SMART CITY PLANNING – DEVELOPING AN URBAN CHARGING INFRASTRUCTURE FOR ELECTRIC VEHICLES

Complete Research

Wagner, Sebastian, University of Freiburg, Germany, sebastian.wagner@is.uni-freiburg.de
Brandt, Tobias, University of Freiburg, Germany, tobias.brandt@is.uni-freiburg.de
Neumann, Dirk, University of Freiburg, Germany, dirk.neumann@is.uni-freiburg.de

Abstract

Electric mobility is considered to be an essential building block of a sustainable transportation paradigm. Particularly for the urban metropolis of the 21st century electric vehicles embody a promise of decreased air pollution and increased quality of living. However, city planners need to muster enormous investments into urban charging infrastructure in advance. We develop a decision support system to help city planners use these funds more efficiently and place charging infrastructure where it is truly needed. We construct a regression model to determine factors that influence the utilization of charge points in Amsterdam, one of the most "electrified" cities in the world. We include data on more than 50,000 points of interest, for instance shopping malls and museums, as well as charging data from 273 charging points with 427 individual outlets in the city of Amsterdam. From these influences we derive a decision support system that places charge points where they are expected to experience the highest utilization.

Keywords: E-Mobility, Smart City Planning, Decision Support, Business Intelligence.
1 Introduction

In recent years, electric mobility has received substantial public as well as academic attention. Considered a key technology of a transition towards a more sustainable way of living, electric vehicles powered by renewable energy allow transportation without environmentally harmful CO2 emissions. The governments of several countries, including the United States, China, and Germany, have already invested billions of dollars in subsidies to promote this technology. However, sales of electric vehicles (EVs) or plugin-hybrid EVs are still substantially below target (Electric Drive Transportation Association, 2013). This appears to be caused by high costs of acquisition on the one hand, but also by a frame of mind called "range anxiety" (Eberle and Helmolt R., 2010) on the other. Lacking novel battery technologies, this anxiety can only be alleviated by an accessible charge point infrastructure that allows drivers to recharge the batteries of their vehicles when necessary.

The benefits of EVs are particularly relevant for cities and urban agglomerations. Distances are short and many urban areas - in developed as well as developing countries - suffer from toxic air pollution, which severely affects health and quality of living (Kraemer et al., 2000, The New York Times, 2013). Transitioning from cars powered by combustion engines to EVs substantially decreases traffic-related air pollution, thereby increasing the appeal of the entire region.

However, even in urban areas range anxiety remains a major issue. Residents refrain from acquiring EVs as long as there is no charge point infrastructure and commercial infrastructure providers (for instance gas stations and utility companies) have no interest to invest in charge points unless there is a sizable number of customers i.e. owners of EVs. City planners around the world are aware of this chicken-egg-problem and invest millions of dollars to provide the initial spark to the rise of electric mobility by erecting an extensive charge point infrastructure.

In this paper, we take the experiences of a city that has been at the forefront of this development for many years, Amsterdam, to develop an application that supports city planners in using these public funds most efficiently. We construct a beta regression model to explain the utilization of more than 250 charge points in Amsterdam. For this we ask the question "Where do residents need public charge points the most?" and focus particularly on points of interest (POIs), for instance shopping malls, schools, and museums. Using the Google Places API we collect data on more than 50,000 such POIs. Borrowing from urban economic theory we use the proximity of these POIs to charge points to explain their respective utilization. The regression produces several POI categories that significantly influence charge point utilization, thereby explaining a large share of total variation. These categories are then used to develop a decision support system that optimally places new charge points based on their expected utilization.

This paper is structured as follows. In the next section, we discuss research related to our work. In Section 3 we present the empirical study with an introduction of the collected data, followed by methodological summary of the beta regression as well as an evaluation of the regression results. Section 4 contains the development of the decision support system based on the empirical investigation. Section 5 concludes and provides an outlook on our future work.

2 Related Work

Urban transportation has become an important subfield of urban economics and engineering research in recent years. As electric mobility becomes increasingly relevant to a sustainable society, optimal placement of the necessary charging infrastructure is a major challenge to city planners in urban centers. A central issue is to align the placement of charge points with behavioural patterns of electric vehicle

---

1 Drivers worrying over the short driving range of electric vehicles due to limited battery capacity.
drivers. For instance, a recent mobility study by Liu et al. (2012) summarizes the main objectives of private transportation as getting to work, shopping, recreational activities, private errands, and private transport. This study also indicates that parking time varies between one and seven hours. Hence, decisions concerning charging infrastructure should be made in context of these patterns.

