Amsterdam University of Applied Sciences

A prediction model for the demand of charging stations for electric vehicles in the municipality of Amsterdam

Steenbrink, M.; Jochemsen, T.; Piersma, N.; van der Mei, R.; Dugunji, E.

[Link to publication](#)

A prediction model for the demand of charging stations for electric vehicles in the municipality of Amsterdam

5 authors, including:

Nanda Piersma
Amsterdam University of Applied Sciences/C…
30 PUBLICATIONS 711 CITATIONS

SEE PROFILE

Rob van der Mei
Centrum Wiskunde & Informatica
209 PUBLICATIONS 792 CITATIONS

SEE PROFILE

Elenna R. Dugundji
Centrum Wiskunde & Informatica
30 PUBLICATIONS 450 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:
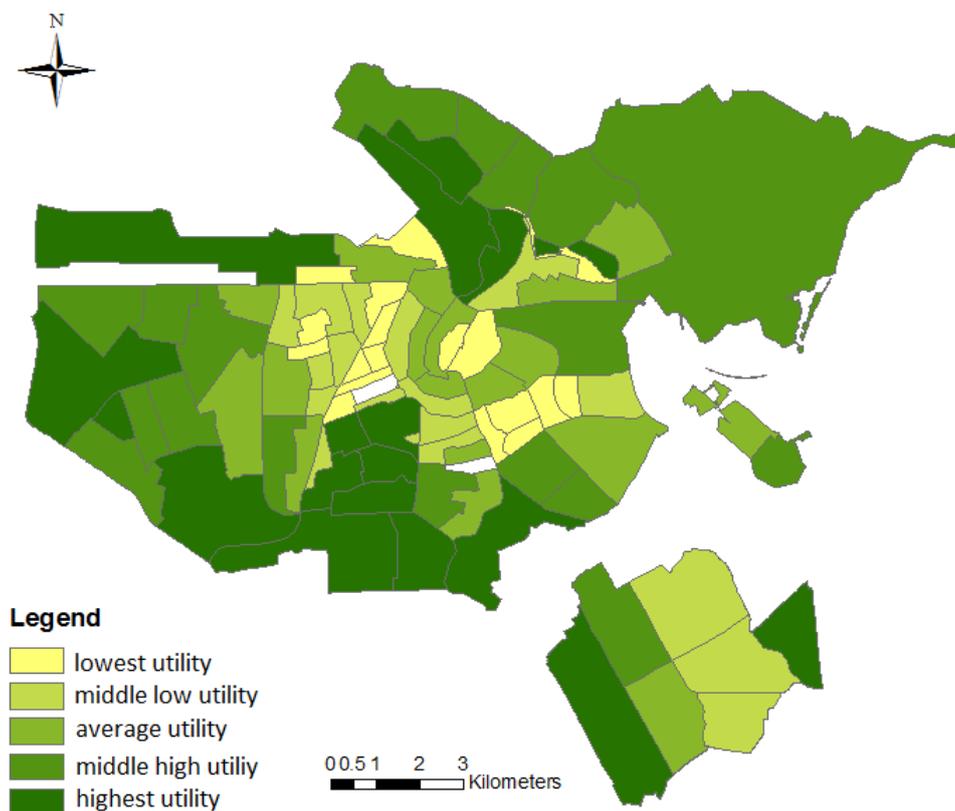
Project Sociodynamic Discrete Choice View project

Project Twitter View project

A prediction model for the demand of charging stations for electric vehicles in the municipality of Amsterdam

Merel Steenbrink, Tirza Jochemsen, Nanda Piersma,

Rob van der Mei, Elenna Dugundji

12 February 2016

A prediction model for the demand of charging stations for electric vehicles in the municipality of Amsterdam

Merel Steenbrink, Tirza Jochemsen, Nanda Piersma,
Rob van der Mei, Elenna Dugundji

12 February 2016

Abstract

To stimulate the usage of electric cars, the municipality of Amsterdam wants to significantly increase the number of charging stations. The question is if it is possible to determine and predict where these charging stations are necessary. Using charging behavioural data from 2012, 2013, and 2014 a discrete choice model was determined. The charging stations were aggregated to neighbourhood level and next a multinomial logit model was developed. This model describes the probability that a car owner will charge his/her car in that neighbourhood, for each neighbourhood. This way it can be determined in which neighbourhoods the most charging station are and will be needed.

