_24_JR',
'P_25_44_JR',
'P_45_64_JR',
'P_65_EO_JR',
'P_ARB_ZS',
'P_HH_M_K',
'P_HUURWON',
'BEV_DICHTH',
'GEM_HH_GR',
'AANT_INW',
'Car_density',
'Dist_shops',
'Solar_power_kw',
'Green_votes_perc']

#create mapping for variable names
columns_to_check_mapping = ['Dist_train_station',
'Dist_hospital',
'Dist_school',
'Perc_low_income',
'Perc_high_income',
'Perc_high_educated',
'Perc_low_educated',
'Perc_15_24_yr',
'Perc_25_44_yr',
'Perc_45_64_yr',
'Perc_65_EO_yr',
'Perc_self_employed',
'Perc_household_child',
'Perc_rental_house',
'Population_density',
'Avg_household_size',
'Num_residents',
'Car_density',
'Dist_shops',
'Solar_power',
'Perc_green_votes']

#make the figures
for index,column in enumerate(columns_to_check):
    fig, ax = plt.subplots(figsize = (8,4))
```

```
    df_selection = df_selection[df_selection[column] >= 0]
    ax = df_selection[column].plot.kde(label = f’MRA {columns_to_check_mapping[index]}’, legend = True)
    df_selection2 = df_selection2[df_selection2[column] >= 0]
    ax = df_selection2[column].plot.kde(label = f’NL {columns_to_check_mapping[index]}’, legend = True)

    #no assumptions on distribution
    ax.plot([], [], ’ ’, label=f”P-value KS-test: {round(ks_2samp(df_selection[column], df_selection2[column])[1],5)}”)

    #test on normality
    if stats.shapiro(df_selection[column])[0] < 0.05 :
        #not normally distributed
        ax.plot([], [], ’ ’, label=f”P-value Wilcoxon-test: {round(stats.ranksums(x=df_selection[column], y=df_selection2[column])[1],5)}”)
    else:
        #normal distribution
        ax.plot([], [], ’ ’, label=f”P-value T-test: {round(stats.ttest_ind(a=df_selection[column], b=df_selection2[column], equal_var=True)[1],5)}”)
    plt.legend()
    fig.savefig(f’figures/{columns_to_check_mapping[index]}.png’)
```

Listing A.11: Code for comparing the distribution between the MRA-Elektrisch region and the rest of the Netherlands.