The actual development of charging infrastructure for EVs has been discussed extensively in recent years, especially under consideration of public budget constraints. This stresses the need for decision support in this matter. Several case studies, for instance in Beijing, Stockholm, or Taiwan, have been conducted with the objective of planning an urban charge point infrastructure. Various programming and optimization schemes were used in order to minimize investments and operational costs (Liu et al., 2012, Long et al., 2012, Wang, 2008, Wang and Liu, 2011). Furthermore, a model introduced by Chen et al. (2013) assigns optimal charge point location by combining regressions to predict parking demand with a facility location problem. The objective function of the model minimizes total cost of access as a function of walking distances between zones weighted by parking duration. Ge et al. (2012) introduce a planning model for a charge point infrastructure, which combines aspects of the road network with traffic flow, structure, and capacity constraints. The model tries to minimize investment and operational costs for all stakeholders. Feng et al. (2012a) and Tang et al. (2011) use a weighted Voronoi diagram, to minimize the power loss of EVs before reaching the next charging station on the one hand and to maximize the annual operating income of charge points on the other hand. Moreover, Feng et al. (2012b) design a charge point infrastructure on trunk roads using queuing theory. The location decision is derived by maximizing the expected number of EVs that need to be charged, subject to service cost and waiting fees for customers.

However, for the development of the optimal charge point infrastructure for a given city, planners need to consider geographical and environmental constraints on trips and charging times of current EV. Hence, Frade et al. (2011) formulate a discrete maximum covering model with decay and capacity restrictions to determine charge point locations. The model considers temperature, daytime, and charging demands. It was tested in a case study for a neighbourhood in Lisbon, Portugal, with the number of charge points to be placed as an exogenous number. Case studies for Chicago, Seattle, and Ohio try to investigate the optimal locations for charging stations, applying integer programming schemes (Xi et al., 2013, Andrews et al., 2013).

Furthermore, Hess et al. (2012) set up a genetic programming model to find charge point locations by minimizing the average trip duration of EVs. The placement of charge points is determined by the expected mobility patterns of EV and the approach includes a depletion and charging model, as well as a general mobility model for route adaption. Also Ip et al. (2010) formulate a linear programming model to optimize charge point allocation. It prepares and aggregates road traffic information into demand clusters through hierarchical analysis.

Additional research by He et al. (2013) deals with a game theoretical approach that examines the interactions among influences such as the availability of public charging opportunities, destination, price of electricity, and route choices of EVs. Optimal allocation of charge points is then conducted by a mathematical program, based on an equilibrium model. However, as the model is of a strategic nature, it does not determine the exact locations and capacities of the placed charging stations. Wirges et al. (2012) formulate a dynamic spatial EV charging infrastructure model for 2020 in the region of Stuttgart.

As can be seen in the literature review above, several approaches towards optimal placement of EV charge points exist. However, this research is rarely based on theoretically justified empirical analyses of real-world data, thereby lacking a link to reality necessary to inform decision-makers. Hence, the central contribution of this work is twofold. On the one hand, we provide a rigorous empirical evaluation of possibly the most comprehensive data set on urban charging patterns in a major city - Amsterdam. On the other hand, this analysis is based on a sound theoretical foundation from transportation research and urban economics. We will explore both aspects in more detail in the next section. The results of our
analysis are subsequently transformed into an application that can inform the decisions of city planners in urban centers around the world.

3 Empirical Analysis of Charge Point Utilization

Setting up a public charge point infrastructure in a major city easily requires several million dollars in implementation costs. In 2009, the city of Amsterdam launched the Amsterdam Electric initiative, which aimed to have 200 charge points by 2012 within city limits. In June 2013, Amsterdam offered several hundred charge points to electric drivers which are used on a daily basis. The utilization data on more than 273 of these charge points is publicly available. In this section we present a regression model that uses this data to determine which factors influence whether a charge point is frequently used or not. In the next section we use the results from the regression analysis to develop a system that optimally places new charge points based on these factors and helps city planners to efficiently use public funds.

3.1 Data Set

The data set used for this research contains three parts. First, we have collected data on the utilization of 273 charge points with a total of 427 outlets in Amsterdam. All outlets are normal Type II chargers with a power supply of \( \leq 22 \) kW. While the city of Amsterdam is responsible for planning the infrastructure, the providers of the charging stations are Essent as part of the RWE AG, Nuon as a part of Vattenfall, and Coulomb Technologies as an electric vehicle infrastructure company. This information is publicly available through a website which returns at any given time the total number of outlets at a specific charge point and the number of outlets currently in use. We collected this information for every charge point every minute for six months to derive the average utilization of every respective charge point outlet during that time. This produces 273 utilization values in the interval \([0, 1]\) which serve as the dependent variable in our regression.

Secondly, we employed the Google Places API to derive coordinates for more than 50,000 POIs in Amsterdam. The POIs include a total of 93 categories, for instance bars, parks, museums, clothing stores, and banks. The respective proximities of charge points to every POI, sorted by categories, are used as explanatory variables in the regression.

Thirdly, we used data from the Dutch central bureau of statistics on population density and economic situation of residents in the various neighbourhoods of Amsterdam (Centraal Bureau voor de Statistiek, 2013). We associated the GPS coordinates of each charge point with the relevant neighbourhood and used population density and average income as control variables in our regression.