1. Introduction

The paper will focus on discrete choice models and their practical use for modelling behaviour involved in the usage of charging stations for electric vehicles in the municipality of Amsterdam. For each neighbourhood the utility of present and possible future charging stations will be determined.

The social relevance of this research will be discussed first, focusing on the vision of the municipality of Amsterdam and the other relevant research in the field. The paper will then digress about discrete choice models, before applying the choice model to describe and predict the utility of charging stations.

1.1. Social relevance and Amsterdam's vision

In 1998 the European Union (EU) signed the Kyoto Protocol, which stated a goal to reduce emissions of greenhouse gasses by 20 percent by 2020. The transport sector is a growing contributor to the emission of greenhouse gasses. In 2010 the sector was responsible for about 23 percent of the total energy-related CO_2 emissions and an increase in income and transport has led to an increase in the use of cars. [17] The car was good for 15 to 30 percent of the total journeys in the Western countries in 2007, with Western Europe at 50 percent and America at a staggering 90 percent. In China the car sales increased from 2.4 million in 2001 to 7.2 million in 2006. [12]

The use of electric vehicles could therefore significantly decrease the emission of greenhouse gasses, reduce our dependence on fossil fuels, and provide the transition to renewable energy sources. [10] Electric vehicles are even less noisy, resulting in less noise pollution.

Amsterdam as a municipality also wants to increase the rate at which the city is getting more sustainable, focusing mostly on clean air and the use of renewable energy. [20] Amsterdam wants to achieve emission-free transport and is setting goals such as a 30 percent decrease of soot emissions and 35 percent decrease of nitrogen dioxide emissions by 2025. [1]They therefore actively stimulate the use of electric vehicles. Amsterdam is a pioneer in electric vehicles and to maintain this status has set the goal of increasing the number of charging stations from 1000 to 4000 by 2018. [1] To realise this vision effectively it is important to have a placing strategy that produces not only an efficient network, but also the best possible support for users.

1.2. Literature overview

In past years many research has been performed towards the optimization of the placement of charging stations, where the loading demand was first estimated to then tackle the optimisation problem of placement. Chen, Kockelman and Khan [6] predicted the loading demand using travel reports of households and individuals in Washington. Considered factors were parking spots, parking duration, and car destination. Frade et. al. [7] made a distinction between day and night time demand and based then loading demand estimate on the number of cars per household, the average travel time, and the percentage of employed individuals. Both papers did not consider the distribution of demand, where Liu et. al. [13] did take this into consideration. Ge, Feng and Liu [8] first partitioned the area, such that each partition had one loading pole. They then used a genetic algorithm to choose the best position within each partition. Long et. al. [14] approached the problem as a weighted graph and optimized the routes towards each loading pole. He et. al. [10] used the assumption that the location of a loading pole influenced the destination of the driver. Consequently, they investigated the equilibrium between loading pole availability, destination choice, and electricity prices to determine the loading pole positions using the active-set method.

The methods discussed are all based on small datasets, excluding more robust methods which require large data sets. In this paper a *discrete choice model* can be applied as a large data set is made available by Amsterdam.

1.3. Research question

In this paper two questions will be answered, firstly "How can behavioural data about the usage of charging stations for electric vehicles in the municipality of Amsterdam be modelled to deduce the demand of existing locations?", and secondly "How can the estimated model be used to predict the demand of future locations?".

2. Approach

As said earlier, the municipality of Amsterdam want to increase the number of charging stations from 1000 now, to 4000 in 2018. The policy in place at the moment is reactive: when a citizen buys an electric vehicle, he or she can apply for a charging station near their house. The application is processed and an investigation is done into possible placements. This procedure takes time. When charging station demand can be predicted beforehand, the process can become proactive and as a result the up-scaling process can be accelerated.

2.1. Method

In order to predict where charging stations are needed, a discrete choice model is applied to the behavioural data about charging sessions in the municipality of Amsterdam. From 2012 onwards Amsterdam collected data from each charging station about who, when and how long someone charged their car. The data set consists of 135,051 observations, where each observation consisted inter alia of the user-id, the charging station-id, the date and the connection-time. Each loading session descries the choice of a car owner (choice maker) for a specific charging station (alternative). This data can be used to determine a utility function. When this is obtained, it is possible to determine the utility of future charging station. This way the charging station with the highest utility can be determined before an application is filed.

