Università di Pisa

Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di laurea magistrale in Informatica Per L'Economia e L'Azienda

Tesi Specialistica

Big Data and Electric Mobility

Candidato:

Ivano Luigi Argano

Relatori:

Prof. Dino Pedreschi

Dr. Salvatore Rinzivillo

Dr. Barbara Furletti

Dr. Lorenzo Gabrielli

Controrelatore:

Prof. Antonio Frangioni

Anno Accademico 2012/2013

Abstract

Nowadays Electric Vehicles are getting more and more important to address modern issues like pollution, economical transportation needs and more efficient and flexible ways of moving. In this thesis we focus on the assessment of an electrification rate of the major urban areas of Tuscany, by simulating the consumption of a real EV on millions of real users trajectories. We propose different usage scenarios, all regarding a different level of sophistication, this to make more reliable evaluations in different environmental conditions, and we study the first, and most important couple of them. Then, we generate the algorithms used for the simulations, and address the challenges met on the path, such as GPS data sampling and elevation extraction issues.

Riassunto

Gli odierni veicoli elettrici stanno assumendo sempre piú importanza come risposta a problemi quali l'inquinamento dell'aria, il bisogno di un mezzo di trasporto piú economico e modalità di spostamento piú efficienti e flessibili. In questa tesi focalizziamo l'attenzione all'individuazione di un tasso di elettrificabilità delle maggiori aree cittadine della Toscana, attraverso una simulazione del consumo di un reale veicolo elettrico su milioni di traiettorie di utenti reali. Proponiamo quindi differenti scenari di uso, tutti riguardanti differenti livelli di sofisticazione, in modo da generare valutazioni piú precise al variare di specifiche condizioni. Studiamo quindi i primi due, e piú importanti scenari. In seguito generiamo gli algoritmi utilizzati per la simulazione, risolvendo tutte le sfide incontrate sul cammino, come il sampling dei dati GPS e i problemi relativi all'estrazione dell'elevazione.

Contents

Introduction

The last few decades have been characterized by a growing interest in renewable energy and the search for more efficient and economical ways of moving, and at the same time, by the availability of a large and ever-growing quantity of data. Those *big data* are complex and unstructured, and are characterized by a high level of detail and their intrinsic difficulty in analyzing them.

The availability of GPS-enabled devices, and the ever-growing presence of GPS devices on board of the vehicles, for insurance purposes for example, makes the location based big data an opportunity to study human mobility behavior and to understand it, in order to find useful information such as the use of city space or the individuation of traffic jams in order to reduce them.

Until now, the oil economy and the inadequate battery performances of the first electric cars have always been a brake for researches on electric vehicles and alternative vehicles technologies, leading the motor companies to invest only on the improvement of internal combustion engines.

However, the panorama is changing fast, and the Electric Vehicles (EVs) technologies are getting more and more important to address modern issues like pollution, economical transportation needs and more efficient and flexible ways of moving.

The first electric cars created were unable to face these goals, but today, with the progress made on battery efficiency, engines efficiency and great optimization, are we able to utilize electric cars for everyday use? Are we able to substitute an internal combustion engine based vehicle with an electric one? This work is intended to shed some light on these questions, by investigating them with the help of location based big data, and by simulating the every day usage of an EV in real life movements.

Contribution and Organization of the Thesis

This work aims at investigating the possibility of using EVs every day. Many works in literature already study this topic, but, differently from this one, they do not use real life trajectories, and do not address all the related issues that may arise. The purpose of this thesis is to study the electrification rate of Tuscany and its major urban areas, running a simulation of the consumption of an EV on millions of real vehicles' trajectories. For this reason, the indicator used here to evaluate the electrification rate is the ability of the EV to reach the end of the day with the battery level $\geq 0\%$: to derive this indicator, we run a simulated discharge and recharge process on the battery, based on the trajectories of users of all types and in real conditions, passing from real roads and being affected by real traffic. In addition, this simulation addresses issues like slope angles and elevation extraction. For this purpose, various usage scenarios are proposed and analyzed, differing from each other in some constraints applied to model different real life conditions.

Here are then summarized the main contributions:

- Creation of a model that simulates the discharging and recharging process of an EV, and application of this model on Tuscany's real users trajectories
- Individuation of the main EVs usage scenarios

- Individuation of an index of electrifiability that shows if the simulation on trajectories fails or not
- Implementation of the simulation using Java programming language

The remainder of the thesis is organized as follows.

In Chapter 1 some of the greatest contributions on EV studies are presented, ranging from Feasibility and Benefit Analysis, passing by Smart Grids and Vehicle2Grid, and up to Connected Vehicles.

Then, in Chapter **??**, the main project's scenarios are explained, scenarios of usage of EV in different conditions and with different constraints applied.

In Chapter 2 the charging and discharging model is presented.

Then, in Chapter 3 the experiments and their results are presented, followed by the conclusions in Chapter 4.

Chapter 1

Literature and Studies on Electric Vehicles

In the last few years, Electric Vehicles have being more and more subject of studies and researches, involving motor companies' research labs and making the governments have a growing interest in this alternative transportation method.

In this chapter we provide an overview of all the studies about Electric Vehicles made until now. In Section 1.1 some related work about the benefits analysis and feasibility studies on EVs are presented, while in Section 1.2 studies on batteries and their State of Charge are analyzed. Then, in Section 1.3 the major studies on power grid management in EVs future scenarios are presented, involving the topics of Vehicle2Grid technologies and Smart Grids, which are the leading arguments of debates regarding the future electric network management in view of the massive introduction of EVs. Following, in Section 1.4 the main contributions on Battery Energy Management and Recharge Modalities are presented, along with some works on future benefits of internet connected vehicles in Section 1.5. Finally, in

Section 1.6 the main work on the topic of human mobility mining is presented, along with the potential of the implemented software described in it.

1.1 Feasibility and Benefits Analysis

Some previous works and researches on Electric Vehicles (EVs) are related to various analysis on the benefits of using EVs and the eventual usage rate of EVs. Here are presented some of these works.

In [1], Yiming Pan et al. want to make an analysis about the widespread use of EVs in the U.S, China, Sweden, and France, countries that are selected as representative, based on power generation, location, economic strength, cultural background and social status. In order to do that, they devise the AGT evaluation method that puts the Analytic Hierarchy Process (AHP) Gray Relation Grade Analysis and TOPSIS together, and *finalize comprehensive evaluation values of promoting electric vehicles in each country.* First of all the authors use the AHP Gray Relation Grade Analysis to execute a hierarchical analysis of Widespread Use, and then they make a comprehensive evaluation value using a ranking method called TOPSIS. This method calculates the distance between the ideal solution and each evaluated object. They conclude that thermal-based countries should promote EVs in advance by issuing subsidies to open the market, and they also suggest that Ethanol-Fuel vehicles are a more practicable response to energy crisis.

In Wenbin Luo et al. [2], the authors use AHP in order to *develop a model for assessing the environmental, social, economic and health impacts of widely using* of EVs. They develop a model in which all the US conventional vehicles are replaced by EVs to assess how much money the country could save by widely using EVs. Then, this model is expanded to estimate how much money the world would save by widely using EVs. The authors finally evaluate the key factors that the governments and vehicle manufacturers may need to consider when determining if and how to support the development and use of the EVs. They *show that when considering environmental and healthy impacts, an EV is more environmental-friendly and healthy-friendly than conventional vehicle. However, when it comes to the economic impacts, conventional vehicle is more acceptable compared to EVs.*

1.2 Battery Simulation and State Of Charge Evaluation

The battery pack is the main energy source for an EV, therefore the estimation of its state of charge is of critical importance. In this section some works related to battery cycle simulation and state of charge estimation are presented.

In [3] the State of Charge (SoC) estimation methodology is studied. The authors begin analyzing various proposed methods of battery SoC estimation, starting from the non-model-based on Coulomb counting method (simple, online but highly sensitive to the current sensor precision) up to Black-box battery models, such as artificial neural networks based models, fuzzy logic models and support vector regression (SVR) based models. These models often provide good results, but they are very computationally heavy and at risk of overfitting. They analyze then a methodology based on extended Kalman filtering (EKF), that has the advantages of being closed-loop (self-corrected), online, and the availability of dynamic SoC estimation error bound. For this reasons the EKF-based model has an increasing popularity, but it has a limited capability of providing robustness against the modeling uncertainty. The authors propose a comparison between a novel robust extended Kalman filter (REKF) and a standard EKF for Li-ion battery SoC estimation using an experimental dataset. They then conclude demonstrating that the REKF-based SoC estimation method has a smaller error bound, and it *has stronger robustness against the noise statistics to some extent, better tolerating the inappropriate tuning of the process and measurement noise covariances in the battery system*.

Another contribution on Battery Simulation has been made by Feng Ju et al. [4]. The authors' aim is to understand in depth the dynamic behavior of batteries and its relationship with manufacturing process. *To achieve this goal, a battery simulation model is needed. Such a model should provide capabilities for performance evaluation and failure prediction, through simulation of cell performance under different conditions.* In such a way, the authors are also enabled in investigating the impacts of changes in working status, temperature, and driving profiles. They also keep into consideration the fact that in a battery all components are correlated with each other, and that the manufacturing cycle is important in determining how the battery will perform. They develop then a simulation framework called *virtual battery*, capable of keeping in consideration internal and external parameters, as well as the manufacturing quality on welding joints (as an additional element).

In the category of Battery Simulation and SoC Evaluation the work by Volker Schwarzer and Reza Ghorbani about Drive Cycle Generation for Design Optimization of Electric Vehicles gives a valuable contribution [5]. The authors propose a Driving Cycle (DCs) Generation Tool that doesn't use a database of recorded data, but creates DCs in *a modular fashion by assigning probabilistic values to each key parameter. The modules are then assembled* *to form a DC according to predetermined rules.* Firstly, they obtain *probability functions for each parameter of a recorded set of DC data*, and then they implement them into a DC generation software tool that produces an unlimited amount of DCs based on the characteristics of the original data set. *The generated DCs are a very precise representation of the original DC data in terms of frequency spectra, speed distribution, acceleration distribution, load characteristics, and occurrence probabilities.* Thus the proposed DC generation methodology is an efficient and highly adjustable tool.

As Sangyoung Park et al. state in [25], the EVs' energy efficiency can be improved in many ways. One of the most important processes that can enhance energy efficiency of EVs is the regenerative braking, that is direct power conversion from the wheel to battery. They also state that the power loss during regenerative braking can be reduced by hybrid energy storage system that use supercapacitors, that can accept high power unlike batteries, which have small rate capability. Their contribution is to introduce systematic enhancement of regenerative braking efficiency for hybrid energy storage systems in EVs, obtaining *19.4% energy efficiency improvement.*

As stated in [26], in order to simulate the behavior of an EV in real-life weather and temperature conditions and its quick charging process, Attila Gollei et al. propose a temperature-dependent model for a simulation of use of a family of batteries used in EVs. The authors in fact want to find a relationship between *the actual magnitude at any instant of the exact charging state and the connection point voltage values as influenced by the deviation of the environmental temperature from the surface temperature of the cell.* They also simulate a *quick charge (about 20 min) process for extending the lifetime of these expensive battery packs*, obtaining *a more precise remaining charge estimation for displaying the remaining distance.*

1.3 Impact on Power Grid and Vehicle to Grid Technologies

In a very near future, where electric cars are being more and more widely adopted, electric power grids are supposed to sustain an increasing demand of energy. For this reason, power grids may face issues related to an overload scenario, where all EVs of a country are being plugged in for recharge. In this section some of the major contributions on this topic are presented.

In [6] Matteo Vasirani and Sascha Ossowski face the problem of the impact on power grid of simultaneous charging of many Plug-in Electric Vehicles (PEVs). In fact, this case *can cause power quality degradation, energy losses and overloads of the distribution substation, leading to overheating and ultimately equipment damage*. The authors define their proposed allocation policy, and analyze it from a gametheoretical point of view. The main idea is that the policy is inspired by lottery scheduling: the resource rights are represented by lottery tickets of equal unitary value, and the process with the winning ticket is granted the resource. Then, the amount of lottery tickets held by a process determines the probability of winning the lottery and therefore being granted the resource. This policy has been proven to be probabilistically fair in the long run. They finally *demonstrate that this allocation policy is capable of balancing allocative efficiency and fairness if PEVs coordinate to play the best equilibrium strategy.*

Shigaku Iwabuchi et al. in [7] face the same problem presented in [6], but in a different point of view. In their scenario, shopping centers are equipped with recharging stations for EVs, powered by fuel cells and solar power. They propose a Behavior Induction based Energy Management System (BIEM) in order to induce changes in the timing with which users charge their EVs. *BIEM affects the timing of the charging by recommending times for EV charging to the user. The times recommended to the user are selected for efficient use of power according to the circumstances of the energy facilities. If the user accepts the recommended charging time, there will be fewer occurrences of insufficient power and convenience will be improved.* They finally propose an algorithm for recharge timing recommendation to the user.

Shiyao Chen et al., in [8] face the issues of charging hundreds of EVs simultaneously. In fact, *charging facilities require properly designed pricing and scheduling that take into account the intermittency of renewable energy, the grid electricity cost, the arrival-departure characteristics, and customer price sensitivity.* They conduct simulation studies for the impact of the operations in two scenarios: monopoly and Bertrand duopoly competition. The result is that the operator can improve profitability if customers are flexible in time, and *the operating cost of charging facility has to be closely monitored and balanced with the pricing.*

In [11], A. Sheikhi et al. study a new Plug-in Hybrid Electric Vehicles (PHEVs) charging scheduling program that aims at optimizing customers charging cost. They consider the generation capacity limitations of a power grid and the dynamics of prices in the different time slots of a day. They propose an algorithm that calculates the near optimum charging schedule for all vehicle owners. The authors also show how the algorithm works in some simulations, and they state that, using this algorithm in a simulation where 70% of homes are equipped with PHEVs, *not only the PHEVs charging are shifted to the off-peak times, but also the load profile becomes considerably smooth.*

In [13], Calnan et al. focus on the impact of EVs on electricity generation in Ireland in 2025. Using a software package by Energy Exemplar named PLEXOS, they proceed with modeling the Irish electricity market. The authors, along with the information on the composition of the five generation portfolios received from the Irish system operator Eirgrid, undertake detailed market simulations in order to assess the impact of government targets for EVs on the generation costs, emissions, generation stack and the cost to load of this additional demand. The results show that *gas will be the dominant source of electricity generation to load EVs and that wind as an electricity source will experience a minor reduction in curtailment, with the least cost charging profile showing a more pronounced reduction.*

As the authors in [13], Borba et al. [14] also study different energy sources. The authors study the electric power system of Brazils Northeast region. With the installation of both wind power plants for 4.0 GW by 2013, and the construction of nuclear and run-of-the-river hydro-electric power plants, they are designing an appropriate modeling of the power system. This system has to take into account the integration of variable and unpredictable generation sources. They consider, in this case, electricity storage technologies, including the promotion of EVs and PHEVs. In this case, the authors study the possibility of using a governmental PHEV fleet as a way to increase the flexibility of the power system. As a result, they estimated that *a fleet of 500 thousand PHEVs in the Northeast region in 2015, and a further 1.5 million in 2030, could be recharged overnight for half the year to use the electricity surpluses of the wind farms planned for the region, thus avoiding the costs of modifying the electricity system.*

In [18], Abouzar Ghavami et al. study the impact of a large number of EVs charging simultaneously on the limited power capacity of the distribution feeders. The authors propose two algorithms to be implemented in a decentralized manner, in order to control the amount of power through each specific distribution feeder to avoid system overloads that may lead to breakdowns. They show that *both approaches converge to attain near-optimal load variance while ensuring that the feeders are not overloaded.*

Fjo De Riddera et al. [19] also study the potential of EVs in the Smart Grid. They propose a centralized charging schedule for EVs, that takes into account local and temporal flexibility and consumers' preferences. The algorithm operates locally on the vehicle, and uses information such as trajectory planning, parking duration and charging controllers to operate. In this manner, the consumers privacy is always insured, and the power constraints of car park are always met. While on board processors have to deal with the coordination algorithms, the parking managers need only to be concerned with the network congestion issues. In case the power constraints at the charging location are violated, vehicle owners are given an incentive to charge at other locations. The authors also propose a first application that focuses on controlling the power flows at the parking locations and on rescheduling the recharge of EVs, and a second one that takes also into account the imbalance costs. The trajectories are computed using an activity based model called FEATHERS. The authors finally simulate the usage of the algorithm, constructing charging schedules on a fleet of 200 EVs. *The charging schedules are constructed day-ahead, given a (time-varying) electricity price, and given a known trip schedule for the following day.*

In [20], Joosung Kang et al. study a concept of real-time scheduling, starting from a centralized approach, for charging EVs, in order to reduce the impacts on the power grids. The authors make some simulations of this concept, showing some advantages in comparison with existing valley-filling techniques, taking also into account timing constraints of EV owners. The proposed method also sets electricity price basing on preferences of EV owners, in order to encourage EVs to follow the schedules. The authors also state that, in a near future, it could be possible to extend the current centralized scheme into a decentralized one, in order to make charging schedules interactions with other system components such as home energy management systems.