3.2 Methodology and Theoretical Background

The objective of our regression model is to explain the variation in utilization of charge points. Utilization captures the share of available time (i.e. excluding maintenance) during which the outlets of a charge point are on average used to charge an electric vehicle. Hence, it is a value in the interval \([0, 1]\) with 0 and 1 indicating a charge point that is never, respectively always, used. Due to the interval restriction a simple linear regression is unsuitable. Similarly, since the variable is continuous on the unit interval a logit or probit regression is cumbersome to use, as well, since both capture the probability of a discrete event occurring. Therefore, we employ a beta regression model as proposed by Ferrari and Cribari-Neto (2004), which has been specifically designed for modelling rates and proportions. The density of the beta function is given by

\[
\pi(u; p, q) = \frac{\Gamma(p + q)}{\Gamma(p)\Gamma(q)} u^{p-1}(1 - u)^{q-1},
\]

\[0 < u < 1\]

(Wagner et al. / Smart City Planning for Electric Vehicles, Twenty Second European Conference on Information Systems, Tel Aviv 2014)
for utilization $u$, parameters $p$ and $q$, and $\Gamma(\cdot)$ as the gamma function. This assumes that the utilization follows a beta law. However, the beta distribution is very flexible and allows for a wide range of shapes, depending on the parameterization. Consider, for instance, Figure 1, which shows the histogram of utilization values. A beta distribution with $p = 1.3$ and $q = 5.95$ appears to approximate the histogram quite well, which strongly indicates that utilization follows a beta law.

The regression model can be obtained by assuming that the expected value of a specific $u_i$ (the expected utilization of a particular charge point $i$) following the density given in Equation 1 can be written as

$$g(E(u_i)) = \sum_{j=1}^{k} x_{ij} \beta_j,$$

(2)

where $g(\cdot)$ is a link function that maps $[0, 1]$ into $\mathbb{R}$, $\beta = (\beta_1, ..., \beta_k)^T$ is a vector of unknown regression parameters, and $x_{i1}, ..., x_{ik}$ are observations on $k$ covariates. The regression estimation is performed by maximum likelihood. Out of several possible link functions we choose the logit transformation $g(z) = \ln\left[\frac{z}{1-z}\right]$, since this provides easily interpretable coefficients. Specifically, for a given utilization $u_i$ the effect of a change in a single covariate $\Delta x_{ij}$ can be explained by the following relationship with $u^*$ as the utilization after the change:

$$e^{\beta_i \Delta x_{ij}} = \frac{u_i^*/(1-u_i^*)}{u_i/(1-u_i)}.$$

(3)

The covariates in the regression model consist of 93 POI categories (including a category for other charge points in the vicinity), average income and population density data on the neighbourhoods where the respective charge point is located, as well as the distance to the center of Amsterdam from the location of the charge point. While the inclusion of the latter covariates into the model is rather straightforward, we will now explain how we included the POI data into the regression model.

Figure 1. Histogram of charge point utilization compared to a beta density function

The "Why?" is in that context the more pressing question - if POIs did not matter, we do not even need to discuss how to include them. The central argument for the importance of POIs derives from one of the most striking differences between electric vehicles and their classical counterparts: refuelling. While the stop at a gas station with a gasoline-powered car rarely takes more than five minutes, refilling an EV may take between 20 minutes (fast-charging station and small battery) up to several hours. As people are unlikely to just sit in their cars for that duration, EV charging becomes more a matter of parking that of refuelling - people will simply run errands while the EV is parked at a charge Point. There have been several publications within the transportation and urban economics literature concerning the determinants of parking decisions made by car drivers, for instance (van der Goot, 1982, Hunt and Teply, 1993, Arnott and Rowse, 1999). The theoretical and empirical models in these studies contain a few common determinants: walking distance from parking lot to final destination, cost of parking, and location of the parking lot with respect to other lots and the route of the car driver.
While these publications consider individual decision-making of drivers, these determinants provide a basis for the aggregated view taken in this work. We want to determine which charge Points – i.e. parking lots with a charging option – are most appealing and expected to experience the highest utilization. In Amsterdam electric vehicles park for free, so cost of parking is not an issue. Our POI-measure captures the other determinants, walking distance and driving routes. POIs are by definition possible final destinations of drivers. Since we use various categories of POIs, we can estimate which types of POI are most appealing to EV drivers – if charge Points close to a shopping center exhibit a high utilization, we deduce that shopping centers are often final destinations of drivers, subject to other possible influences.