2.2. Aggregation of alternatives

When the choice set consists of a large number of alternatives two problems can arise. Firstly, a large number of alternatives can lead to a large computing time. Secondly, it is possible that not all the characteristics of the alternatives are available for each alternative. For example: when the choice set consists of locations in a city, some attributes may only be available per neighbourhood. These problems can be solved by using an aggregated set, instead of the universal set of alternatives.
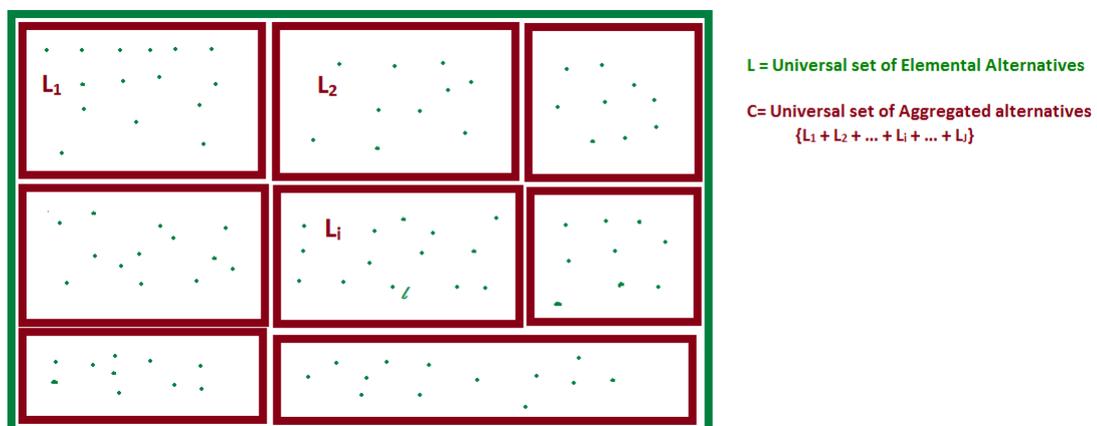

L = Universal set of Elemental Alternatives

C= Universal set of Aggregated alternatives
$\{L_1 + L_2 + ... + L_i + ... + L_J\}$

Figure 2.4.: Elementary and aggregated alternatives

2.3. Sampling of alternatives

Another way to deal with a large data set and the corresponding long computing time is the sampling of data. There are two methods for sampling data, either the observations or the alternatives can be sampled. If a decent number of observations are available more information can be gathered from a data set by using many observations with little alternatives, then by using a small amount of observations with many alternatives. This mostly applies to socio-economic properties [3]. As such, in this paper it was decided to sample the alternatives.

3. Application

3.1. Model

Due to the large number of charging station and the lack of specific characteristics for each charging station, the decision was made to aggregate the choice options to a neighbourhood level. A choice for a charging station is now a choice for a neighbourhood. The neighbourhoods are assigned as the municipality of Amsterdam assigns them and consists of 97 different neighbourhoods. The data is made anonymous, resulting in the choice maker having no specific characteristics. This means the overall utility and not the utility for one specific user is measured. As 97 neighbourhood is still a large choice set and it is not possible to determine a specific subset for every choice maker, it was chosen to sample the choice set. In the first model, to every observation with chosen neighbourhood, a random sample of five neighbourhoods was added. Also, when a neighbourhood did not contain a single charging station, the neighbourhood was not a member of the choice set. However, when the five neigbourhoods are chosen randomly, the mutual correlation between the neighbourhoods is not taken into account. In the second model it is. To do so, the division of the city into 6 "stadsdelen", city districts, was used. To every observation we now added one random sampled neighbourhood of each district. An observation consists of the chosen neighbourhood and five other neighboorhouds, each situated in another district. Again a Multinomial Logit Model was estimated, resulting in the second model. Afterwards, with the same data setup, a cross-nested logit model was estimated, resulting in model 3.

To determine the utility function, the program Biogeme [4] was used. On the basis of a list of observations the program determines the β's of the systematic component of the utility function. Each observation here consisting of the chosen neighbourhood with its characteristics and five randomly sampled neighbourhoods with their characteristics.