As Joosung Kang et al. do in [20], in [21] the authors face the electric grid capacity issues in a near future, where EVs could represent a big problem when connected to the grid. They address the problem by considering a mathematical function that minimizes the system power losses, eg. a nonlinear optimization model.

From a different point of view it's interesting to describe the problem faced by Xiaomin Xi et al. in [22], i.e. the location of EV chargers in order to maximize their use by private EV owners. In the first phase, the authors determine where EV owners live, then they use linear programming to discover the optimal location and size of charging stations. The model has been applied to the central-Ohio region demonstrating that a combination of level-one and level-two chargers maximize the charging energy available, where level-one chargers are 110V and level-two are 220V.

The impact of EVs seen as power loss on the power grid is not the only type of impact that EVs could have. In fact, in [24] Jeffrey S. Marshall et al. study the problem related to the heat transfer around underground cables, like thermal degradation. The authors stated that, with just a 30% of EVs penetration, vehicle charging is found to rise the peak temperature of the cables' surfaces, *increase the daily variance in cable temperatures, and significantly decrease the estimated time to failure for cables with thermally sensitive insulation.*

As we can observe reading all the related researches, the electric grid is going to be a more and more crucial infrastructure for future energy management. It's supposed to react at changes, making use for example of cloud computing, and optimizing energy supply in all of the possible scenarios. One of these scenarios, widely discussed in literature, is the possibility of using power stored in EV's batteries to address Smart Grids' peak loads by injecting power into the grid: this is the so called Vehicle to Grid (V2G) scenario.

Willett Kempton and Jasna Tomic [16] face the V2G opportunities by discussing what type of vehicles are the most suitable for the V2G, and to what markets they can sell energy, focusing on capacity, cost and revenue of electricity coming from Electric-Driving Vehicles (EDVs). They move from the consideration that the electric grid does not have backup batteries, while EVs instead are potential backup batteries for the electric grid because they are used for an average of 4% of the vehicle's lifetime. There are 3 types of EDVs able to produce V2G, that are Battery Electric Vehicles (EVs), Plug-in hybrid EVs (PHEVs) and Fuel Cell Vehicles, and four types of power markets, distinguished by different control regimes. Then an analysis on eventual power capacity of V2G is conducted, based on 3 factors: *the current-carrying capacity of the wires and other circuitry connecting the vehicle through the building to the grid, (2) the energy stored in the vehicle divided by the time, and (3) the rated maximum power of the vehicles power electronics*. By estimating revenues and costs of V2G, the authors conclude that *V2G would improve the reliability and reduce the costs of the electric system*.

Lassila et al. [17], face the same topic taking into consideration the point of view of an electricity distribution company, and assessing the economic impact of V2G. They focus on the vehicle's discharging perspective, aiming at presenting a methodological framework that could help distribution system planners to estimate the preliminary feasibility of energy storages. Although the model consists of calculations and parameters that involve many assumptions and uncertainty, *the study shows the importance of understanding the correlation between the distribution network value, network capacity, and energy storage systems.* In this context it could be possible to cut the distribution fees charged to end-users with the large-scale adoption of EVs if the issue can be taken into account during the system planning. Finally, the authors state that *the shape of the base load curve and the peak operating time affects strongly to the feasibility of energy storages.*

Gallardo-Lozano et al. [15], discuss on EV's on-board battery chargers compatible with smart grids. Nowadays, currently used EV's battery chargers are high power and non linear devices, and they generate significant amount of current harmonics. In the future of smart grids, EVs are going to be always connected components, therefore their power quality impact has to be analyzed and optimized. The authors present, in this paper, a three phase on-board battery charger compatible with Smart Grids, and enabled to V2G operations. This battery charger is characterized by the ability of recharging the batteries during peak-off times, and delivering the energy back during peak times of electrical consume. The focus of the work is on the control strategy that enables the bi-directional operations. This control strategy *tries to fulfill the recent IEEE Standard 1459-2010, with the objective of maximizing the use/injection of Alternating Current (AC) power from/into the grid, and reducing the load harmonic factor and load unbalanced factor.*

In [23], M.A. Lopez et al. study the congestion management in a microgrid with high penetration of EVs. The authors formulate a model that takes into account both the technical and economic aspects of the integration of EVs in a power grid. Then, they use V2G to address congestion issues, and they propose an algorithm based on power distribution factors (DFs). DFs are used to determine the amount of energy that a specific EV should contribute to alleviate the congestion in a line. *The algorithm has proved to be effective if the congestion level is not very high.*

1.4 Battery Energy Management and Recharge Modalities

In Plug-in Electric Vehicles (PEVs) the battery pack is a critical energy source, and it currently represents the performance bottleneck. In fact, daily driving involves complex vehicle operations, and for this reason a Battery Management System (BMS) is required. The BMS' aims at ensuring safe and reliable operations on batteries, and it provides precise information about the battery, such as battery's SoC. In this section will be analyzed all contributions regarding Battery Energy Management and the Recharging modes available or still in development.

In Hamid Khayyam et al. paper [10], like in [8] and [11], the authors face the problem of charging hundreds of EVs simultaneously, but from a different point of view. They *propose a new intelligent battery energy management and control scheduling service charging that uses Cloud computing networks.* The authors make some experimental analysis of the proposed scheduling service and compare them to a traditional scheduling service, through simulations. They derive that *the Cloud computing intelligent vehicle-to-grid (V2G) scheduling service offers the computational scalability required to make the decisions necessary to manage V2G systems as the number of PEVs and intelligent charging devices increases.* They show that, with the proposed methods, the interactions between PEVs and parking lots and grid are reduced, and the load demand can be predicted.

As previously mentioned, battery packs are the core energy source for an EV. For this reason, the main issue related to energy supply is when and where to recharge vehicles' batteries. In [12], Shin et al. present the design and implementation of a wireless power transfer system for moving EVs. The authors are convinced about the possibility of supplying energy to a moving EV in wireless mode. They design and test their idea using a wireless power system that uses an inductive coupling. As a result, *the system provides 100-kW power with over 80% power transfer efficiency at 26-cm air gap*, and shows that wireless power transfer systems are a feasible way of recharge vehicles' batteries.

1.5 Connected Vehicles

EVs are being more and more connected to the net, in order to transmit various information such as system failures, location information, energy management, charging station location and vehicle performance.

A contribution on this topic is made by Ovidiu Vermesan et al. [9]. The authors discuss about the trend and opportunities deriving from the introduction of connected EVs, and future Energy Management solutions. For the authors, some of the possible uses of the Internet of Energy are power distribution, energy storage, grid monitoring and communication. There are also 4 different generation of EVs, distinguished by performance and complexity. They define *The Internet of Energy concept as a dynamic network infrastructure based on standard and interoperable communication protocols that interconnect the energy network with the Internet allowing units of energy* *(locally generated, stored, and forwarded) to be dispatched when and where it is needed.*

1.6 M-Atlas: Mining Human Mobility

M-Atlas is a querying and mining language and system centered on the concept of trajectory. It is an important tile of this entire work, because it enables the trajectory processing inside the simulation, and it is an excellent instrument to understand trajectories' and users' behavior. As described in [27], *the mobility knowledge discovery process can be specified by M-Atlas queries that realize data transformations, data-driven estimation of the parameters of the mining methods, the quality assessment of the obtained results, the quantitative and visual exploration of the discovered behavioral patterns and models, the composition of mined patterns, models and data with further analyses and mining, and the incremental mining strategies to address scalability.*

M-Atlas has mechanisms for mining trajectory patterns and models that, in turn, can be stored and queried, and supports various functionalities such as:

- trajectory data creation, storage and query through spatio temporal primitives;
- trajectory models and patterns representing collective behavior extraction using trajectory mining algorithms;
- representation and storage of such patterns and models in order to be re-used or combined.

All these functions are combined through an innovative Data Mining Query Language (DMQL). This language can be used to express the whole knowledge discovery process as a sequence of queries to be submitted to the system. M-Atlas supports three types of data: purely spatial data, purely temporal data, and moving points or trajectories. Plus, the M-Atlas system integrates a set of analytical and data mining tools such as the construction of Origins-Destinations Matrix, the construction of georeferenced density maps according to different measures, extracting of T-Patterns, T-Clustering, T-Flocks and T-Flows.

The O/D Matrix can be used to discover the common exits of a city, and then to extract the set of trajectories which is part of a selected flow. T-Clustering can be used to group together similar trajectories in order to discover common behaviors using different methods such as *Route Similarity*. A T-Pattern is a concise description of frequent behaviors, in terms of both space and time, while T-Flocks represent a spatio-temporal coincidence of a group of moving points. This spatio-temporal coincidence defines a common behavior of the people which move together for a certain time interval. Finally, a T-Flow represents a flow of trajectories moving from a region to another one.

In order to start using M-Atlas, a starting dataset of raw GPS data must be available. Then, a trajectory construction function has to be used, in order to pass from raw GPS data to trajectory data. Many parameters can be supplied to the trajectory reconstruction function, and some are essential for the success of the task. They are the minimum time between two consequent trajectories (MAX_TIME_GAP), and the minimum space between two consequent trajectories (MIN_SPACE_GAP). The trajectory construction function takes as input a dataset containing the following fields:

- *id* of the user;
- *lat* value of the current point;
- *lon* value of the current point;
- *timestamp* value of the current point.

and finally gives in output a dataset containing the main following fields:

- *uid*, the user id referring to the id value in GPS_raw;
- *tid*, the trajectory identifier, enumerated from the first of the user to the last, in order of time;
- *the_traj*, the geometry object that contains the information related to every point, made of x,y and z coordinates, which refer to latitude, longitude and timestamp;
- *time_start*, the timestamp of the first point of the trajectory;
- *id*, the identifier of the row in the entire table.

A trajectory, reconstructed by a mining process, is useful for understanding personal mobility and for having a first look on a map of how it behaves. In Figure 1.1 is shown an example for a single user's trajectories.

The authors then conduct an analysis on two massive real life GPS data sets, one containing $\approx$ 17000 vehicles tracked in Milan in one week (April 1st through April 7, 2007) and consisting of a total of $\approx$ 2 Million observations, and another one containing $\approx$ 40000 vehicles tracked in Pisa in 5 weeks (from June 14th through July 18, 2011) and consisting of a total of $\approx$ 20 Million observations. Each dataset is in raw GPS format, thus being in the format of a quadruple with values of id, lat, lon and timestamp. They use

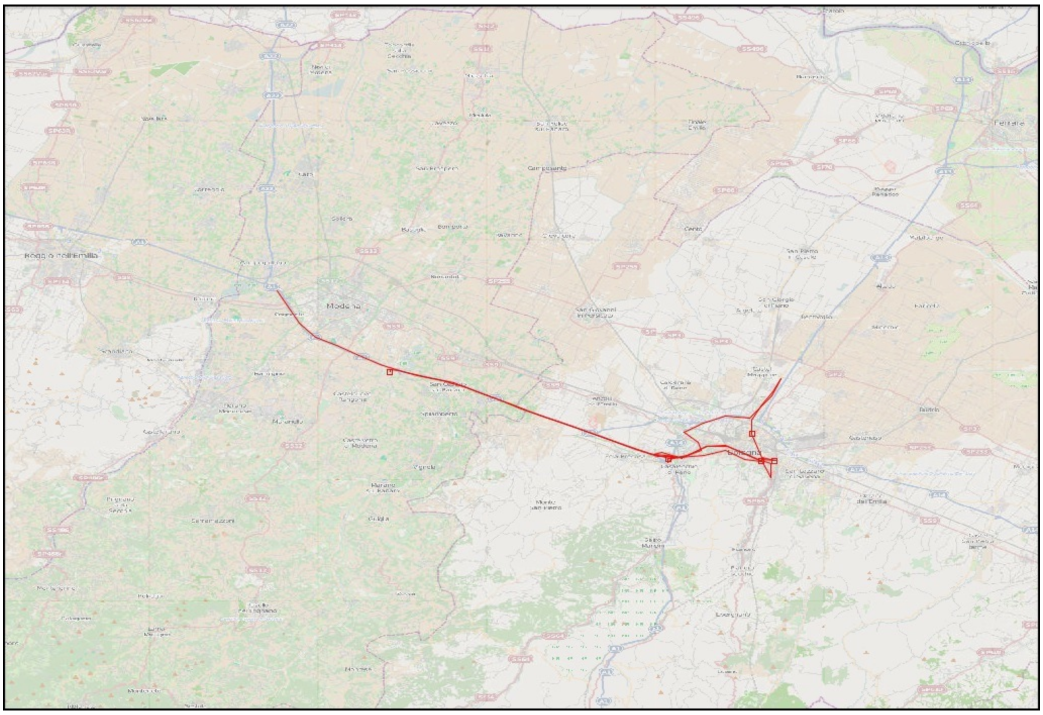

Figure 1.1: User trajectories view on the map

the trajectory reconstruction function to chain together all the observations of the same car id over the entire observation period in increasing temporal order into a global trajectory of car id., and then split the global trajectory into several sub-trajectories, corresponding to trips or travels, by using a cut-off threshold of 30 min. The result is $\approx$ 200000 travels for Milan and $\approx$ 1500000 travels for Pisa.

Their analysis comprehend the movement distribution in the city, which is an analysis on the number of moving vehicles at different hours of the day. Figure 1.2 shows this analysis, and it is important to notice that, in working days, people tend to move together in some precise hours, probably the ones in which commuters move for going to work. The authors also compare this plot with the one obtained from a survey by the Milano municipality on a period of five years, showing *not only that the results are coherent, but also that the survey distribution is known to underestimate the movements where the mismatch occurs*, this because *GPS data also capture nonsystematic movements, while survey data do not, as interviewed people tend not to report their occasional mobility, such as going to the dentist or visiting a friend.*

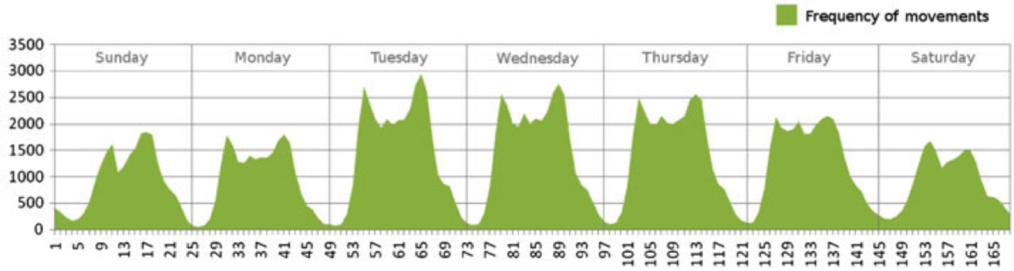

Figure 1.2: Movement distribution of the entire week in Milan - Figure by [27]

They show then a presence distribution, i.e. the number of people stationary in the same places in every hour of the day. Comparing this distribution to the one of the survey, they demonstrate that the two distributions match in most locations.

The dataset object of their studies is also used for basic statistic studies, such as lengths of trips, duration of trips, correlation of length and speed of trips, the radius of gyration (the average distance of a vehicle from its most likely location) and the density of vehicles in space and time. By analyzing trip length and duration they realize that mobility is a complex phenomenon that cannot be characterized by any simple notion of average behavior. The analysis on the radius of gyration shows how vehicle insist on their preferred locations. By computing the radius of gyration, it is possible to compute, for every vehicle, its most likely locations and the general law of the power of its attraction. After all analysis, they conclude that *there is a huge complexity represented in the data, a wide variability of individual mobility behaviors that cannot be fully understood in its diversity by looking only at macroscopic, global measures and laws.* Their goal is then to try to discover different subgroups of vehicles and travels characterized by some common movement behavior.