To transfer these insights into our empirical investigation we need to model how the relevance of a POI changes with distance. For instance, for parking decisions in Amsterdam the Van Gogh Museum across the street is more relevant than the Museum of Modern Art in New York. Although exaggerated, this example clearly shows that the relevance of a POI decreases with distance. Van der Goot (1982) uses a linear transformation to model this relationship, mapping the distance to the final destination on the unit interval. This implies that the appeal of a parking lot decreases the farther it is from the final destination. We simply invert this relationship, arguing that the relevance of a final destination (POI) decreases the farther it is from the parking lot - the driver did not decide to park at a specific lot because of a POI that is very far away; hence, this POI becomes irrelevant. To model the relevance \( r_{ki} \) of a specific POI \( k \) to charge point \( i \) we use a generic exponential function

\[
r_{ki} = e^{-bd_{ki}},
\]

with \( d_{ki} \) as the distance between the two points in kilometers and \( b \) a constant. Hence, the relevance of a POI that is at the location of the charge point is 1. The distance is decreasing and approaching zero the farther the POI is away from the charge point, with \( b \) determining the impact of the distance. For instance, a relevance of 0.5 would be reached at 346 meters and 173 meters for \( b = 2 \) and \( b = 4 \), respectively. Naturally, the choice of \( b \) has a strong impact on the regression results and is subject to differences across cultures – as mentioned before, it estimates the “willingness to walk”. In the model by van der Goot (1982) an upper bound of 40 minutes for walking time is set as a determinant of parking spot selection. We tested several values for Amsterdam and eventually used a value of \( b = 2 \) in our regression model, which corresponds to a walk of approximately 25 minutes.

The value of the covariate \( x_{ij} \) for POIs is the sum of all relevance values for a particular POI category.

\[
x_{ij} = \rho_{ij} = \sum_{p_k \in P_j} r_{kij} \tag{5}
\]

with \( p_k \) as a specific POI and \( P_j \) as the set of POIs that belong to category \( j \). Hence, we refer to \( \rho_{ij} \) as the density of POI category \( j \) at point \( i \), for instance “bank-density” and “park-density”.