3.2. Characteristics

The model uses four characteristics to describe neighbourhoods. Three of the characteristics are founded on Statistics Netherlands (CBS) data and one characteristic is deduced from the behavioural data. Furthermore, the number of charging stations per neighbourhood is also taken into account as the data is aggregated.

3.2.1. CBS data

Three characteristics of neighbourhoods have been founded on key figures of 2012, 2013 and 2014 of the CBS [2]. The characteristics in question are: the average number of cars per household, the percentage of privately owned homes and percentage of Western citizens (for a more detailed description of the characteristics see the appendix). The characteristics were chosen as they had the biggest impact on the charging station choice. The choice was made to not include more characteristics as many characteristics showed a high correlation. The chosen characteristics did not show a strong correlation (see table 3.1). As the average income was not known for every neighbourhood, it was not used as a characteristic in the model. Furthermore, investigation also showed a correlation (Correlation of -0.486 with t-test 0.58) between average income and percentage of Western citizens, meaning that the characteristic was considered indirectly. See table 3.1 for correlations between the chosen characteristics.

Table 3.1.: Correlation between the characteristics

Coefficient 1	Coefficient 2	Correlation
Laadgraad	Westers	0.195
Koopwoningen	Laadgraad	0.125
Auto	Laadgraad	-0.00754
Koopwoningen	Westers	-0.0784
Auto	Westers	0.215
Auto	Koopwoning	-0.476

3.2.2. Charging degree

For each year the average charging degree per neighbourhood has been calculated. First, the charging degree per charging station was calculated by dividing the total connection time by the total existence time of the station times two (as each charging station has two connections). This calculation specifically uses the connection time and not the charging time, as the first represents the time the charging station is occupied. Afterwards, the average charging degree was calculated.

3.2.3. Number of stations

The number of charging stations was documented for each year. If this number was equal to zero, the neighbourhood was not a member of the choice set. The factor is M_i as described in 2.6, meaning the logarithm of the number of stations was considered and used to find the coefficient $\frac{1}{\mu}$..

3.3. Results

As said before, the results of the estimations can be found in table 3.2. As each characteristic has a different unit, the value of the characteristic says very little about the importance of the characteristic. The different characteristics are therefore almost

incomparable in a way. If for instance the unit of a characteristic changes resulting in the characteristic being multiplied by ten, the coefficient will just divide by ten. In this model the goal was to keep the coefficients between -10 and 10. The t-statistic is a better indicator of the importance of a characteristic and it can be seen that the average number of cars and the percentage of privately owned homes are the most important factors. The first characteristic has a positive coefficient, meaning that a neighbourhood with a high number of cars per household has an increased probability for charging activity. This is intuitively clear, as there are, on average, more cars in the neighbourhood, the chances there are more electric cars in the neighbourhood which need to be charged. The percentage of privately owned homes has a negative coefficient however, which indicates that as the number of privately owned homes increases the charging probability decreases. This can be explained by the fact that home owners sometimes have the possibility to build a private charging station on their own driveway, and therefore less public charging stations are needed.

Table 3.2.: Estimation results Model 1

Variable abbreviation	Variable description	Coefficient estimation	t-statistic
AUTO	Average number of cars per household	7,43	289,44
KOOPWNG	percentage of privately owned homes	-0,0762	-160.13
WST	percentage of Western citizens	0,0558	76,41
LDGR	Charging degree	0,801	10,39
$1/\mu$	Logarithm of the number of poles	0.919	96.19

Summary statistics	Model 1
number of observations	134985
$L(0)$	-241978,908
ρ^2	0,655
$\bar{\rho}^2$	0,655

3.4. Prediction

With these results it is possible to calculate the probability for each neighbourhood that a car owner chooses to charge there. This way the neighbourhoods can be ordered by probability, low to high. In the neighbourhoods with the highest probabilities the largest number of charging stations are needed.

A prediction can be made by entering the characteristics of the neighbourhoods. When the data in 2015 becomes available, the chances can be calculated using

$$P_n(i) = \frac{e^{V_{in} + 0,909 \log(M_i)}}{\sum_{j \in C_n} e^{V_{in} + 0,909 \log(M_i)}},$$

where

$V_{in} = 7,43\text{AUTO}_i - 0.0762\text{KOOPWNG}_i + 0.0558\text{WST}_i + 0.801\text{LDGR}_i$

M_i = number of charging stations in the neighbourhood

such that the neighbourhoods with the highest probability are the neighbourhoods where the most charging stations are needed.