So, they start using M-Atlas in order to master the complexity of the knowledge discovery process in its more critical issues, such as the definition of complex interactive and iterative analysis, the estimation of algorithm parameters, and the validation of the models. They begin by characterizing the main flows from the city center toward the suburbs. They use the administrative borders of Milan as input for the T-O/D Matrix model constructor, obtaining a high-level description of the flows between each pair of regions. With the help of a visual interface, the analyst can interact with the model. They firstly focus on the T-Flows leaving the city of Milan toward the north-east suburbs, obtaining the visualization in Figure 1.3 (left). Then, a clustering algorithm is applied, in order to find similar routes, and the result is visible in Figure 1.3 (right), where different colors define each cluster. The function used here to cluster the trajectories is the *route similarity* distance function.

Other analysis are then conducted by the authors. One example is the accessibility to key mobility attractors, like the top accessed parking lots (where Linate airport parking is the top accessed one of the city). Another example is the identification of extraordinary events that could have large impacts on mobility, such as concerts and sport competitions: here, the event location is the destination of many individual trips and it is a small area.

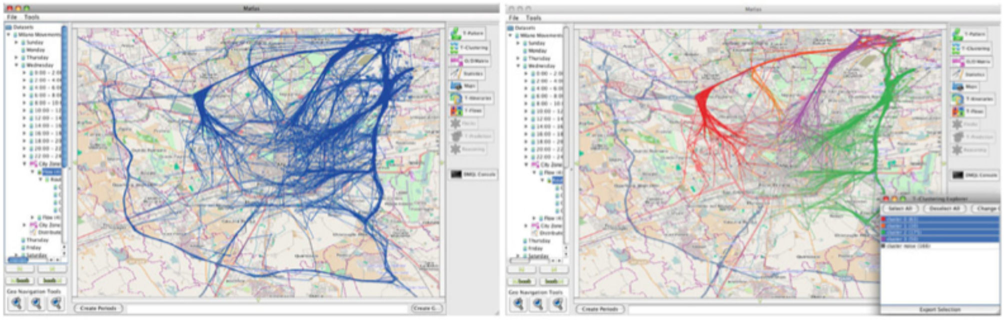

Figure 1.3: T-Flows in Milan - Figure by [27]

Then, after the event, that location turns into the starting point for many return trips. *Even if not known a priori, big events can be easily detected by localizing exceptionally high concentrations of presence in specific areas at specific time intervals.* Other analysis comprehend mobility predictions, that are very useful to predict traffic congestions, and traffic jams detection.

Chapter 2

Discharging/charging model and Scenarios

This thesis aims at the individuation of a key parameter that could lead us to understand when, and under which conditions, it is possible to change from a gasoline powered vehicle to an electric one. This key parameter has been devised in the so called electrifiability rate: it represents the percentage of a single users' journeys that is possible to cover with an EV. This is achieved by simulating the consumption of an EV on real every-day users' trajectories. The starting dataset will be introduced in Section 3.1.

As it is possible to imagine, every vehicle has to deal with more or less complicated physical forces during every day usage. In order to simulate a generic vehicle's consumption, these forces are to be considered and modeled. Moreover, as we are going to study the consumption of an EV, other additional parameters are to be kept into consideration also. All these forces and parameters are going to represent the core model of the simulation, and will be presented further on.

In modeling this simulated reality, also some usage scenarios will be devised and analyzed. These scenario represent a simplification of the reality, useful for better setting the simulation in its parameters, and are a good starting point to evaluate the basic discharging and charging model made in this work. They will also be presented in this Chapter.

2.1 EV discharging/charging model

In this Section the discharge method and the Electric Vehicle Specs used for the simulation will be introduced. Both model and parameters presented here have been discussed in a report by Jesùs Fraile-Ardanuy et al. [28]. The first scenario is the basic one, and all the simulation consider it as the basis to start from: it includes the basic algorithm developed for calculating the consumption, and also if we consider other scenarios, the core of this algorithm, albeit modified, stands still under the hood. The second scenario is also of primary importance, because the algorithm developed for its calculations is also fundamental for all the others, since they all consider the recharging phase in them. Therefore, the algorithms used for the simulations are then presented.

2.1.1 Discharge Model and Nissan Leaf Parameters

An EV is a complicated dynamic system, composed of many subsystems that work all together, like electric motor, battery and so on. This specifications and the formulas will be used for simulating the vehicles behaving on millions of trajectories, thereby they were simplified as much as possible, keeping though a high correlation with the real vehicle specs supplied by the manufacturer of the vehicle used for the simulation.

As we can imagine, during the real driving the vehicle has to deal with forces like gravity and friction forces like wind, tires rolling resistance and internal moving parts. Plus, in an EV we have to keep into consideration all the components consumptions and efficiencies. All these forces and components' consumptions have to be considered in calculating the EVs overall consumption. The model presented here is to be considered valid for every type of vehicle considered. In this way every vehicle can be used for consumption calculation simply changing the vehicle's parameters, that we present here after the general forces explanation.

Here is presented the Vehicle model.
The forces that a general vehicle has to deal with are the following:

- Rolling resistance:

$$F_{rr} = R(M_{car} + M_d)g\cos\alpha \tag{2.1}$$

- Aerodynamic Drag:

$$F_a = \frac{1}{2}AC_d\rho v^2 \tag{2.2}$$

- Gravity (vehicle's weight component):

$$F_{hc} = (M_{car} + M_d)g\sin\alpha \tag{2.3}$$

- Inertial force:

$$F_{la} = 1.05(M_{car} + M_d)a \tag{2.4}$$

where:
$R[-]$ is the tire rolling resistance coefficient;
$M_{car}[kg]$ is the mass of the vehicle;
$M_d[kg]$ is the mass of the driver;
$g = 9.81[m/s^2]$ is the gravity acceleration;

$\alpha[rad]$ is the angle of the driving surface;
$A[m^2]$ is the front area of the vehicle;
$C_d[-]$ is the aerodynamic drag coefficient;
$\rho = 1.2041[kg/m^3]$ is the air density of dry air at 20°C;
$v[m/s]$ is the speed of the vehicle;
$a[m/s^2]$ is the acceleration of the vehicle;

The inertial force has two components: the force required to give linear acceleration and the one required to give rotational acceleration to the traction motor. Since the motor's moment of inertia is difficult to know, it is reasonable to increase the vehicle's mass by 5% in 2.4.

The total Traction Force can be expressed as:

$$F_{te} = F_{rr} + F_a + F_{hc} + F_{la} \tag{2.5}$$

A graphic representation can be seen in Figure 2.1[1].

Then, the mechanical tractive power is the product of tractive force and the average speed of the vehicle. It depends on the power of the engine and the efficiency of the transmission, and it is:

$$P_{te} = F_{te} v \tag{2.6}$$

This power is then transferred to the wheel, and assuming a constant gear efficiency η_{gear} , the power that enters in the gear system block is:

$$P_{mot\_out} = \frac{P_{te}}{\eta_{gear}} \tag{2.7}$$

Is then considered the electric machine efficiency, η_{mot} , and the power that enters in the electric machine block is:

[1] Image from [28]

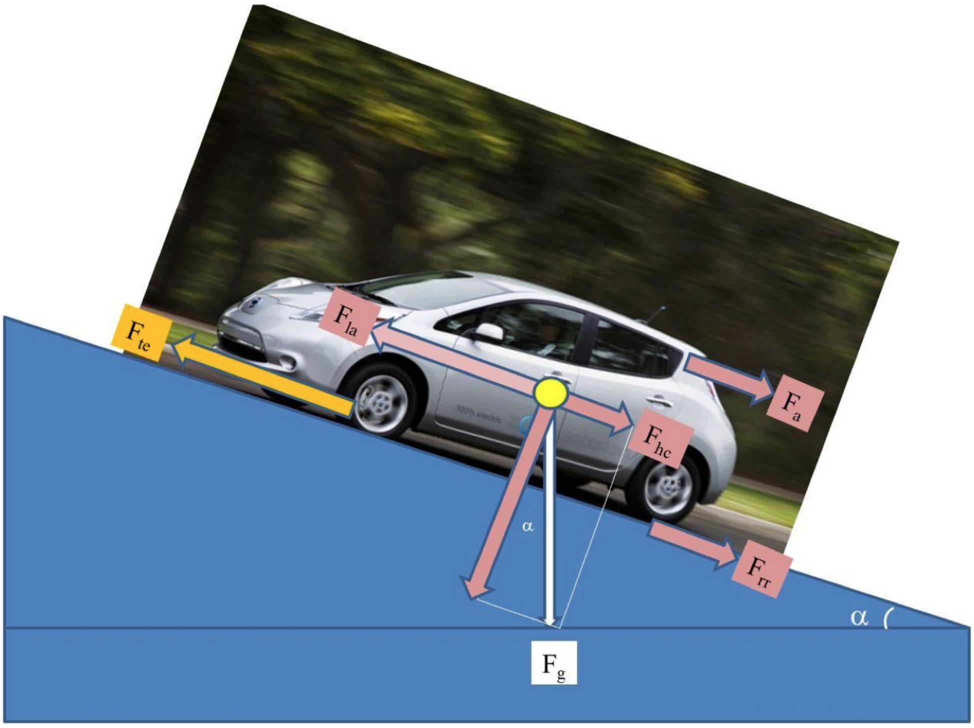

Figure 2.1: Nissan Leaf Traction Forces, Technical Report DATASIM

$$P_{mot\_in} = \frac{P_{mot\_out}}{\eta_{mot}} \tag{2.8}$$

Then, the auxiliary power P_{aux} , that represents electric loads such as lights, wipers, horn, indicators, radio, air conditioning or heating, etc. is considered, and the total power required from the battery is:

$$P_{bat} = P_{mot\_in} + P_{aux} \tag{2.9}$$

During breaking, EVs convert a fraction of kinetic energy that flows from the wheels to the motor, and then to the battery pack in form of electric energy. The fraction is represented here as regeneration factor R_{gen_ratio} . Here the mechanical power P_{te} is negative, and the power regenerated is:

$$P_{te\_reg} = R_{gen\_ratio} P_{te} \quad (2.10)$$

This power flows back through the transmission:

$$P_{mot\_out} = \eta_{gear} P_{te\_reg} \quad (2.11)$$

and, again, through the electric machine:

$$P_{mot\_in} = \eta_{mot} P_{mot\_out} \quad (2.12)$$

Finally, the auxiliary power is added to the motor power to give the total power required from the battery. Note that here $P_{mot\_in} < 0$.

$$P_{bat} = P_{mot\_in} + P_{aux} \quad (2.13)$$

The EV used for this simulation is the 2012 Nissan Leaf version (Figure 2.2). The vehicle parameters of the Nissan Leaf are then presented in Table 2.1, along with the car's efficiencies in Table 2.2 and are useful to fully simulate this vehicle's behavior on the trajectories.

Cross sectional area	$2.27m^2$
Curb weight	1521 kg
Driver weight	90 kg
Cd (drag coefficient)	0.29
μ (coefficient of rolling resistance)	0.012
Regeneration ratio	0.25 normal mode/0.35 ECO mode
Battery storage capacity	24kW
Low Battery Limit	8-10%

Table 2.1: Electric vehicle parameters

Figure 2.2: 2012 Nissan Leaf

Gear efficiency	0.95
Electric Machine & Power Elect. Efficiency	0.98
Charging battery efficiency	0.95
Discharging battery efficiency	0.98
Self-discharging ratio	3% monthly (0 in the simulations)

Table 2.2: Electric vehicle efficiencies

Although the low battery limit is indicated in 8-10%, all the simulations presented in this work are set to completely use the battery charge, so the vehicle stops when reaches 0% battery SoC, instead of 8-10% battery SoC. This to evaluate the total battery capacity of an EV, useful for understanding its full range. Moreover, is to be considered that in real life, an EV has a little power loss also when stopped, this because of always plugged battery. This situation has not been modeled here, and it has been considered a power loss of 0 in the simulation. What said for the self discharging ratio is also valid for P_{aux}: in fact, the auxiliary power required to use lights, horns, electric glasses, air conditioning/heating and radio are assumed to be 0 in this simulations, for simplification purposes.

2.2 EV Usage Scenarios

The entire project consists of various usage scenarios: they are meant to represent various levels of constraints applied to the model. These constraints stand for the possible limitations due to state's funds, infrastructures and users' behavior: for example, such limitations may be the lack of public charging stations, the impossibility of recharge the vehicle due to the lack of presence of a minimum stopping time by the user, or the low voltage recharge of the vehicle. All these conditions have been analyzed and classified into those usage scenarios, covering then all the possibilities that could be met in real life. In this section will be presented the different scenarios that involve the usage of EVs.

This scenarios are to be considered a simplification of the reality, in which the main idea is to check if it is possible to change users' vehicles today without making the user change his behaviors. These scenarios are then a good starting point to evaluate the basic discharging and charging model made in this work.

2.2.1 Scenario 1: charging at home

This is the first and very basic scenario. Here, EVs can only charge at home with low voltage chargers, since there are no other charging stations available for the user. This can represent the common users' behavior: in fact, users start their vehicles at home, and, after their whole day, return home. The simulation applied here aims at identify if the vehicle is able to return back home with some charge available or not, without having to recharge batteries in the middle of the day. An EV's battery pack, at the time of writing, takes a long time to recharge, and needs specific predefined recharge spots. This leads to the formulation of two constraints: the location of the charging spots (in this case they are the users' houses) and the hours at which every EV can recharge. A third constraint of the current scenario can derive from the first one: because of the lack of infrastructures, EVs home chargers can only charge at the speed of 3 kW/h. This scenario comes in play in situations where there are no investments on the power grid, and where there is no possibility to recharge the vehicles in places other than home, i.e. public charging stations.

Figure 2.3 shows a graph that explains the possible simulation. In the first path, the vehicle moves at a certain speed, and the battery charge level decreases accordingly. After a stop, during which the battery charge does not decrease, the vehicle restarts. During Path 2 the vehicle increases its speed many times causing a more rapid decrease of the battery charge than the previous path.

This model applies basically to all scenarios, because the main calculations on consumption are made here. So, the next scenarios will be designed starting from this one and enriching it with the specific constraints required.

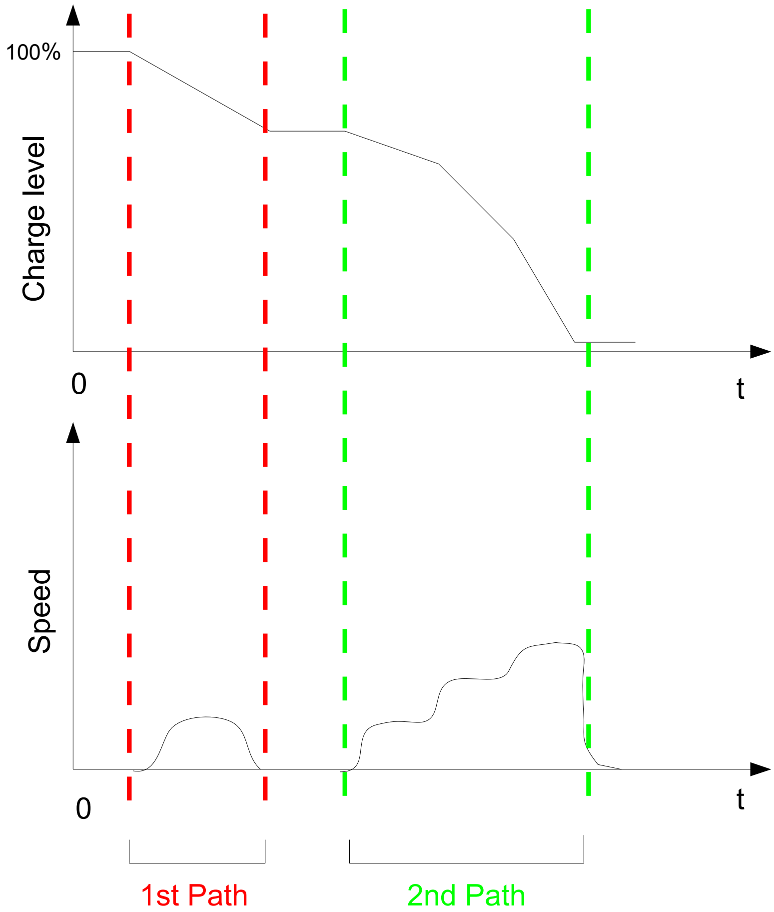

Figure 2.3: Scenario 1

2.2.2 Scenario 2: charging everywhere

To derive the second scenario, the first one is enriched with some new constraints. In fact, here vehicles not only charge batteries at home during the night, but can also do it during the day everywhere. Moreover, a stopping time constraint is introduced here, and in our experiments, it is used a 2 hours minimum stop constraint that, if true, makes the vehicle recharge the batteries. In this scenario, the availability of funds for the construction of infrastructures like public charging spots is considered. In fact, it is considered feasible the eventual recharge of a vehicle wherever it stops, meaning that charging spots are available everywhere.