### 3.3 Regression Results and Evaluation

In the following section, we conducted the beta regression using the betareg-package for R (Francisco Cribari-Neto and Achim Zeileis, 2010). Initially we included all covariates into the regression: the population density (PopDens), the logarithmic average income per person (LogIncPP), the distance from the center of Amsterdam (CenterDist), the charge point density (CPDensity), and the densities of all 93 POI categories. While the calculated pseudo-R2 value for this model was 0.410 (adjusted 0.238), it did not work well to explain utilization. The central reason for this is that the regression under these initial specifications in driven by a few POI-categories with only very few elements - for instance, casinos, bowling alleys, or zoos. Overall, there were 22 categories with less than 20 elements each. This means that in the entire city of Amsterdam there are less than 20 places with this particular tag. For the regression, this poses the problem of spurious correlations. Thereby, the regression attributes an effect falsely to such a rare category because it is at a particular place purely by coincidence. Through this, the
much more pronounced effect of, for instance, the central business district is diminished. There is also
a more subtle problem with multicollinearity, since several categories are very similar or even identical.
For instance, POIs belonging to the category "restaurant" in most cases also belong to the category
"café". Hence, densities for these categories are very similar and, as is generally the case with
multicollinearity, coefficients and significance levels can be deceiving. However, this is only
problematic for categories that are more or less identical. In all other cases, coefficients describe
interactions between POI-categories. For instance, a restaurant that is also tagged as a café may have a
different influence on charge point utilization than one that is not. We derived our final regression
through an iterative fitting procedure. Starting with the control variables introduced above we iteratively
added selected POI covariates to maximize the adjusted pseudo-R2 of the model. This model contains
41 POI covariates for a pseudo-R2 of 0.319 and an adjusted pseudo-R2 of 0.239. The fact that the
adjusted value actually increases compared to the initial specification is further confirmation that some
small categories masked the effect of the categories that actually drive utilization. The regression results
are summarized in Table 1.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-1.946</td>
<td>(1.121)</td>
<td>*</td>
</tr>
<tr>
<td>CenterDist</td>
<td>-0.021</td>
<td>(0.067)</td>
<td></td>
</tr>
<tr>
<td>PopDens</td>
<td>0.000</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>LogIncPP</td>
<td>0.122</td>
<td>(0.301)</td>
<td></td>
</tr>
<tr>
<td>CP_Density</td>
<td>-0.088</td>
<td>(0.042)</td>
<td>**</td>
</tr>
<tr>
<td>Accounting</td>
<td>0.205</td>
<td>(0.080)</td>
<td>**</td>
</tr>
<tr>
<td>Art Gallery</td>
<td>-0.136</td>
<td>(0.039)</td>
<td>***</td>
</tr>
<tr>
<td>Bakery</td>
<td>-0.163</td>
<td>(0.076)</td>
<td>**</td>
</tr>
<tr>
<td>Bank</td>
<td>0.315</td>
<td>(0.085)</td>
<td>***</td>
</tr>
<tr>
<td>Beauty Salon</td>
<td>0.085</td>
<td>(0.043)</td>
<td>**</td>
</tr>
<tr>
<td>Bicycle Store</td>
<td>0.504</td>
<td>(0.141)</td>
<td>***</td>
</tr>
<tr>
<td>Book Store</td>
<td>-0.125</td>
<td>(0.084)</td>
<td></td>
</tr>
<tr>
<td>Café</td>
<td>-0.026</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>Car Dealer</td>
<td>-0.198</td>
<td>(0.072)</td>
<td>***</td>
</tr>
<tr>
<td>Car Wash</td>
<td>-0.364</td>
<td>(0.281)</td>
<td></td>
</tr>
<tr>
<td>Cemetery</td>
<td>0.630</td>
<td>(0.427)</td>
<td></td>
</tr>
<tr>
<td>Church</td>
<td>0.723</td>
<td>(0.418)</td>
<td>*</td>
</tr>
<tr>
<td>Clothing Store</td>
<td>0.052</td>
<td>(0.037)</td>
<td></td>
</tr>
<tr>
<td>Dentist</td>
<td>0.079</td>
<td>(0.038)</td>
<td>**</td>
</tr>
<tr>
<td>Embassy</td>
<td>-0.160</td>
<td>(0.209)</td>
<td></td>
</tr>
<tr>
<td>Establishment</td>
<td>0.003</td>
<td>(0.002)</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>Standard Error</td>
<td>Significance</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Finance</td>
<td>-0.193</td>
<td>(0.041)</td>
<td>***</td>
</tr>
<tr>
<td>Gas Station</td>
<td>0.280</td>
<td>(0.131)</td>
<td>**</td>
</tr>
<tr>
<td>General Contractor</td>
<td>0.107</td>
<td>(0.059)</td>
<td>*</td>
</tr>
<tr>
<td>Gym</td>
<td>0.366</td>
<td>(0.101)</td>
<td>***</td>
</tr>
<tr>
<td>Home Goods Store</td>
<td>0.097</td>
<td>(0.053)</td>
<td>*</td>
</tr>
<tr>
<td>Hospital</td>
<td>-0.361</td>
<td>(0.101)</td>
<td>***</td>
</tr>
<tr>
<td>Laundry</td>
<td>-0.110</td>
<td>(0.103)</td>
<td></td>
</tr>
<tr>
<td>Lawyer</td>
<td>0.082</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td>0.072</td>
<td>(0.098)</td>
<td></td>
</tr>
<tr>
<td>Locksmith</td>
<td>-0.605</td>
<td>(0.250)</td>
<td>**</td>
</tr>
<tr>
<td>Lodging</td>
<td>0.047</td>
<td>(0.017)</td>
<td>***</td>
</tr>
<tr>
<td>Movie Rental</td>
<td>0.372</td>
<td>(0.207)</td>
<td>*</td>
</tr>
<tr>
<td>Museum</td>
<td>0.200</td>
<td>(0.051)</td>
<td>***</td>
</tr>
<tr>
<td>Night Club</td>
<td>-0.217</td>
<td>(0.122)</td>
<td>*</td>
</tr>
<tr>
<td>Park</td>
<td>-0.135</td>
<td>(0.137)</td>
<td></td>
</tr>
<tr>
<td>Pharmacy</td>
<td>-0.111</td>
<td>(0.106)</td>
<td></td>
</tr>
<tr>
<td>Place of Worship</td>
<td>-0.786</td>
<td>(0.342)</td>
<td>**</td>
</tr>
<tr>
<td>Plumber</td>
<td>-0.108</td>
<td>(0.077)</td>
<td></td>
</tr>
<tr>
<td>Police</td>
<td>-0.661</td>
<td>(0.249)</td>
<td>***</td>
</tr>
<tr>
<td>School</td>
<td>0.016</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Shoe Store</td>
<td>0.129</td>
<td>(0.047)</td>
<td>***</td>
</tr>
<tr>
<td>Spa</td>
<td>0.301</td>
<td>(0.254)</td>
<td></td>
</tr>
<tr>
<td>Travel Agency</td>
<td>-0.207</td>
<td>(0.050)</td>
<td>***</td>
</tr>
<tr>
<td>University</td>
<td>-0.107</td>
<td>(0.137)</td>
<td></td>
</tr>
<tr>
<td>Veterinary Care</td>
<td>0.369</td>
<td>(0.167)</td>
<td>**</td>
</tr>
</tbody>
</table>