The data from 2014 supply an ordering by probabilities. The actual results can be found in the appendix, where the neighbourhoods are ordered from highest to lowest probability. These results are presented visually in figure 3.1. The darker the neighbourhood is, the higher the probability.
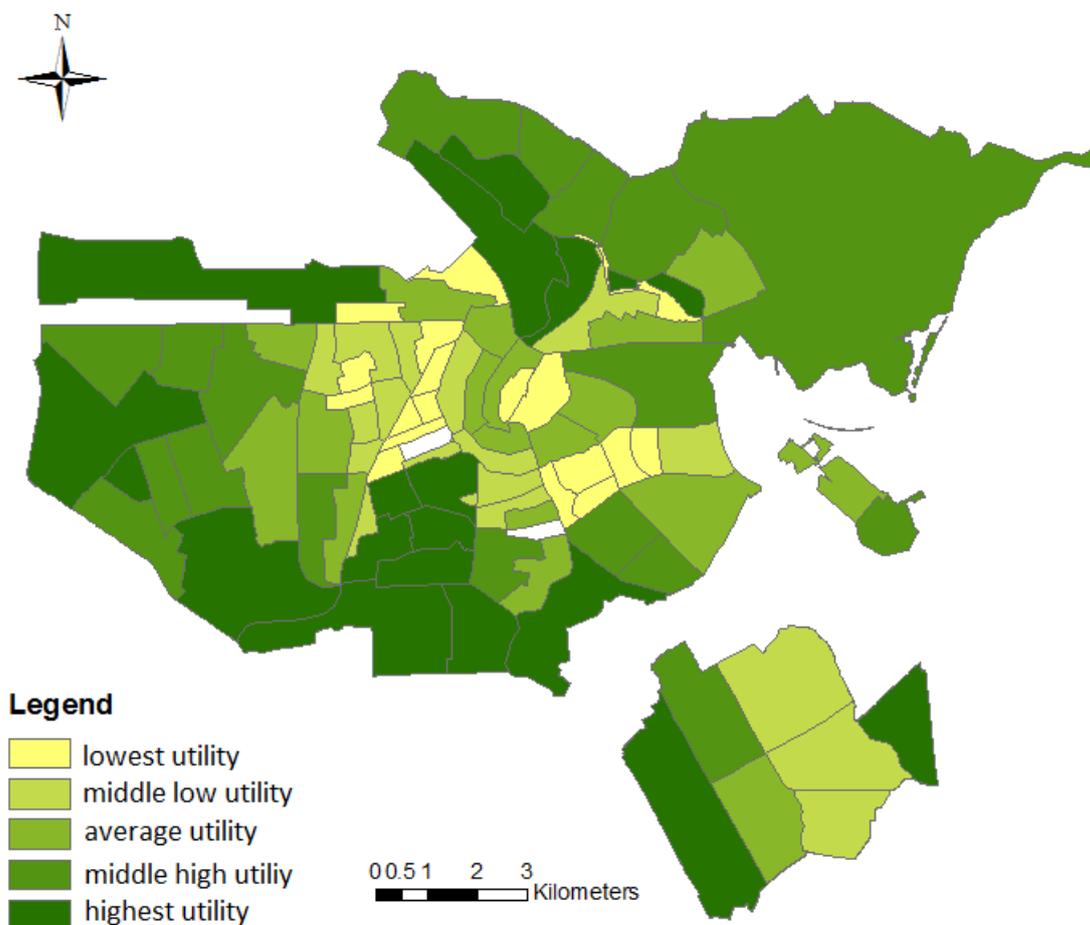

Figure 3.1.: Results, Utility 2014

4. Conclusion

This paper investigated how behavioural data about the usage of charging stations for electric vehicles in the municipality of Amsterdam can be modelled and used to predict the future demand on the charging stations. Due to the large number of data an approach using a discrete choice model was implemented. The model describes the probability that a car owner chooses to charge at a certain station. The choice options are aggregated to neighbourhood level, meaning a choice for a charging station is a choice for a neighbourhood. Using this system an ordering can be made as to which neighbourhoods have the highest probability of being chosen. The probability is a function of several characteristics, namely the average number of cars per household, the percentage of privately owned homes, the percentages of Western citizens and the charging degree. As the choice options are aggregated, the number of stations per neighbourhood is also considered. Using these characteristics a probability for each neighbourhood can be calculated, indicating the chance a car owner will charge in that neighbourhood. Using these results it can be predicted where the demand for charging stations is the highest, namely in the neighbourhoods with the highest probability.