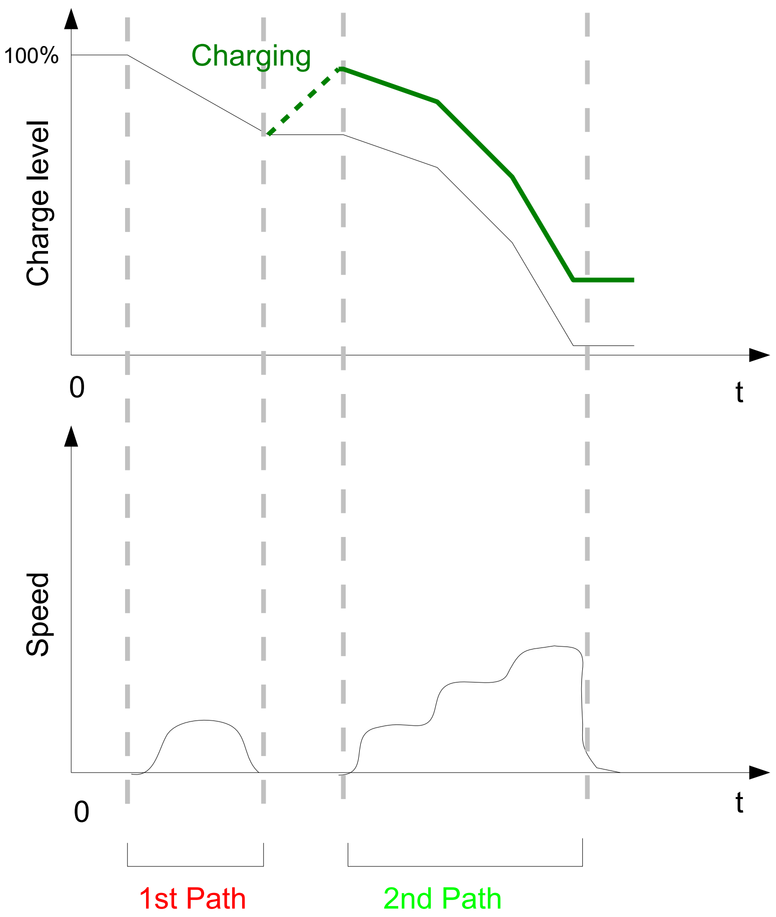

Figure 2.4: Scenario 2

Figure 2.4 shows how the simulation could work. In addition to the first scenario, it is possible to recharge batteries if the stopping time is at least greater than a certain threshold. The threshold used in this scenario, coming from our experiments, is 2 hours. It means that there is the possibility to increase the vehicle's autonomy in presence of long stops. An example of this type may be when people stop the vehicle during work time, that is a long stop period in which EVs can recharge.

2.2.3 Other possible scenarios

The first two scenarios can represent a good starting point for more complex analysis. In fact, by adding constraints to the studied models, it is possible to derive some other scenarios that may better represent real life conditions, like the lack of ever-present infrastructures, fast charging possibilities or Vehicle to Grid EV's usage hypothesis. All these future works are presented in Chapter 4.

2.3 Data Sampling and Interpolation

Along with the main model creation and the main scenarios representation, also an important consideration on the available data is mandatory.

In real life, the consumption of a vehicle is strictly connected to various factors such as slopes, speed, accelerations etc. In few words, they all can be reduced to two main factors: the driving style of the user, and the morphology of the streets the users travels on. Those two elements have to be deeply analyzed, because it is important to check if they can be easily represented by actual data or not.

Let's start analyzing the morphology of the street. In this case, the main factor to consider is the elevation: in order to be realistic, a consumption simulation should take into account the variations of elevations for every street, and for every second. So, the GPS data we start from have to be sampled in the shortest time possible. In fact, if the data sampling is low, some fluctuations in elevations could not be considered, thus changing the final consumption result.

Regarding the driving style of the user, we should consider that some harder accelerations may have a greater impact on final consumption. Again, this can be well represented in the data if the sampling is high enough.

It is important, then, to have an high sampled data to start from. If this condition is not met, it may be a solution to interpolate the data available in order to obtain the required information, for example by having a 1 second interpolation if the original data has more than one minute time lags between two consecutive points.

In this work, because of low sampled data, firstly a simulation on original data has been conducted, and then an interpolation of the data has been taken into account for a second running of the simulation, in order to take into account elevation changes between distant points. It is to be considered though that constant acceleration has been used for this interpolation, thus not considering the driving style factor.

2.4 Scenarios 1 and 2 Algorithms

In this section the main algorithms used for calculating the consumption are presented.

In the first Scenario, as previously mentioned, the discharge process of the vehicle is simulated. EVs start with full battery charge with the first trajectory of the day. Here, all the users journeys are used as basis to run the simulation: the goal is to find and indicator that tells us that an EV can be used by the user instead of a gasoline powered engine. The indicator used is then found to be the total consumption of the vehicle at the end of the day: if the vehicle's total consumption is lower than 100% of battery charge, that user's daily journeys pass the simulation; if, on the contrary, the total consumption of the vehicle at the end of the day is greater than 100%, that user's daily journeys fail the simulation.

An observation is to be made: the actual sampling of the GPS data is not as high as we could expect. For this reason, we initially run a simulation on the actual sampled GPS data, in order to simulating the consumption on just the available data. The consumption is calculated using the formulas provided above in Section 2.1.1.

Here is presented the algorithm in pseudo-code, in order to facilitate the comprehension.

Algorithm 1: First consumption algorithm, no interpolation

```
for every user, every day do
    for every trajectory do
        Get the trajectory and extract its points;
        for every point of the trajectory do
            Get the elevation, angle of road for incremental elevation,
            speed, acceleration and distance from previous point;
            Calculate consumption;
        Write consumption for the trajectory on DB;
```

After the first run, because of the data sampling, a consideration has been made: between every point, the actual changes in elevation could have been not considered, thus changing the final results. For this reason, the algorithm is modified to apply 1 second interpolation on GPS data, in order to consider the elevation fluctuations between two points of the trajectory (deeply analyzed in Section 3.3). An observation is mandatory: the interpolation made here uses constant accelerations, this for consider only the elevation change. In fact, in real world, when we drive a car, we firstly have a big acceleration, and then we maintain a more or less constant speed until we brake. This type of precision is not achieved yet in these simulations, leaving then this implementation for future works.

The algorithm is shown in Algorithm 2.

In Algorithm 3 the recharge algorithm is presented, which is the one at the basis of the computation.

Algorithm 2: Second consumption algorithm, **with** interpolation

---

```
for every user, every day do
    for every trajectory do
        Get the trajectory and scan it;
        for every point of the trajectory do
            Get the time between current point and the next one;
            Get 1 second points;
            for every point in the interval do
                Get the elevation, angle of road for incremental
                elevation, speed, acceleration and distance from
                previous point;
                Calculate consumption;
        Write consumption for the trajectory on DB;
```

Algorithm 3: Recharge algorithm

for *every user, every day* **do**
    for *every trajectory* **do**
        Get the trajectory time_start and time_end;
        Calculate recharge if stop time > 2 hours
        Write eventual recharge and battery SoC at the end of the trajectory on DB;

Chapter 3

Experiments and Implementation

In this chapter will be presented the core work of this thesis, which is the experimental one and implementation of the simulation. But before starting with the explanation of the algorithms and solutions used here, a first presentation of the starting dataset used along with some previous data understanding and analysis is to be done.

3.1 Understanding GPS Mobility Data

Thanks to the collaboration with *Octo Telematics Italia S.r.l.*, CNR, the Italian National Research Institute, obtained a database containing the GPS data of $\approx$ 160.000 vehicles that stayed, or at least passed by Tuscany in the month of June 2011. The owners of these cars are subscribers of a pay-as-you-drive car insurance contract, under which the tracked trajectories of each vehicle are periodically sent (through the GSM network) to a central server for anti-fraud and anti-theft purposes.

	id numeric	point geometry	ts timestamp with time zone	vel numeric	heading numeric	quality numeric	status numeric	delta numeric
1	84	0101000	2011-05-07 06:28:22	0	0	1	0	0
2	84	0101000	2011-05-07 06:30:25	120	156	3	1	2117
3	84	0101000	2011-05-07 06:31:21	140	156	3	1	2148
4	84	0101000	2011-05-07 06:32:11	148	120	3	1	2099
5	84	0101000	2011-05-07 06:32:58	128	120	3	1	2017
6	84	0101000	2011-05-07 06:33:54	146	120	3	1	2157
7	84	0101000	2011-05-07 06:34:46	144	120	3	1	2055
8	84	0101000	2011-05-07 06:35:37	138	110	3	1	2037
9	84	0101000	2011-05-07 06:36:34	146	106	3	1	2247
10	84	0101000	2011-05-07 06:37:29	144	106	3	1	2167

Figure 3.1: GPS_raw table sample

The GPS data sampling is between one and three minutes on average. It means that every point is recorded in such temporal window, and then it can be very close to the previous one, or even very distant, depending on the speed of the vehicle. The overall work has been made using PostgreSQL as Database Software, with the PostGIS Extension, which provides spatial objects for the PostgreSQL database. The starting table contains all basic GPS information. Every row contains the latitude and longitude, the timestamp of sampling, and the id of the vehicle. A little sample of the raw GPS data table is shown in Figure 3.1.

In this representation there is not a sort of division, or a grouping of points, that indicates the single trajectory. This information has to be discovered, and it is done, in this case, by using M-Atlas software to reconstruct the users' trajectories (as seen in Section 1.6). In this case, the trajectory reconstruction function was previously used to derive the trajectories table, that is the starting dataset of this work. The parameters used here are 20 minutes for the MAX_TIME_GAP and 50 meters for the MIN_SPACE_GAP (for the explanation of these parameters, see Section 1.6).

The resulting table has 7 millions rows, and consists of these, important, columns:

- *uid*, the user id referring to the id value in GPS_raw;
- *tid*, the trajectory identifier, enumerated from the first of the user to the last, in order of time;
- *the_traj*, the geometry object that contains the information related to every point, made of x,y and z coordinates, which refer to latitude, longitude and timestamp;
- *n_points*, the number of points of the trajectory;
- *length* of the trajectory, in meters;
- *duration* of the trajectory, in seconds;
- *time_start*, the timestamp of the first point of the trajectory;
- *id*, the identifier of the row in the entire table.

Figure 3.2 shows a sample of the table. This table is the starting point of every consumption calculation program, as it contains all the basic informations of every journey needed: the position of the points, the time between them and the timestamps of the first and the last point.

3.1.1 Preliminary Explorations

For this work, trajectories will be used to simulate, on each trip of the user, the battery consumption based on many factors, such as vehicle mass, speed, acceleration and elevation. Before the real consumption simulation, some previous analysis are presented, in order to better understand the population of the dataset.

	uid numeric	tid numeric	the_traj geometry	n_points numeric	length numeric	duration numeric	time_start timestamp without time zone	id integer
1	84	6	0102000	3	1241.4	872	2011-05-09 10:14:35	6
2	84	7	0102000	5	4702.9	1646	2011-05-09 13:04:02	7
3	84	1	0102000	21	40592.	1509	2011-05-07 06:28:22	1
4	84	2	0102000	6	8195.4	1657	2011-05-07 07:33:18	2
5	84	3	0102000	24	36489.	5311	2011-05-09 06:46:53	3
6	84	4	0102000	3	3928.0	1143	2011-05-09 08:37:32	4
7	84	5	0102000	2	0.0	0	2011-05-09 09:44:43	5
8	84	8	0102000	19	34168.	2028	2011-05-10 09:55:17	8

Figure 3.2: Resulting table sample

The following analysis are a "picture" of the dataset that makes easy understanding, for example, the distribution of the users between the various cities, the kilometers driven for each city or the difference between weekends and working days in driven kilometers. These analysis are then a starting point for measuring some previous statistical data, that in most cases is being the only analysis conducted. Figure 3.3 shows the cumulative distribution of driven kms for various cities of Tuscany. As it is possible to see in the picture, the cities with more users in absolute that have short trips are Florence and Prato, both with a big gap between them and the others. Then, in the right side of the plot is possible to see that the difference between the cities decreases. At first sight we can observe that these two cities, unlike the others, have a greater number of vehicles in comparison to the smaller cities.

Another interesting analysis is about Working Days w.r.t. Weekends average daily trips. In fact, one could expect that in weekends people go somewhere else than in the city, or at least drive more than in working days. This is not the case, because in this sample, people, in weekends, at least drive as long as in working days. The result of the analysis can be seen in Figure 3.4.

This result may lead us to think that people drive less in the weekends,

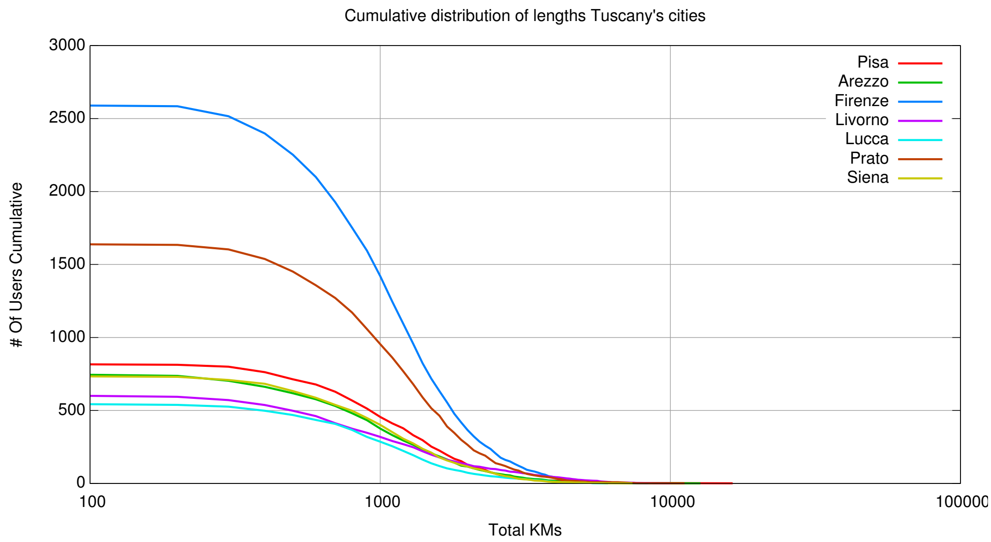

Figure 3.3: Cumulative distribution of driven kilometers of Tuscany's cities

but we should not generalize. Actually, it only means that in the month of June 2011 people drove less in weekends than in week days (that means we should have more data to generalize).

Having the target of the comprehension of the rate of electrifiability of urban mobility, another analysis is conducted on the possible rate of electrifiability that this work is going to assess. Even if it is a very initial analysis, made without taking into consideration the multitude of parameters that affect the EVs consumption, this analysis can explain what we are going to see further on. It is based on the length of the trips, and in the specific, only trips under the 100 km threshold were kept into consideration, thus making this one a very simplistic analysis. The plot is shown in Figure 3.5.

In the picture, on the X axis is shown the percentage of the trips under 100 km, where 100 indicates all the trips (always under 100 km); on the left Y axis is shown the number of users, and on the right Y axis is shown the

Figure 3.4: Working days vs weekends

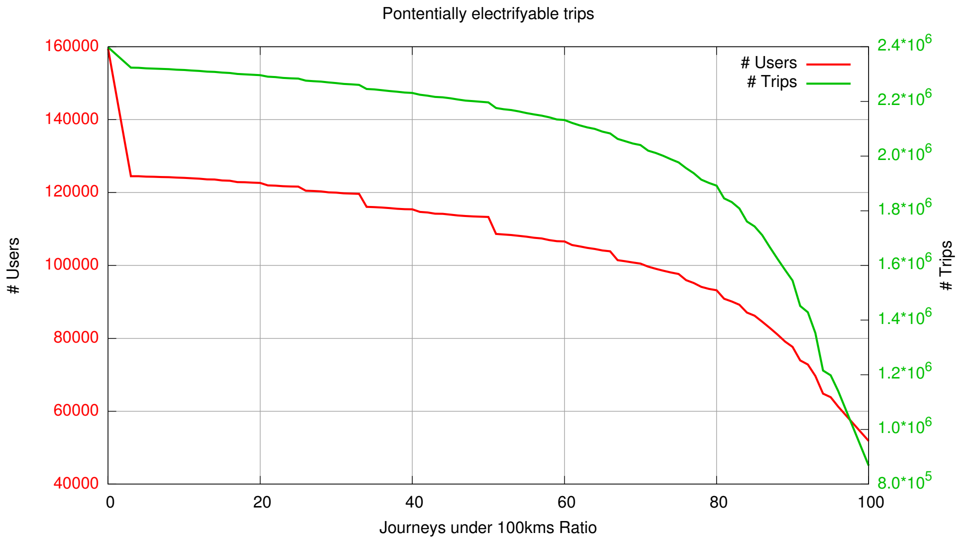

Figure 3.5: Potentially Electrifiable trips

number of trips. For example, the 100% of the trips under 100 km (X axis) is a total of ≈ 0.9 Millions trajectories (right Y axis) and is being performed by more than 50.000 people (left Y axis).