Significance indicated at 10% (*), 5% (**), and 1% (***) levels

Table 1. Final Regression Results

Out of 45 covariates, 28 have an influence on utilization at a significance level of at least 10%. The direction of the effect often makes intuitively sense. For instance, the different types of stores have a positive influence, because people drive there to go shopping and park their cars for an extended period of time. Similarly, gas stations are usually close to streets with intense traffic, also suggesting a positive influence. Car dealerships, on the other hand, appear to be unappealing targets for drivers of electric cars. Interpretation is less clear when categories overlap, for instance finance, bank, and accounting. The negative influence of finance is only valid when the POI is neither tagged as a bank nor as accounting. Intuitively, banks and accounting firms have a positive influence, because people may stay there for an extended period of time or they are close to shopping and business districts. Finally, the density of charge points has a significantly negative impact. Naturally, placing a charge point at a location decreases the
expected utilization of further charge points at that location. This is particularly relevant, since this allows us to use the coefficient of charge point density to estimate the impact of existing charge points when exploring the best location to place a new one. We can now use the significant coefficients to create a map of Amsterdam with the most attractive places for charge points given any number of existing charge points. For instance, Figure 2 illustrates the situation for a "clean sheet" Amsterdam, i.e. a situation where no charge points have been placed yet. High utilization is indicated by green shades, low utilization by red shades. The map illustrates the attractiveness of downtown Amsterdam, as well as the area around the Vondelpark or De Pijp. Unsurprisingly, the waterways are less attractive, as is the port area in the northwest and some regions in the southeast.

![Figure 2](image)

*Figure 2. Expected charge point utilization in Amsterdam provided no charge points ranging from 6% (dark red) to 37% (dark green)*

In the next section we will use the information gained from the regression analysis to construct an application that supports city planners in optimal placement of charge points based on heat maps as illustrated in Figure 2.

4 Decision Support System for Optimal Charge Point Placement

In this section we introduce a decision support system for the placement of charge point given the regression coefficients from the analysis summarized in Table 1. We use Python as an object oriented programming scheme to transfer the statistical analysis to a placement scheme for charging infrastructure. The program interacts with different PostgreSQL databases containing all relevant information concerning POIs, charge point usage data, and control variables like population density and average income of the respective area (see Section 3.1). To find the optimal location for a charge point, the algorithm determines the area within the city with the highest expected utilization. At this stage this is the optimal location for a new charge point and the remaining charge points are placed iteratively. However, with every iteration utilization values are recalculated given the density of charge points that have been placed before. This approach can be used to establish a charging infrastructure for any city, provided that the determinants of charging patterns among the population are similar to Amsterdam. It can also use an existing infrastructure and extend it by determining the optimal location of additional charging stations.

We now proceed by briefly summarizing the development and the functionality of the system, followed by a discussion of simulation results.
4.1 System Development

The regression in Section 3 was conducted for the city of Amsterdam, as it provides an extensive data set on charging patterns due to its pioneering role with respect to urban electric mobility. Hence, in this paper we discuss the application of our decision support system to the city of Amsterdam. We will investigate the transferability to other cities in our future work.

We define the total planning area $\mathbb{A}$ as a set of tiles $a_{ij}$ (subareas) as the following matrix

$$
\mathbb{A} = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
$$

(6)

The center of each tile $a_{ij} \in \mathbb{A}$ serves as a possible charge point location, while each tile itself is defined as a triple

$$a_{ij} = (\phi_{ij}, \lambda_{ij}, c_{ij}),$$

(7)

with $\phi_{ij}$ as latitude coordinate, $\lambda_{ij}$ as longitude coordinate, and $c_{ij}$ as the number of charge points within this tile. The size of $\mathbb{A}$ and in particular the number of tiles determents the granularity of the planning approach. Further, we have to initially set a total number of CPs $\gamma$ to be placed by the planning algorithm. In each tile $a_{ij} \in \mathbb{A}$ the approach is able to place at most $\gamma$ charge points, thus, $c_{ij} \in [0, 1, \ldots, \gamma]$. As a consequence, the numbers of charge points for all tiles within the predefined area $\mathbb{A}$ is defined as the following set

$$\mathbb{C} = \{c_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq n}. $$

(8)

This only approximates the ground situation in a city, since city planners need to take into account various circumstances, for instance access to the power supply or a connection to a road with an associated parking spot. However, within the area of such a tile expected utilization is assumed to be constant, with respect to its size. The subsequent planning steps of fine-tuning the location of the tile or more explicit the CP to be set in this subarea has to be carried out by the responsible division of the city administration, for instance building authorities, and is not part of our research. This is also the reason why a more finely-granulated grid is unnecessary, as tiles will increasingly lack access to power or roads.

Eventually, the charge point optimization scheme seeks to maximize the utilization $u_{ij}$ for all tiles $a_{ij} \in \mathbb{A}$ given an initially predefined number of charge points $\gamma$ to be placed. Hence, the objective function is defined as

$$
\max_{\mathbb{C}} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \cdot u_{ij},
$$

(9)

with the following constraint

$$
\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \leq \gamma.
$$

(10)

Since the utilization of each tile $a_{ij}$ is determined by its vector of densities for all covariates $x_{ij} = (x_{ij1}, \ldots, x_{ijk})$ and the only density which changes during runtime is the charge point density $x_{ij1}$, this specific density actually determines the overall output of the objective Function 9. Therefore, we are able to split the optimization formula into two parts, part 1 for all covariates except the CP density (POI categories, population density, etc., respectively $\sum_{k=2}^{K} x_{ijk}$) and part 2 for the charge point density $x_{ij1}$ itself. By substituting $u_{ij}$ with the logit-link function in Equation 2, the optimization model can be expressed as
\[
\max \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \cdot g^{-1} \left( \frac{1}{x_{ij}} \sum_{k=2}^{K} \beta_k \cdot x_{ijk} + \beta_1 \left( \sum_{r=1}^{m} \sum_{s=1}^{n} c_{rs} e^{-b \cdot d(\phi_{ij} \lambda_{ij} \phi_{rs} \lambda_{rs})} \right) \right)
\]

subject to \(\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \leq \gamma\).