5. Popular summary

Electric vehicles emit less environmentally harmful gasses than regular cars and make a contribution to the transition into renewable resources. Furthermore, they are less noisy then regular cars. This fits perfectly in the vision of the municipality of Amsterdam, who sees this transition as a priority, as well as cleaner air. Therefore it doesn't come as a surprise that the municipality is stimulating the use of electric vehicles actively. An example of this is placing a big amount of charging stations in Amsterdam: right now, there are more than 1000 charging stations in Amsterdam. The neighbourhood Amsterdam Zuid even has the highest density of charging stations in the world! The municipality wants to maintain this leading position and therefore wants to upscale the number of stations to 4.000 in 2018. To do this deliberately it is important to think about a strategy where on the one hand the user is optimally supported, and on the other hand an efficient network is created. From 2012 on, the activity at every charging station was recorded. This thesis dealt with the question of how this rich behavioural data could be modelled to predict where future stations would be most needed. To model the data a Discrete Choice Model was used. This model consists of four components: the choice maker, alternatives, attributes and the decision rule. The decision rule describes the choice mechanism, and therefore which alternative the choice maker chooses. A widely used decision rule is the Utility function, U_{in}, which describes the utility of every alternative (i) for a certain choice maker (n). The utility function consists of two components: the systematic component (V_{in}) and the disturbance factor (ε_{in}). The systematic component is a linear function of the attributes, the disturbance factor compensates for measurement errors and unobserved attributes. This results in:

$$U_{in} = V_{in} + \varepsilon_{in} \tag{5.1}$$

$$V_{in} = \beta_1 x_{in1} + \beta_2 x_{in2} + \cdots + \beta_k x_{inK} \tag{5.2}$$

The choice maker will choose the alternative that has the highest utility, so the highest value of Uin. However it may happen that a choice maker chooses inconsistently, or that two choice makers with similar attributes will choose differently. Therefore you may consider the choice process as a stochastic process. The probability of choice maker n choosing alternative i is the probability of the utility of this option being higher than all other options:

$$Pn(i) = P(U_{in} U_{jn} \text{for all options j}) \tag{5.3}$$

In this thesis the choice maker was the driver of the electric vehicle. The alternatives were the different charging stations in the city. The choice to be made was: at which stations will the driver charge? If every station would be considered as one alternative, the choice set would be way too big. Furthermore, not every charging station had unique, distinguishing attributes. Therefore we chose to aggregate the stations to the level of neighbourhoods. Every neighbourhood is now one alternative. The utility function describes the utility of the choice for a charging station in a certain neighbourhood. The utility is a function of the following characteristics: average number of cars per household, percentage of apartments, percentage of Western inhabitants and the average charging degree. Besides this, a factor is added with the number of charging stations to compensate for the fact that a neighbourhood with more stations will be chosen more frequently. With this model the probability that a driver chooses to charge their car in a certain neighbourhood can be calculated. In this way the need for new charging stations can be calculated: in the neighbourhoods with the highest probability, new stations are most needed.

A. CBS: explanation of characteristics

A.1. Percentage of Western citizens

Total Western citizens [%]: The number of immigrants on January 1st, expressed in an integer percentage of the total number of citizens. This number is deduced from the structural census of municipal basis administration. Someone belong to the category "total Western citizens" if they originate from Europe North-America, Oceania, Indonesia and Japan. The percentage is stated when there are 50 or more citizens in the neighbourhood.

A.2. Percentage of privately owned homes

Privately owned homes [%], reference date: January 1st), as a percentage of the total number of homes and only stated when there are 20 homes or more per neighbourhood and when the number of homes with an unknown ownership is below 50 percent.