This last analysis relies on the hypothesis that, on average, an EV runs for 100 km before the battery is completely discharged. So, if the hypothesis is correct, ≈50.000 of users could benefit of an EV as their only vehicle. This value represents about the 31% of the population. Further on, with the simulation of an EV consumption on all these trajectories, this result will be validated.

All these analysis are only a first sight on the problem, and can not tell us relevant informations on possible electrifiability rates, as they not consider may factors, such as, for example, the terrain effects on consumption, but can, at least, improve our comprehension of the next part of experiments.

3.2 Java Implementation

The programming language used to create and run the simulation is Java 7, with the libraries needed for querying the PostgreSQL database where the actual starting table resides.

The program consists of more classes. The main class can be represented by the algorithms described in Section 2.4. An example of the algorithm is shown.

Algorithm 4: First consumption algorithm, no interpolation

for *every user, every day* **do**
    for *every trajectory* **do**
        Get the trajectory and extract its points;
        for *every point of the trajectory* **do**
            Get the elevation, angle of road for incremental elevation, speed, acceleration and distance from previous point;
            Calculate consumption;
        Write consumption for the trajectory on DB;

It connects to the DB and performs all the operations described in the algorithm, except for two of them: the elevation extraction and the consumption calculations, both operations made in other 2 classes. These classes are the elevation class and the consumption class, that contains all the formulas and vehicle's specs described in Chapter 2 and calculates the consumption referred to the parameters provided to it (that are distance from previous point, time between current and previous point, speed, acceleration and angle of the road, as seen in algorithms described in Section 2.4).

When a trajectory gets analyzed by the simulator, its single points are extracted in their temporal order. In this way, every latitude and longitude informations are used to query the elevation value from the elevation extraction class, that provides the altitude for that point. Then, for every couple of consecutive points are calculated some values like the time difference between them, the distance between them and the slope angle. In this way, it is possible to derive speed and accelerations, and all these information are passed to the consumption calculation class that calculates the final value for that segment. Every consumption value for the segments is then summarized, in order to obtain the final consumption value for the trajectory.

For a better comprehension, Figure 3.6 shows a simple diagram of the classes used for the simulation.

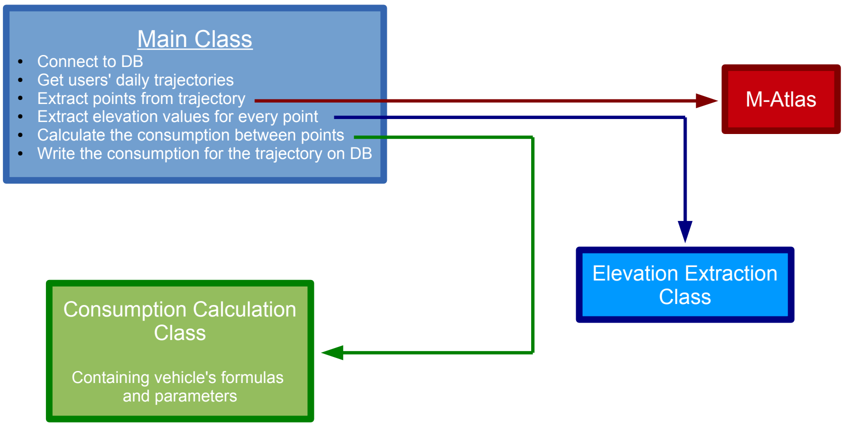

Figure 3.6: Diagram of the classes used in the simulation

As mentioned in Section 2.1.1, an EV, as every other vehicle, has to deal with some forces during every day usage: physical forces like wind, gravity and traction forces are fundamental in calculating the overall consumption. In fact, as every vehicle, an EV has to deal with the rolling resistance force of the tires, the aerodynamic drag due to the front surface area of the vehicle, gravity and inertial force. But an EV's consumption is subject also to some other variables like the efficiency of the transmission, the efficiency of the electric machine and the gear efficiency. Then, as previously mentioned, other variables should be considered in the overall consumption, such as the self discharge rate and the auxiliary power used for electric and electronic extra equipment (radio, wipers, air conditioning etc.). These two last parameters were not considered here, and were set to 0. Then, in case of braking, the vehicle is also capable of regenerative braking, thus regenerating a little part of the energy that recharges the battery. In this simulation the regenerative breaking was set to standard mode, which is a value of 0.2 (i.e. the 20% of the energy flows back to the battery).

Then, the elevation class is described. The elevation data is gathered from the CGIAR Consortium for Spatial Information (CGIAR-CSI) web site. On their website it is possible to choose the elevation data type by resolution. The data type chosen is the Shuttle Radar Topography Mission (SRTM) 90m Digital Elevation Data, the more precise data source available. The format of the data is GeoTIFF, a multi-layer Tiff image that retains altitude information for every pixel. The GeoTIFF is in the format of 6001*6001 pixels, every one having a $90m \times 90m$ side, and covers a little square part of the globe. This "squares" are not big enough to cover the whole Italian country inside one single file; for this reason, the elevation class has to choose which of the 7 GeoTIFF that contains the Italian country is good to make the altitude request. Then, after the choice, an external bash code is executed in order to extract the final value. The program used in bash is called GDal, and its called routine is *gdallocationinfo*, with some extra parameters used to ask only the value of altitude.

Some important observations have to be made. The elevation values extracted refer to the terrain, and the elevation between two adjacent spots can differ of at least 1 meter. This leads to two considerations: the first one is that there is no direct reference to the road elevation in the GeoTIFF. It means that bridges and tunnels are not considered in it, but the elevation considered refers to real terrain elevation. Although this is considered an issue that could make this elevation sources unreliable, no better solutions for the road elevation extraction were available at the time of writing. The second consideration is that in real driving, the road level does not vary in 1 meter steps, but varies gradually. For this reason, in order to simulate a real slope angle, a moving average of the 10 passed elevations is applied over each point, so to smooth the elevation curve and make it more realistic.

3.3 Scenario 1 Experiments

In this section the experiments on Scenario 1 are presented. The first java simulation, based on Algorithm 1 is implemented in such a way that, for every point, the elevation extraction subroutine (bash command *gdallocationinfo*) has to be called. That elevation extraction subroutine is soon proved to be too slow, as it would have taken $\approx$ 58 days to complete the consumption calculations on the starting table, that, as previously mentioned, consisted of$\approx$ 7 millions trajectories. Then, we managed to keep the result of each queried pixel elevation in memory, in order to ask only once the elevation for every pixel, only if needed. In such a way, the computation initially is slow, but after some extractions begins to speed up, and finally takes just 4 days to finish. In this first run we do not consider any interpolation, therefore we use the actual points recorded from the GPS device for the calculation, with their resolution between $\approx$ 1 and $\approx$ 3 minutes.

The results obtained are then queried and interpreted. As it is possible to see in Figure 3.7, the X axis represents the percentage of days (where 100% relates to all 31 days covered for June 2011), and the Y axis represents the percentage of users. With this graphic representation we can see how the percentage of users varies on the change of the percentage of daily trips. This is the plot that summarizes the key index used for the evaluation of the results.

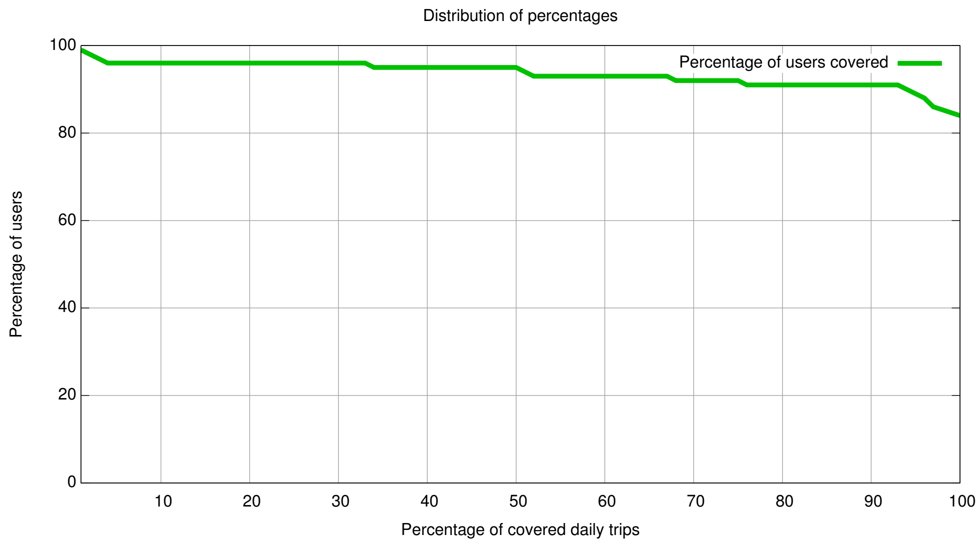

Figure 3.7: Electrification Rate

The results are that a 83% of users are fully covered in their daily trips by an EV.

From this analysis, there are some considerations. The time between every recorded point is very long, considering the fact that this data were not supposed to be used for a simulation, but just to insurance purposes. This may be a problem, because between one point and the next one could exist some big and unseen variations in elevation, thus modifying consequently slope degrees and the final consumption. For this reason, the same calculations are made, but this time including also an interpolation process that brings data in 1 second resolution. As can be seen in Algorithm 2, the algorithm is changed to consider interpolation.

The consumption, as stated above, can differ very much from the previously calculated one, due to the higher elevation relevance. In fact, Figure 3.8 shows that the elevation between two distant points can fluctuate very much, thus changing the estimated consumption.

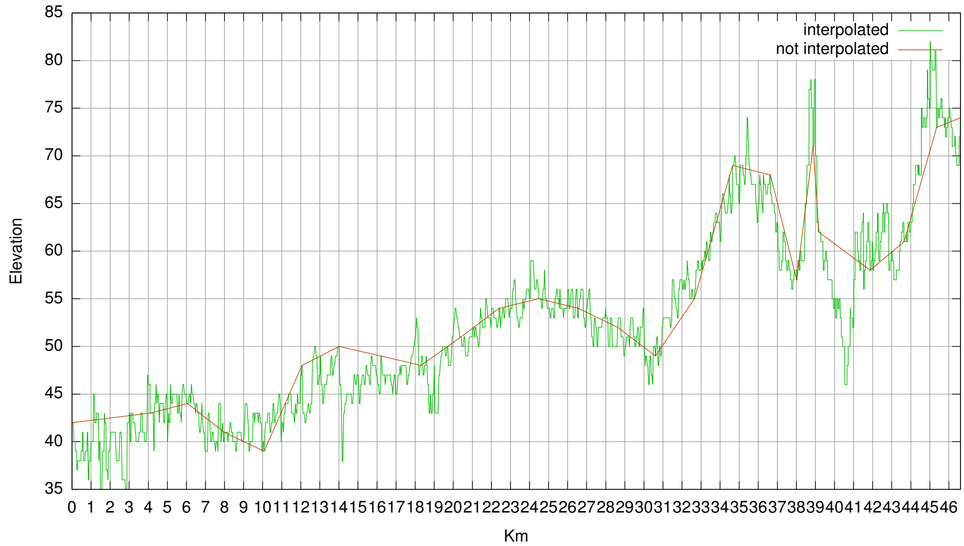

Figure 3.8: The elevation fluctuation put to light by interpolation

In Figure 3.8 is possible to see, in X axis, the meters, and in Y axis the altitude of that point. In the specific, on X axis every rectangle represents a km, for a total of 47 km trip.

In order to see to what extent the consumption changes, Figure 3.9 shows a scattered plot for a sample of 10.000 users of the database. X axis represents consumption before interpolating, and Y axis consumption after the interpolation. Here, every dot represents a single user, and its position determines if the consumption has increased or not. As it is possible to see, the overall consumption tends to move up in a significant way.

Then, as described in Section 3.2, the elevation curve is smoothed, this to make the slope angle variations more realistic. In fact, as described in Section 3.2, the elevation extraction can only give values in a 1 meter resolution, thus simulating an EV that changes the elevation of 1 meter in one second

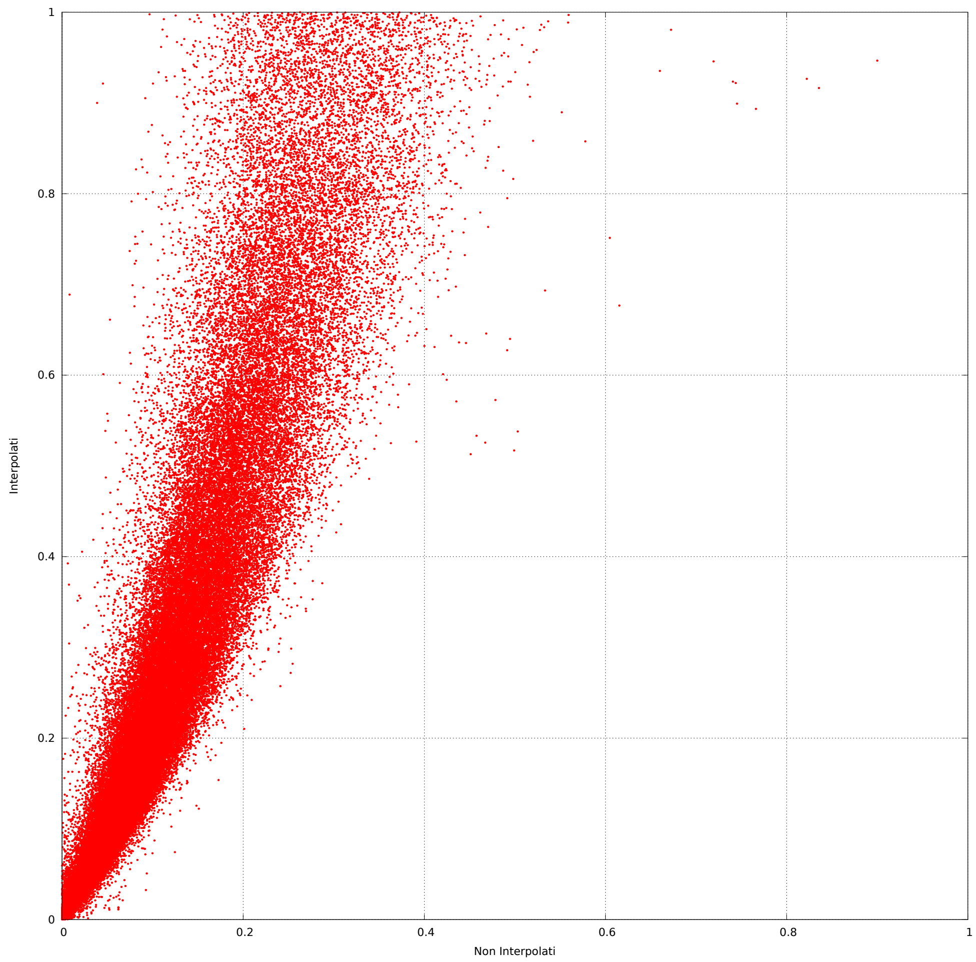

Figure 3.9: A scattered plot for consumption change

(that is pretty unrealistic). For this reason, a moving average of 10 previous elevations is applied to every point, and the result is shown in Figure 3.10.