Due to its exponential nature, it is not possible to solve the optimization problem (11) in an adequate time frame, even with a large number of high-performance computers. However, since the only density that changes during runtime is \(x_{ij}\), we are able to determine the remaining \(x_{ij}, ..., x_{ijK}\) densities for all covariates \(K\) and all tiles \(i, j \in A\) in a pre-calculation step. Thus, we are up to substitute part 1 by an initially calculated constant \(\epsilon_{ij}\) that determines a factor for all covariates \(\sum_{k=2}^{K} x_{ijk}\). Hence, the above optimization model 11 can be simplified to

\[
\max \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \cdot g^{-1} \left( \epsilon_{ij} + \beta_1 \left( \sum_{r=1}^{m} \sum_{s=1}^{n} c_{rs} e^{-b \cdot d(\phi_{ij} \lambda_{ij} \phi_{rs} \lambda_{rs})} \right) \right)
\]

subject to \(\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} \leq \gamma\).

However, the complexity of the problem is still too high to calculate an optimal solution in a reasonable time, given an appropriate granularity for the initial planning area \(A\). Since the complexity is dependent on the number of tiles of the predefined area, \(\|A\|\) determines the number of calculations and, thus, the runtime of the optimization scheme. Therefore, we decided to calculate the optimal locations for each charge point station iteratively, which greatly reduces the complexity and enables a reasonable runtime. Further, to place CPs one by one is an appropriate assumption, because current CP infrastructures all over the world were constructed in stages as well.

Subsequently, we explain the iterative implementation of the charge point planning approach. In a first step the application calculates the respective densities of each (significant) POI category, by applying Equation 4 to every relevant POI. Thereby, we use a maximum radius of influence of \(\sigma = 2\) km, meaning that the relevance of a specific POI is set equal to zero for any distance greater than \(\sigma\). The only purpose of this limit is to reduce computational load, since the effect on utilization is marginal - according to Equation 4, the relevance of a POI on a location farther away than 2 kilometers is less than 0.02.

The initial value for the charge point density is zero at the beginning of the planning algorithm, as we assume that there are no initial charging stations. We essentially want to find out where Amsterdam would have placed its charging stations using the application introduced in this paper. Therefore, we predefine a number of charge points to be placed \(\gamma\) and run the algorithm as long as this number is reached. The relevance values \(r_{kaij}\) for every POI \(k\) are then added within their respective category for each tile \(a_{ij}\). Thereby, we produce the values for the significant covariates from Table 1 for each tile. Every significant covariate is then multiplied by its respective coefficient and added to produce the logit of the expected utilization. By applying the inverse logit function the application eventually calculates the expected utilization for each tile. The planning scheme chooses now the location with the highest expected utilization \(a_{\text{max}}\) to place the charge point at the center of this location. Further, we have to recalculate the charge point density of all tiles with a distance to \(a_{\text{max}}\) of less than \(\sigma\). Since a newly placed CP increases the charge point density in the surrounding tiles and the significance level of the covariate \(CP\_Density\) is highly negative, the expected utilization for all tiles \(a_{ij} \in A'\) decreases after each placing step. Thereby, the expected utilization for affected locations decreases according to Equation 3, since the coefficient is significant with a value of 0.088. Thus, with an increasing number of charging stations, the surrounding locations become less attractive. In other words, as soon as the number of CPs \(c_{ij}\) in a specific area \(a_{ij}\) increases the CP density \(x_{ij}\) increases too, while the expected
utilization $u_{ij}$ for this area decreases at the same time. As a result, even in popular areas the attractiveness decreases after a few iteration steps, thus, the algorithm selects automatically other subareas solely on the basis of statistical analysis. The application repeats these iteration steps until the predefined number of charging stations $\gamma$ has been reached.

In future simulation runs we use the current charging infrastructure of Amsterdam as initial state and analyze the optimal locations for additional charge points.