A.3. Cars per household

The number of cars per household on January 1st). The cars are regionally classified using the license plate registration. Cars that are registered to the address of a lease or rent company can therefore distort the results. The number of cars per household is stated with a minimum of 50 households and a value of at most 2.5 cars per household

Bibliography

[1] Infographic duurzaam amsterdam. `https://www.amsterdam.nl/gemeente/volg-beleid/agenda-duurzaamheid/infographic-duurzaam/`. Accessed: 2015-10-20.

[2] Kerncijfers wijken en buurten 2004-2015. `http://www.cbs.nl/nl-NL/menu/themas/dossiers/nederland-regionaal/cijfers/incidenteel/maatwerk/wijk-buurtstatistieken/kwb-recent/default.htm`. Accessed: 2015-11-20.

[3] M. Ben-Akiva and S. R. Lerman. *Discrete Choice Analysis, Theory and Aplication to Travel demand*. MIT PRess, Cambridge, Massachusetts, 1985.

[4] M. Bierlaire. Biogeme: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland*, 2003.

[5] P.H.L. Bovy, S. Bekhor, and C.G Prato. The factor of revisited path size,alternative derivation. *Transportation Research Record: Journal of the Transportation Research Board*, (2076):132140, 2008.

[6] T. Chen, K. Kockelman, and M. Khan. The electric vehicle charging station location problem: A parking-based assignment method for seattle. *Transportation Research Record*, 2013.

[7] I. Frade, A. Ribeiro, G. Goncalves, and A. Pais Antunes. Optimal location of charging stations for electric vehicles in a neighborhood in lisbon, portugal. *Transportations Research Record: Journal of the Transportation Research Board*, (2252):91–98, 2011).

[8] S. Ge, L. Feng, and H Liu. The planning of electric vehicle charging station based on grid partition method. *IEEE, Electrical and Control Engineering Conference, Yichang, China*, 2011.

[9] C. A. Guevara-Cue. Endogeneity and sampling of alternatives in spatial choice models. *Massachusetts Institute of Technology*, 2010.

[10] F. He, D. Wu, Y. Yin, and Y. Guan. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transportations Research Part B*, (47):87–101, 2012.

[11] N. Johnson and S. Kotz. *Continuous Univariate Distributions, Volume 1*. 1994.

[12] S. Kahn Ribeiro, M. S. Kobayashi, J. Beuthe, D. Gasca, D. S. Greene, Y. Lee, P. J. Muromachi, S. Newton, D. Plotkin, R. Sperling, P. J. Wit, and Zhou. Transport and its infrastructure. *Transportations Research Part B*, (47):87–101, 2007.

[13] Z. Liu, Wen, W. Fushuan, and G. Ledwich. Optimal planning of electric-vehicle charging stations in distribution systems. *Transactions on Powder Delivery, IEEE*, 28(1):102–110, 2013.

[14] J.I.A. Long, H.U. Zechun, S.O.N.G Yonghua, and Zhuowei Luo. Optimal siting and sizing of electric vehicle charging stations. *Electric Vehicle Conference (IEVC), IEEE International*, 2012.

[15] D. McFadden. Conditional logit anaysis of qualitative choice behavior. *Frontiers in Econometics*, 4:105–142, 1974.

[16] D. McFadden. Moddelling the choice of residential location. *Transportations Research Part B*, (47):87–101, 1978.

[17] R.R. Sims, F. Schaeffer, X. Creutzig, M. Cruz-Nez, D. DAgosto, M.J. Dimitriu, L. Figueroa Meza, S. Fulton, O. Kobayashi, A. Lah, P. McKinnon, M. Newman, J.J. Ouyang, D. Schauer, Sperling, and G. Tiwari. Transport. *Mitigation of Climate Change,Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.*, pages 599–690, 2014.

[18] J. Stada. Discrete keuze theorie. `https://www.mech.kuleuven.be/cib/verkeer/dwn/cursusteksten-voor-2011/les-9-discrete-keuzetheorie-2011.pdf`.

[19] G. Troncoso Parady. Introduction to discrete choice models. *The 14th Behavior Modeling Summer School, September 25-27*, 2015. Accessed: 2015-10-10.

[20] Gemeenteraad van Amsterdam. Duurzaam amsterdam: Agenda voor duurzame energie, schone lucht, een circulaire economie en een klimaatbestendige stad. *Spatial Interacion Theory and Planning Models*, pages 75–96, 11 maart 2015.