As final result, this time a 30% of users are fully covered in their daily trips by an EV (Figure 3.11). This is a really important result, and several conclusions can be done. First of all, with this plot we can see how the percentage of users covered grows with the reduction of the index of electrifiability: if we want to reach an index of 100% electrifiable journeys, we reach just the 30% of the entire population. But if we can decrease the index, let's

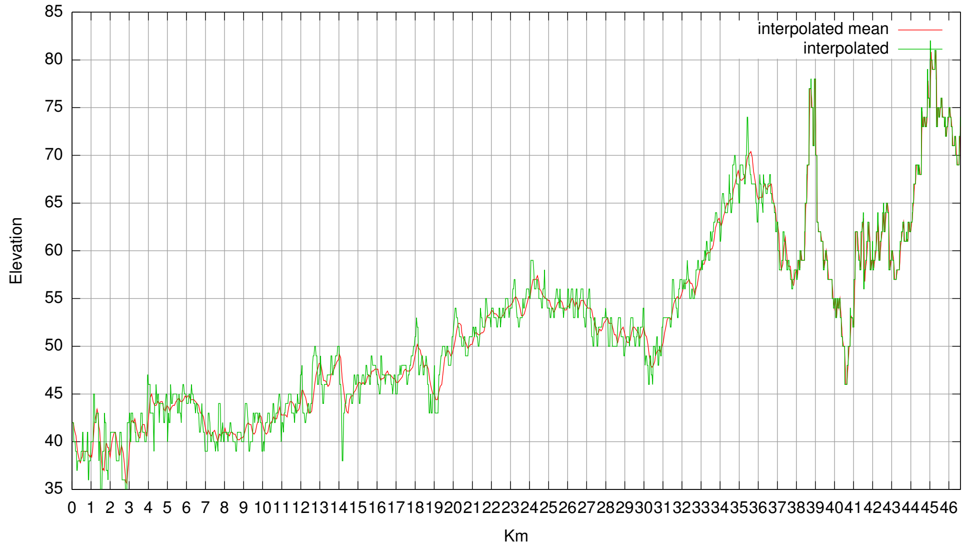

Figure 3.10: The smoothed elevation curve

say, at 80%, the users covered grow up to ≈ 55%. Another point we can state is that with this result we cannot ever reach the 80% of the population, because the curve has a too plain shape.

At this point, the question is: is this the final result we really searched for? Is this 30% of totally electrifiable users the real value of electrification rate of Tuscany we wanted to find? To answer these questions, another analysis is conducted upon obtained data, as represented in Figure 3.12.

It is a density plot where we can see, in X axis, the total days that a user was seen in the database (i.e. he/she took the car), while on the Y axis the number of days covered by the EV for that user. On the diagonal there are all the users for which there is a 100% coverage, and every color represents the number of users in that point of the plot.

It's interesting to note that, in the lower-left side, there is a high number of users that appear for just 3 days (or less) and are not completely covered.

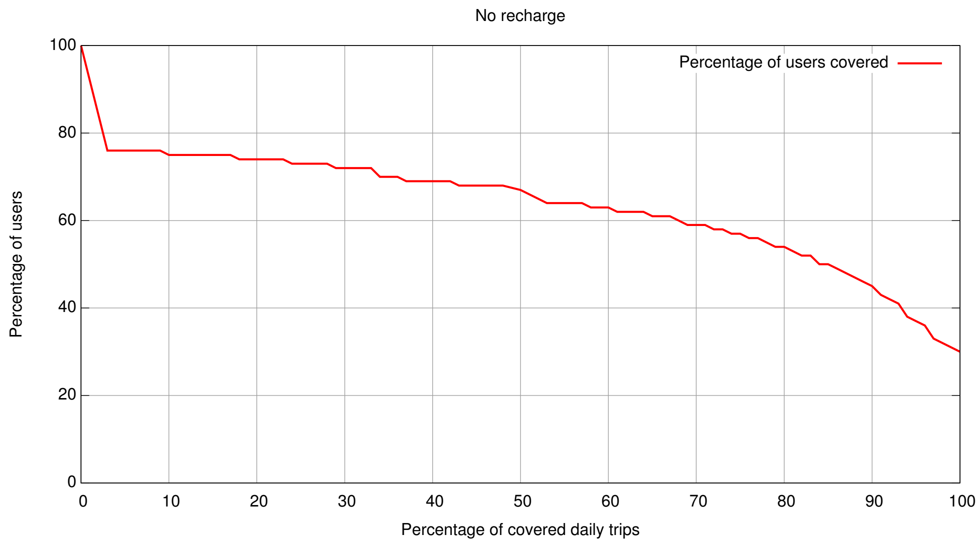

Figure 3.11: Electrification Rate - After Interpolation

This could mean that these users were only passing by Tuscany in the month of June 2011, thus implying that those users are not really Tuscany people. In order to investigate these hypothesis, these users were separated from the others, and a distribution of lengths of the two resulting groups were made. As it's possible to see in Figure 3.13, this *3 days' people* group is very different from the other one, because of the big difference of traveled kilometers. In fact, there is a great number of users (approximately one third of the population) that covers a lot more kilometers than the other group. We could imagine that those people, having so much more driven kilometers in just a few days (not more than 3), is people passing by Tuscany, that leads, in this case, a misrepresentation of the Tuscany people: in other words, they should not be included in the electrifiability analysis.

For this reason, another plot has been calculated, considering, this time, only Tuscany's people. The result is shown in Figure 3.14, where the 100%

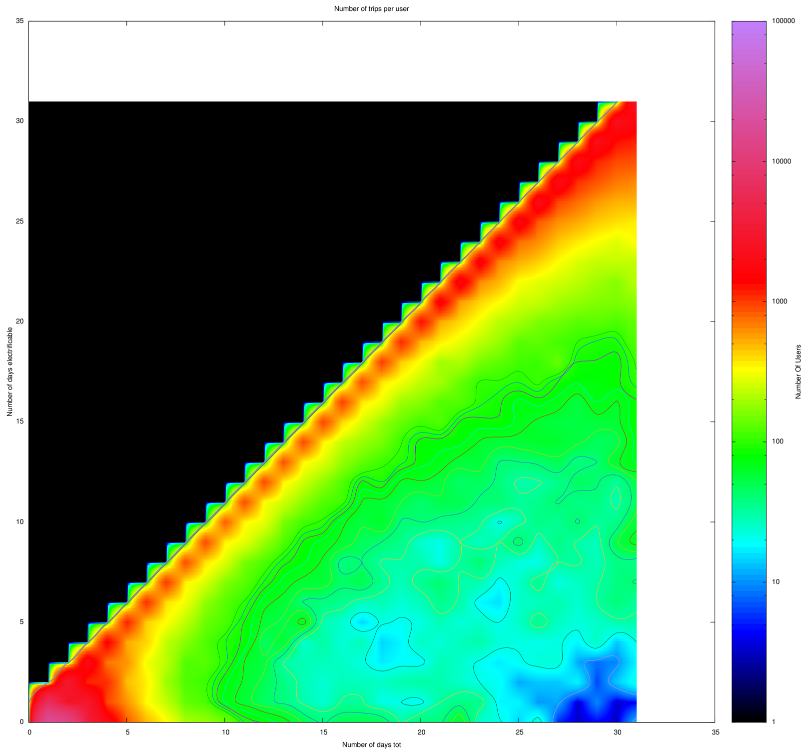

Figure 3.12: Density plot scenario 1

of users consists of 36435 users this time. As it is possible to see, the final percentage for 100% covered trips is the same (29%) but the rest of the curve reaches more quickly high percentage of users reducing the index of electrifiability. This is a far better result than the one shown in Figure 3.11: the total number of users covered is far higher than the previous one: it is reasonable to think that not all the users will be covered, and so, if we would like to choose a target level of electrifiability to be reached, it maybe would be the 80%, because, at this level we would be able to reach more than 70%

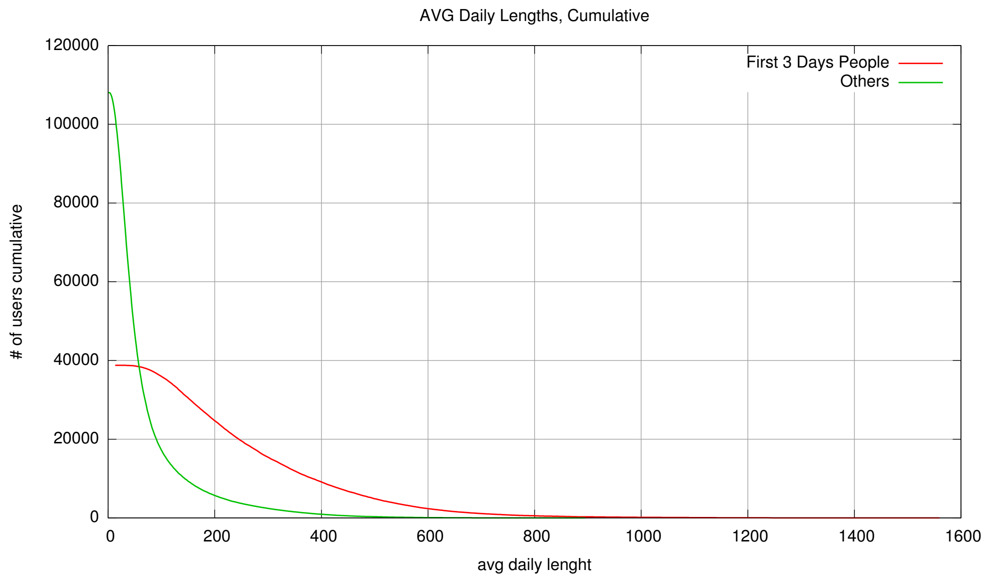

Figure 3.13: Difference between the two identified groups in traveled kms

of Tuscany people. To better show this fact, in Figure 3.15 is presented the comparison between the two results. The figure shows, in fact, that a bigger area is covered by the new Tuscany index plot.

It is then interesting to assess the electrification rate of the main Tuscany's cities for this scenario. Figure 3.16 shows this analysis for the main cities. The cities took into exam are Pisa, Florence, Livorno, Prato, Arezzo, Lucca and Siena. In the picture it is possible to see that Arezzo reaches the biggest covered users' percentage, upon and index of electrifiability of 100%, followed by Lucca and Pisa. An interesting behavior is shown in Livorno, that is for the most part slightly separated from the other cities behaviors (in a negative sense), while, coming up to higher index values, it intersects with Florence and Siena and reaches higher values of those 2 cities with high levels of the index.

To better analyze the most crowded cities, Figure 3.17 focuses the atten-

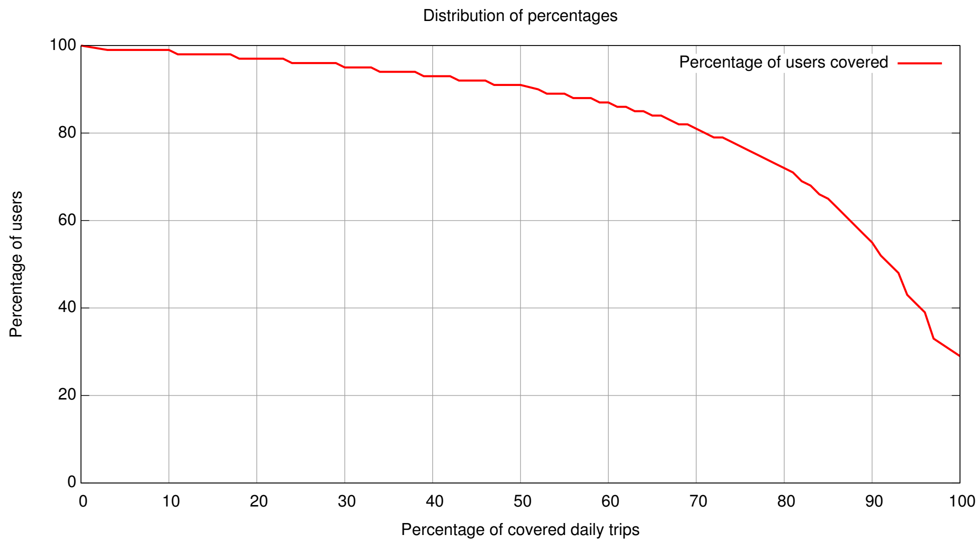

Figure 3.14: Tuscany's users' rate of electrifiability

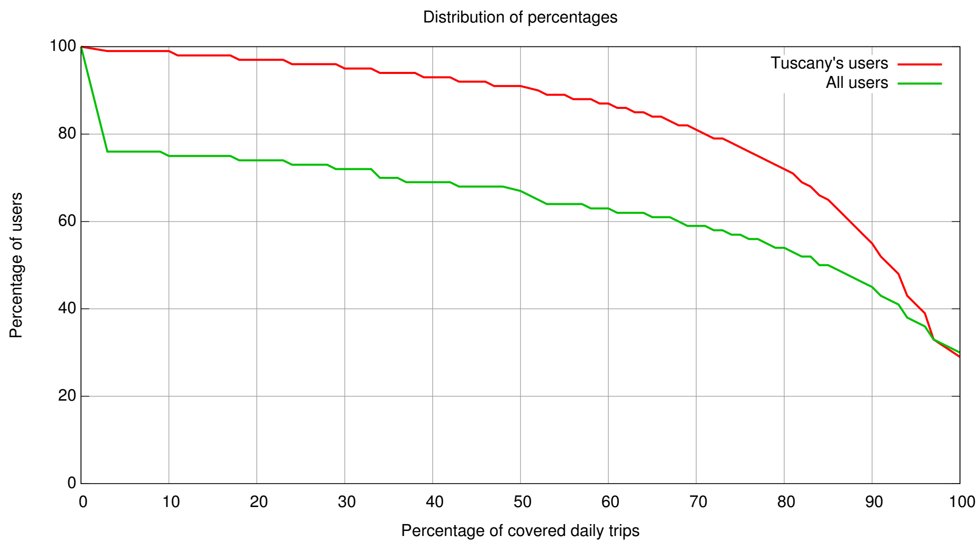

Figure 3.15: Tuscany's users VS all users' rate of electrifiability

tion on the 4 biggest cities, which are Pisa, Florence, Livorno and Prato. also here it is possible to notice the different behavior of Livorno, stating that for this city, the users tend to move outside it to a greater extent in comparison with other places. Here is then possible to notice that, in this plot, Pisa has the highest level of total electrifiability, while Florence is the last city.

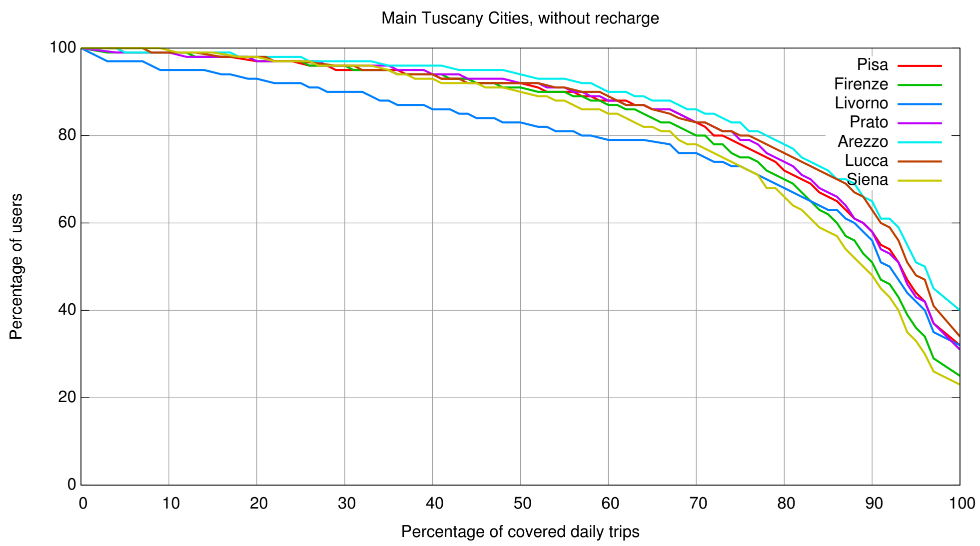

Figure 3.16: Electrification rate of the main cities of Tuscany, scenario 1

Some preliminary observations can be done: considering Tuscany, and then the single cities, it is already possible, for this scenario, to find a good index of electrifiability, meaning that, if we want to cover the 80% of total journeys, we could reach the 70% of people just today.