### 4.2 Simulation Results

The application algorithm introduced in the previous subsection not only identifies promising locations for charge points, but can also iteratively determine a citywide charging infrastructure. Hence, we apply the approach to the city of Amsterdam by defining a planning area $\mathbb{A}$ around the city center with an edge length of 9 kilometers in a first step. For each tile $a_{ij} \in \mathbb{A}$ we set an edge length of 100 meters, whereby the application divides this area $\mathbb{A}$ into a matrix of 8,100 tiles. Further, we initially set the number of CPs to be placed to $\gamma = 400$. However, this does not necessarily mean that 400 charging stations will be set, since it is possible that the algorithm selects a specific tile multiple times. Accordingly, the planning approach places CP-outlets, rather than CP-stations. Figure 3 illustrates the simulation results for the predefined 81 square kilometer area surrounding downtown Amsterdam, given the initial heat map as illustrated in Figure 2. The blue markers are the locations for charging stations the planning algorithm has determined. Naturally, the charge points are placed in the green areas with downtown Amsterdam, the Vondelpark and De Pijp particularly prominent. However, some charge points are also placed at locations in green shaded areas farther away from the center.

![Figure 3. Simulation results for the city of Amsterdam placing 400 CP-outlets](image1)

![Figure 4. Current charge point locations in Amsterdam](image2)

Now consider Figure 4 which depicts the locations of actual charge points in Amsterdam. Charge point locations are far more distributed over the area than our application would have permitted. Note that in either figure, charge points may contain multiple outlets. This illustrates that city planners in Amsterdam have not only been concerned with utilization, but also other objectives. For instance, they may have intended to reduce range anxiety in EV drivers. Also, many charge points have been set up at locations that are very visible to promote electric mobility even if they are impractical for charging.

To validate our regression and simulation results we perform several additional calculations. First, we place the existing charging points within the “clean sheet” illustrated by Figure 2 to determine the
deviation between the regression results and the real world situation. While the average CP utilization of Amsterdam is approximately 13%, the calculated infrastructure is at ≈16%. The small difference is likely caused by other influences that were not included in our regression model. Furthermore, we place an equal number of charge points in the same planning area. The comparison shows that our stations are almost twice as highly utilized as the real world stations, even by taking into account the above deviation of 3% (real world infrastructure: ≈13%; simulation output: ≈28%).

Certainly, we do not suggest for Amsterdam to tear down its existing charging infrastructure. However, once objectives like visibility and reliability have been achieved the city may increase its focus on utilization. It is clear that the application introduced in this paper can only advise city planners concerning placement decisions of charge points. Particularly as long as electric mobility is still struggling for popular acceptance, other objectives beyond utilization may be just as valid. However, as adoption rates increase and charge points turn into profitable investments, utilization becomes the predominant factor in placement decisions. Thus, the insights gained from our model may provide valuable information not only to city planners, but also to businesses.

5 Conclusion

Electric mobility is one of the most sustainable means of transportation and has received substantial attention in recent years. However, the key to convincing people to switch to a CO2-free technology for driving and to increasing the sales figure of electric vehicles is an adequate charging infrastructure. As the necessary financial investments to build up such an infrastructure easily run into millions, city planners need a way to identify locations where charge points provide the highest value to electric vehicle drivers.

In this paper we have developed a decision support system based on an extensive real-world data set to advise city planners. This system determines the optimal locations for charge points within a given area based on expected utilization. Initially, we have conducted a regression analysis using detailed information on charge point usage of 273 charge points with 427 individual outlets within the city of Amsterdam. In addition, we have collected data on more than 50,000 points of interests in and around Amsterdam to explain the variance in charge point utilization. We have determined the impact of various categories of points of interests using a beta regression. The regression output reveals the impact of specific sites like banks, museums, or night clubs on the utilization. We found that various types of stores have a significantly high influence, likely because people spend a lot of time at these locations. In contrast, other points of interest, for instance bakeries, have a negative influence, since people stay there for too short a time to warrant plugging the vehicle into a charge point outlet.

In a next step we were able to use the categories with a significant influence to derive a city planning application that places charging stations at those locations with the highest expected utilization. The application divides a predefined planning area of 81 square kilometers surrounding the center of Amsterdam into 8,100 tiles and determines the expected charge point utilization for each of these. By applying an iterative placement approach the algorithm constructs an optimal charging infrastructure for electric vehicles - based exclusively on statistical analysis of real world data. However, the collected data set provides just a snapshot of recent charge point usage of electric vehicles in Amsterdam. To validate these results we need to continue monitoring charging events to identify changes in charging behaviour as the popular acceptance of electric mobility and the number of electric vehicles increases.

In our future research we will, furthermore, incorporate additional data sources into the regression model, for instance data on traffic flow and more detailed information on local conditions such as power access, streets, buildings, and pedestrian zones. We will also enhance our decision support system by clustering some categories of points of interests, such as restaurants. These can be further analysed using reviews and ratings in social media. Also further constrains like at least one station every 300 meter will be taken into consideration to reduce range anxiety.
References


Feng, Liang; Ge, Shaoyun; Liu, Hong; Wong Long; Feng, Yan (2012): The planning of charging stations on the urban trunk road. In: IEEE: Innovative Smart Grid Technologies - Asia (ISGT Asia), 2012, pp. 1–4.