3.4 Scenario 2 Experiments

In this section the experiments on Scenario 2 are presented, in which the user can recharge the EV's battery everywhere if it stops for at least 2 hours. The java simulation is based on Algorithm 3, and takes 3 days to complete. This time, no elevation extraction functions are called, because the starting dataset is the table created within Scenario 1, where the consumption values

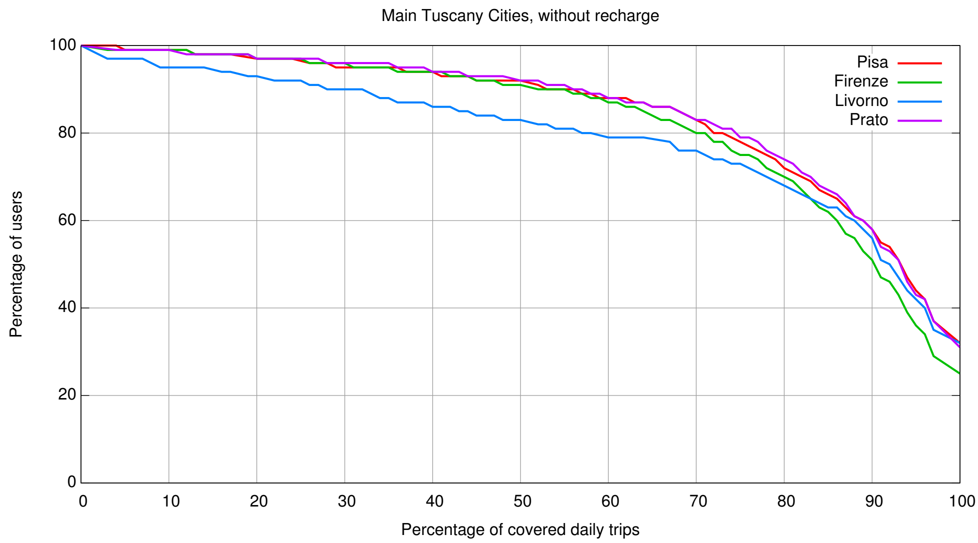

Figure 3.17: Electrification rate of the main cities of Tuscany, scenario 1

for each trajectory are already present. The algorithm, shown in Section 2.4, states if the vehicle stopped for the minimal required time to recharge batteries, and if that condition is true, it recharges the battery at the rate of 3 kW/h. It stops if 100% SoC is reached or, if the time before complete charge is higher than the time of stop, it stops charging before the total recharge is completed.

After the recharge program run, the starting table was enriched with some new calculated columns needed for keeping the information of residual battery SoC at the end of every trajectory.

It is possible to see the results of this simulation for the entire dataset in Figure 3.18, and the comparison between this result and the previous one in Figure 3.19. As it is possible to notice, the results are very similar, in curve's shape, as the one obtained with the entire dataset in scenario 1 (visible in Figure 3.11). In fact, also here it is not possible to reach the 80% of users' coverage,along with the substantially plain curve, that does not allow to gain a big number of users by reducing the index of electrifiability.

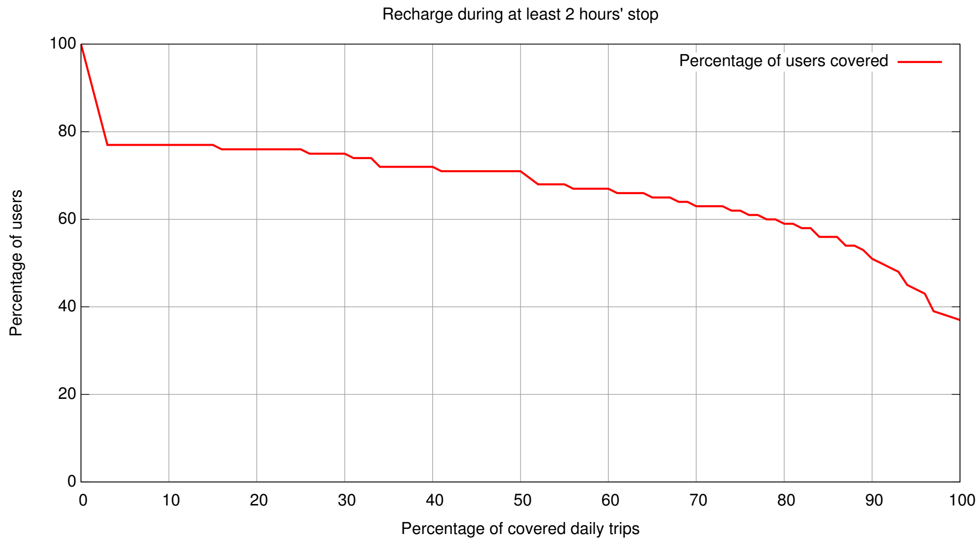

Figure 3.18: Electrification rate after recharge

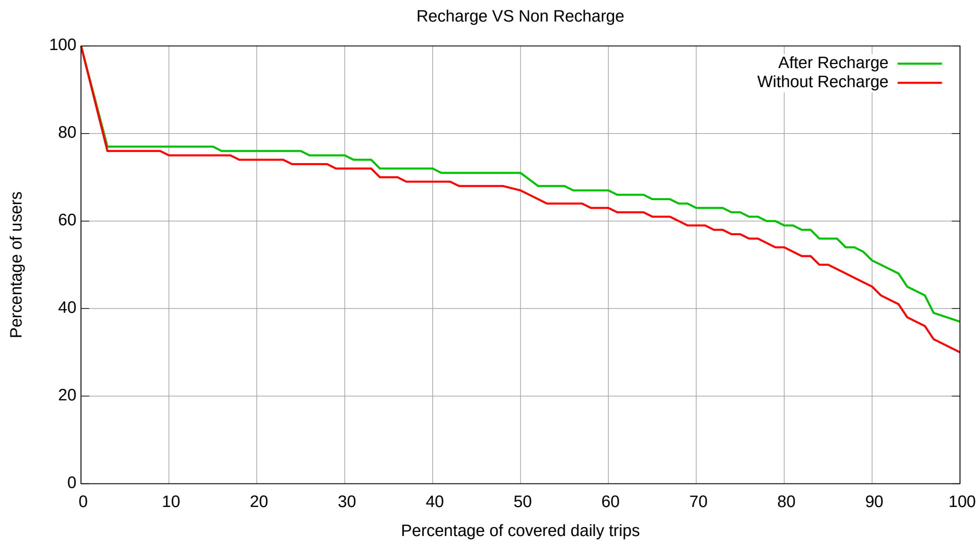

Figure 3.19: Comparison between the 2 scenarios for all dataset users

The result is an increment up to the 37% of users, for the entire dataset, that have their total daily trips covered. As for Scenario 1, also here the same considerations are valid: with the density plot, it is visible the same group of the so called *3 days' people*, but before investigating the analysis on just Tuscany users, an observation must be made. As it is possible to see in Figure 3.20, people, after recharge, tend to move in direction of the diagonal, meaning that more people are being covered by an EV than before.

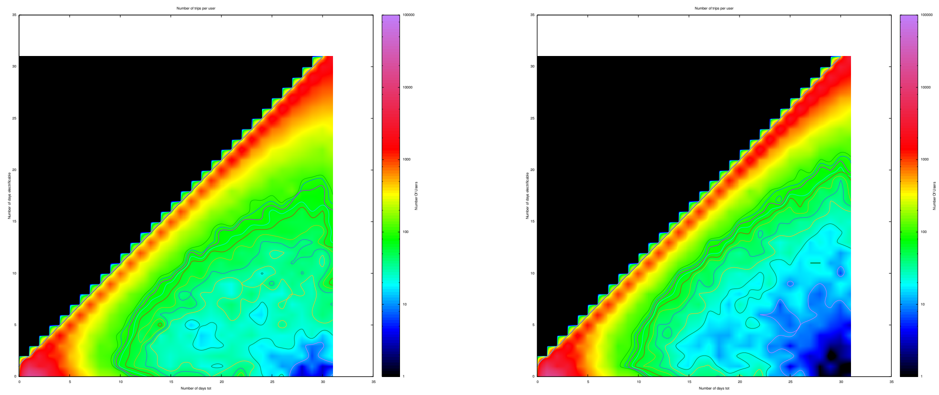

Figure 3.20: Comparison between densities without recharge (left) and with recharge (right)

Going on with the analysis, it is possible to check the electrification rate for Tuscany's users in Figure 3.21, while it is compared to the electrification rate for all users in Figure 3.22. In Tuscany, with the recharge, the increment respect all users simulation is greater, reaching 40% of users that have their total daily trips covered. But a more important result is worth mentioning: the curve of electrifiability grows more rapidly reducing the electrifiability index. This means that, compared with scenario 1, with a choice of the 80% of covered users we can reach more than the 80% of people covered by an

EV: this is an increment of more than the 10%. In addition, as it is possible to notice in Figure 3.22, the curve covers a greater area of the plot, thus covering a greater amount of users in percentage.

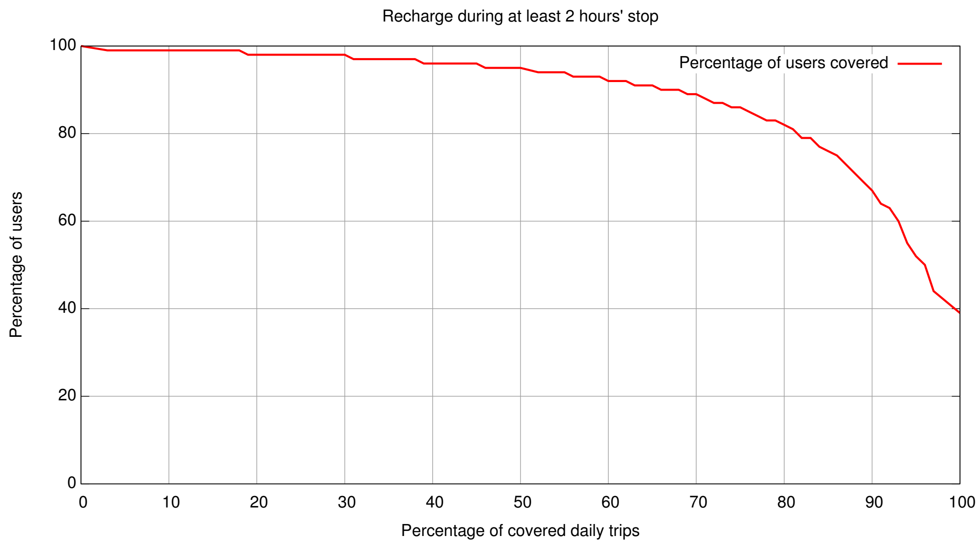

Figure 3.21: Electrification rate after recharge for Tuscany's users

Last but not least, an analysis over the main cities of Tuscany is conducted: as it's possible to see in Figures 3.23 and 3.24, the electrification rates of Pisa and Arezzo are the highest in Tuscany, and considering only Figure 3.24, where just the main 4 cities are considered (the biggest ones), Pisa has the highest level of electrification available, with 43%. Also here it is noticeable the behavior of the Livorno's electrification rate: it is initially slightly separated from the others for a long interval (in a negative way), then it finishes on top of Florence, meaning that, for 100% electrifiability index, covered users are a higher percentage than Florence, but when it comes to consider lower levels of electrifiability, Florence has an higher percentage of covered users. it is possible to meet this behavior also in Section 3.3.

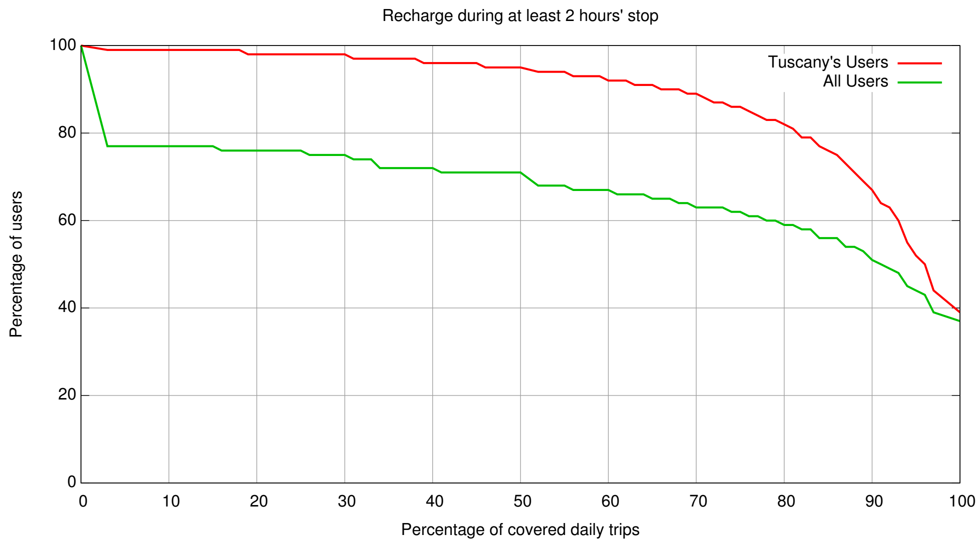

Figure 3.22: Electrification rate after recharge for Tuscany's users VS all users

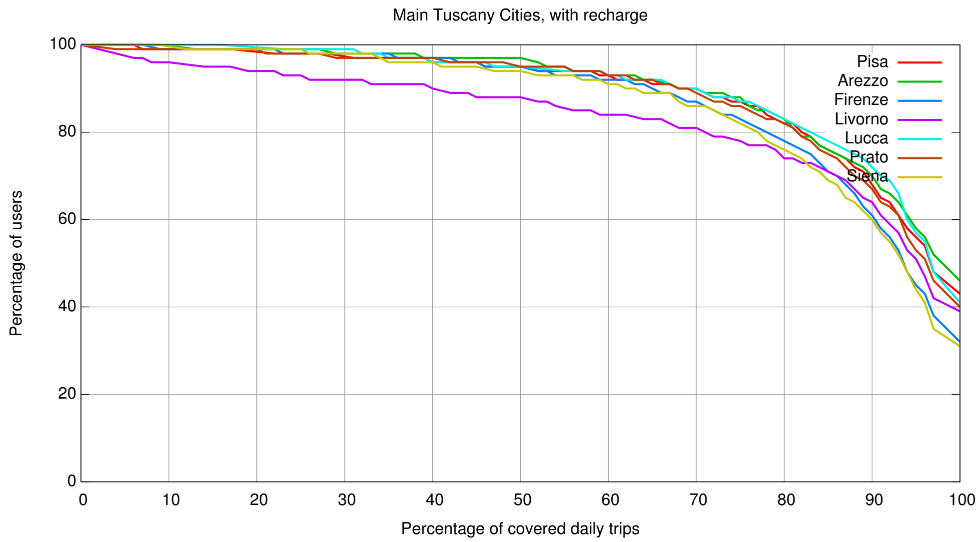

Figure 3.23: Electrification rate of the main cities of Tuscany, scenario 2

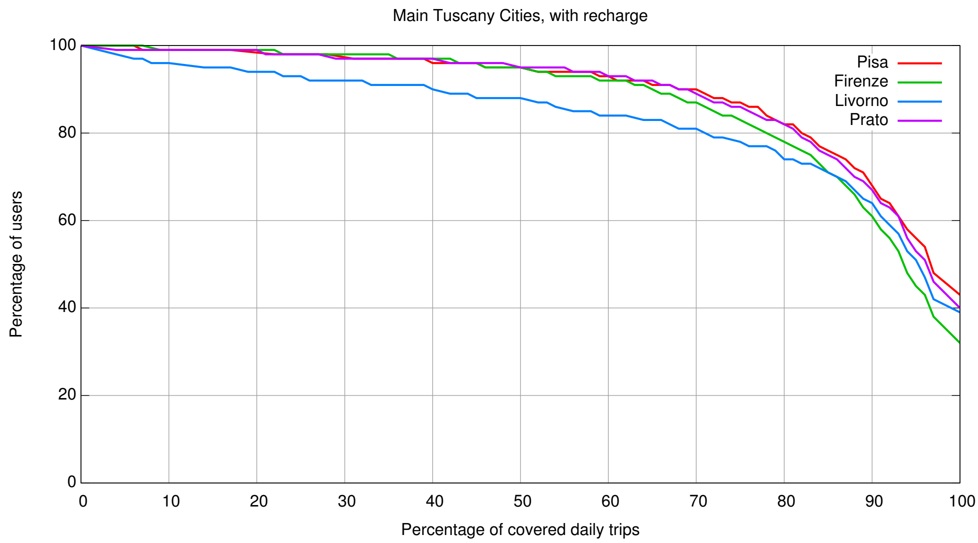

Figure 3.24: Electrification rate of the 4 main cities of Tuscany, scenario 2

Also here some preliminary considerations can be done: the main cities level of electrifiability is high enough to find a threshold that let's us to reach a great part of population. In fact, if we assume as good an electrifiability index of 80%, we could reach almost in any city the 80% of users.

Chapter 4

Conclusions

The future mobility, as everybody begin to imagine today, is going to be electric. This is because problems like air pollution, gasoline price and environmental impact are topics getting more and more visibility at everyone's eyes. But also if the need for green and sustainable mobility is getting higher, some relevant problems are still far from being resolved, such as the limited EVs kms range and the long times required for battery charging.

These problems are still not addressed by today's EV's technologies: although studies and researches are being conducted on batteries enhancement, there are still big progresses to be made, in order to guarantee on EVs the same advantages had today in gasoline powered vehicles. For this reason, the scientific community is very interested on how to address these issues with today's technologies, and make the actual generation of EVs an affordable and desirable alternative to gasoline powered vehicles for the most of users.

After having discussed some important and recent contributions in the literature, we introduced M-Atlas, the Mobility data mining platform we used for the data analysis. Then we presented the core work of this thesis: the simulation of consumption of an EV on real users trajectories, in order to find the electrification rates for specifical areas. The main scenarios were described, starting from the first and easiest, charging at home (where no charging locations other than home are available), and the second one, charging everywhere (where charging locations are available everywhere, and the EVs can charge if stopped for at least 2 hours).

The theoretical scenarios devised and formalized have been an important step in view of their implementation: in fact, we firstly set up the charge/discharge model of an EV's battery, and then we applied it on the real journeys of thousands of users of Tuscany by utilizing a real vehicle's specs in the model.

Some simplifications have been taken into consideration before proceeding with the simulations, and regard different topics. First of all, the model's simplifications are summarized: the EV simulated had no self discharge rate, and no auxiliary power modeled (i.e. these parameters were set to 0). Then, the interpolation took into account the elevation fluctuations, but did not consider any interpolation of accelerations, as it was set to be constant. Last, but not least, the scenarios were intended to be a simplification of the reality, thus enabling us to model it.

In scenario 1 we firstly analyzed the consumption for the actual data sampling of the dataset: this in order to analyze the electrificability rate on just the available data. As a result, we achieved an 83% of total electrificable trips, but this result, because of the low sampling of the data, did not take into consideration the variation of elevation between every distant point. For this reason we interpolated the original data in order to gain 1 second data sampling, in order to consider the fluctuations of elevations previously ignored. We found that the entire dataset has a total electrification rate of 30%, and it is never able to reach 80% even with low electrificability rate levels, such as 10% for example: by stopping to an electrificability rate of 80% for example, we can reach only the 55% of users covered. Then we stated that, taking only people coming from Tuscany into exam, the total electrifiability rate was still the 30%, but the curve was characterized by a more inclined behavior: by stopping to an electrifiability rate of 80% for example, we are able to reach more than the 70% of covered users. Similar considerations can be done for the main cities of Tuscany, analyzed in Section 3.3, where all scenario 1 analysis are conducted.

In scenario 2 there is the possibility of recharging the EV's batteries wherever the user stops, if the stopping time is greater than 2 hours. The simulation run over the entire dataset gave, as a result, an increment in reached users up to the 37% for totally covered trips. The same considerations made in scenario 1 are valid here: the electrification rate is never able to reach 80% in any level. Taking only Tuscany's people into exam, the simulation showed a similar behavior as in scenario 1: the total electrifiability rate is 39%, but with a more inclined curve here too. In fact, by reducing the target of electrifiability rate, there is the possibility to reach more rapidly a growing number of users. Just for example, by stopping to an electrifiability rate of 80%, we can reach up to the 80% of users.

Some considerations have to be kept in mind evaluating the final results: the elevation extraction has to be implemented considering tunnels and bridges, in order to make vehicles' consumption be more realistic, but this is an upper bound error, so at least vehicles' consumption is overestimated. Moreover, interpolation of data has to be implemented also on real accelerations, instead of using constant acceleration like did here (for the reasons explained in Section 2.4), but with 1 second sampled GPS data this problem could be addressed. Keeping in mind these considerations, the results can be considered a good starting point in order to understand the level of possible electrifiability of the main cities: for some of the cities took into exam, it's already possible to convert a pretty big percentage of vehicles into EVs, without making users change their habits.

This thesis presents 4 main contributions: (a) the creation of a model that simulates the discharging and recharging process of an EV, and application of this model on Tuscany's real users trajectories; (b) the identification of the main EVs usage scenarios, applying a simplification of the reality; (c) the identification of an index of electrifiability that shows if the simulation on trajectories fails or not; (d) the implementation of the simulation.

Additionally, future analysis are to be conducted on the remaining scenarios. Those scenarios are fundamental to assess the electrification rates of a city or a region in a more detailed way, as changes between different environments, such as infrastructures differences for example, can occur. All the possible future scenarios to be considered are presented in Section 4.1.

4.1 Future works and scenarios

In this section are presented some of the scenarios that could be analyzed in future works. They represent the addition of more complex constraints, in order to model, in a more precise way, the different conditions that can be met in real life. They are:

- scenario 3: charging spots, where a constraint on charging locations is inserted;
- scenario 4: fast charging, where the possibility of fast charging the EV is added;

- scenario 5: Vehicle2Grid, where the possibility of using EVs as batteries for injecting back energy to the grid is added.

4.1.1 Scenario 3: charging spots

In Scenario 3 a new constraint is introduced: EVs can recharge their batteries only at the charging spots and if they stop for at least 2 hours. This is a more realistic example of usage of EVs, and its result of electrifiability rate is meant to be situated between the scenarios 1 and 2.

In a real upcoming future, the most real situation sees the presence of multiple charging stations, situated in the main parkings. The location of the charging spots can also be derived and calculated with an optimization algorithm (not yet discussed here) that finds the most frequented locations where users stop their vehicles. This is an interesting scenario, because the charging stations are not available everywhere: in fact, economic, structural or territorial reasons are a limit for the widespread diffusion of charging stations. Figure 4.1 shows a map of the locations of currently installed public charging stations in Pisa. As it is possible to see in the figure, charging stations are not present in every street, and this is caused by the above mentioned limitations.

4.1.2 Scenario 4: fast charging

This scenario adds the possibility of fast charge, but only near specific charging stations, because of the technically advanced equipment that can be found only in them. This mode, called Level 3 charging, allows to recharge an EV at a rate of 50 kW/h in Direct Current. This situation is of great interest of study, because it could bring the EVs time of re-usability very close to that one of gasoline powered vehicles. In fact, one of the reasons that EVs

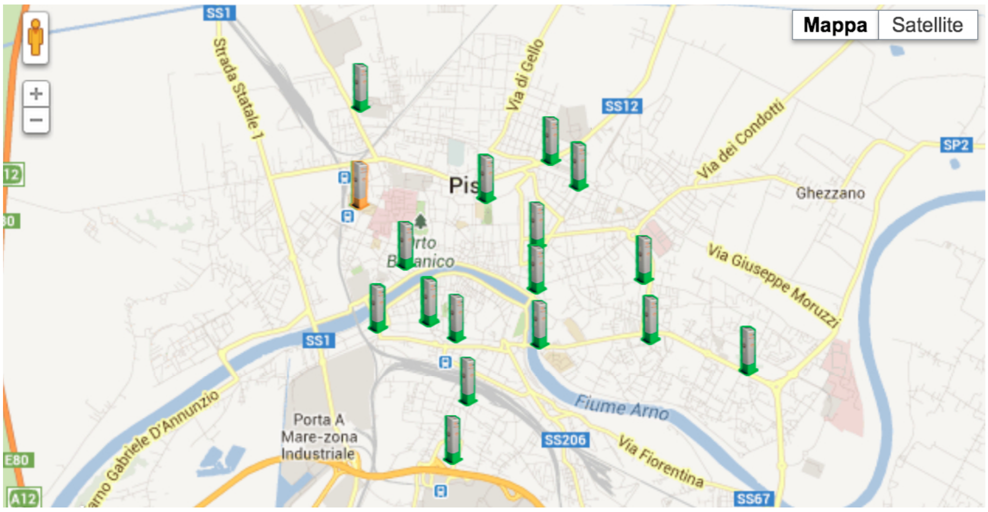

Figure 4.1: Map of some charging spots in Pisa

are still not mature enough today to take the place of gasoline vehicles is the difficulty to recharge batteries in a short time, making the EV available to start again after some hours. The future wide-spreading of EVs passes also (and in great part) from this point: the speed of the recharge process. If this issue still remains misaddressed, EVs will never be an attractive choice for customers. For this reasons, future works will take in high consideration this scenario.

4.1.3 Scenario 5: Vehicle2Grid

The fifth and last scenario is of great importance. The future of the electric grid is of great interest of study. When the time of simultaneous charging of hundreds, or thousands, of EVs will come, the grid will have to face issues like overloads of cables and power transmitters, that could lead to infrastructure damages and, in the worst case, to an energetic crisis. For this reason, many studies on the electric grid's ability to adjust itself, and to make EVs start charging in specific time intervals are conducted. This is the main reason why it will be called *Smart* Grid.

The possibility of having, in EVs, hundreds of batteries connected for long periods of time to the grid could be an answer to all these problems: in fact, by using EVs stocked energy it could be possible to reduce peak loads, and to fill valleys of unused energy to recharge those batteries. Figure 4.2 shows a possible simulation: in red, it is added, to the basic consumption simulation based on the trajectory, the possibility of giving some energy back to the grid, but still preserving the minimum amount of energy useful to get back to home.

This is the final experiment that future works will conduct, having, in this way, simulated all possibilities of using an EV in the near future. At the end of this project, it will be possible to apply the simulations to every type of city, and/or in every possible country, meeting all the requirements in terms of constraints that every zone could face, keeping though the possibility of adding some new unpredictable constraints to the model.

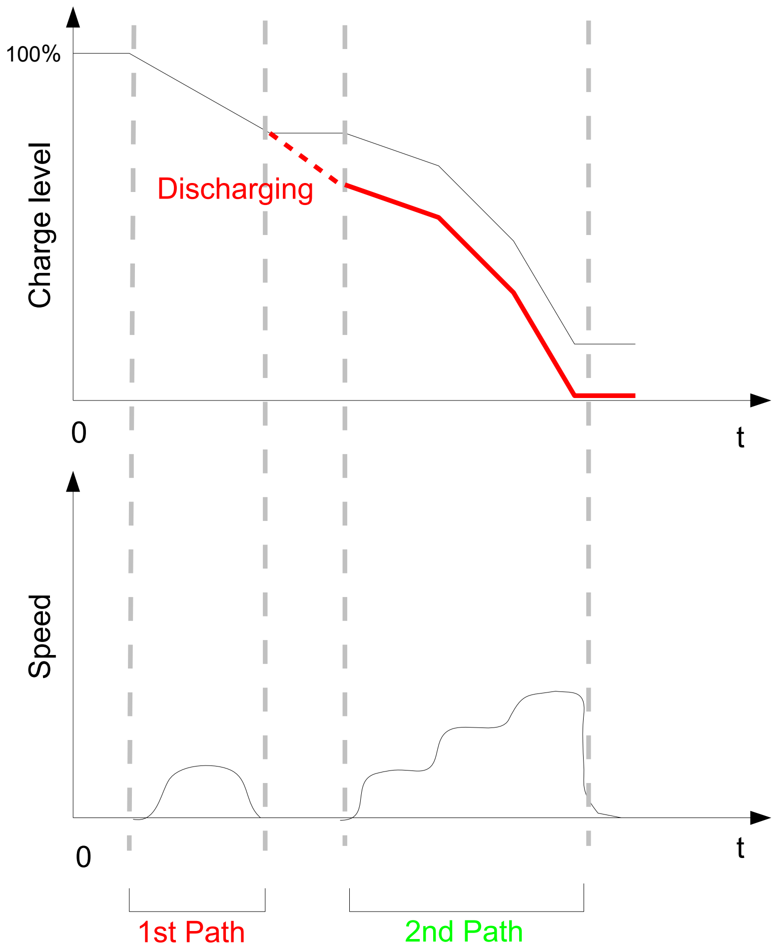

Figure 4.2: Scenario 5

Acknowledgments

Il primo ringraziamento va alla mia famiglia, che mi ha sempre sostenuto e spronato a dare il meglio di me, anche nel nuovo percorso universitario che ho intrapreso. Ai miei genitori, senza i quali oggi non sarei qui, e a mia sorella e suo marito, che mi hanno sempre supportato in quello che faccio (e che, per questo, mi hanno sempre chiesto di risolvere un problema al loro pc!) e che stanno per donarmi un bel nipotino. Ad Irene, che mi ha aiutato in momenti difficili ed è stata sempre presente in quelli belli, e, soprattutto, mi ha pazientemente ascoltato nella presentazione del mio lavoro!

Voglio inoltre ringraziare tutte le persone del KDD Lab che mi hanno dato consigli e suggerimenti preziosissimi, tra cui i miei relatori Salvo, che mi ha dato modo di tirare fuori il meglio di me, Barbara, che ha sacrificato dei weekend per correggere i miei errori, e Lorenzo, che mi ha dato preziosissimi suggerimenti in campi lavorativi per me poco conosciuti. Inoltre, il mio amico Luca che ha sempre avuto una parola di conforto e di sostegno, e tutti coloro che non ho modo di nominare e che in laboratorio hanno avuto un minuto per darmi una informazione su un problema o un momento di relax e di risate. Ultimo, ma non ultimo, il Professor Pedreschi, il mio primo relatore, grazie al quale ho intrapreso lo studio del data-mining, per le sue idee, la sua disponibilità e per l'entusiasmo che mi ha trasmesso per questo campo di studio.

Bibliography

[1] Yiming Pan, Yuanbiao Zhang, Zhang Zhang and Qinglong Cao, *Striving for Synthetic Benefits: Is your Country Suitable for the Widespread Use of Electric Vehicles*

[2] Wenbin Luo, Guanghui Liu and Xingjun Hu, *The Assessment of the Feasibility of Widely Using Electric Vehicles Based on the Analytic Hierarchy Process*, 2012 International Conference on Computer Distributed Control and Intelligent Enviromental Monitoring

[3] Xiaosong Hu, Fengchun Sun, Yuan Zou, *Comparison between two model-based algorithms for Li-ion battery SOC estimation in electric vehicles*

[4] Feng Ju,*Student Member, IEEE*, Junwen Wang, *Student Member, IEEE*, Jingshan Li, *Senior Member, IEEE*,Guoxian Xiao, and Stephan Biller, *Member, IEEE*, *Virtual Battery: A Battery Simulation Framework for Electric Vehicles*, IEEE Transactions on Automation Science and Engineering, Vol. 10, no. 1, January 2013

[5] Volker Schwarzer and Reza Ghorbani, *Drive Cycle Generation for Design Optimization of Electric Vehicles*, IEEE Transactions on Vehicular Technology, VOL. 62, no. 1, January 2013

[6] Matteo Vasirani and Sascha Ossowski, *A proportional share allocation mechanism for coordination of plug-in electric vehicle charging*, Engineering Applications of Artificial Intelligence 26 (2013) 1185 - 1197

[7] Shigaku Iwabuchi, Soh Masuko and Junichi Hoshino, *Energy Management System for Electric Vehicle based on User Behavior Induction*, CHI 2013 Extended Abstracts, April 27May 2, 2013, Paris, France

[8] Shiyao Chen, Timothy Mount and Lang Tong, *Optimizing Operations for Large Scale Charging of Electric Vehicles*, 2013 46th Hawaii International Conference on System Sciences

[9] Ovidiu Vermesan, Lars-Cyril Blystad, Reiner John, Peter Hank, Roy Bahr and Alessandro Moscatelli, *Smart, Connected and Mobile: Architecting future electric mobility ecosystems*, Design, Automation and Test in Europe 2013, 1740 - 1744

[10] Hamid Khayyam, Jemal Abawajy, Bahman Javadi, Andrzej Goscinski, Alex Stojcevski and Alireza Bab-Hadiashar, *Intelligent battery energy management and control for vehicle-to-grid via cloud computing network*, Applied Energy 111 (2013) 971 - 981

[11] A. Sheikhi, Sh. Bahrami, A.M. Ranjbar and H. Oraee, *Strategic charging method for plugged in hybrid electric vehicles in smart grids; a game theoretic approach*, Electrical Power and Energy Systems 53 (2013) 499 - 506

[12] Jaegue Shin, Seungyong Shin, Yangsu Kim, Seungyoung Ahn, Seokhwan Lee, Guho Jung, Seong-Jeub Jeon, and Dong-Ho Cho, *Design and Implementation of Shaped Magnetic-Resonance-Based Wireless Power Trans-*

fer System for Roadway-Powered Moving Electric Vehicles, IEEE Transactions on Industrial Electronics, Vol. 61, No. 3, March 2014

[13] P. Calnan, J.P. Deane , B.P. O Gallachoir, *Modelling the impact of EVs on electricity generation, costs and CO2 emissions. Assessing the impact of different charging regimes and future generation profiles for Ireland in 2025*, Energy Policy 61 (2013) 230 - 237

[14] Bruno Soares M.C. Borba, Alexandre Szklo and Roberto Schaeffer, *Plug-in hybrid electric vehicles as a way to maximize the integration of variable renewable energy in power systems: The case of wind generation in northeastern Brazil*, Energy 37 (2012) 469 - 481

[15] Javier Gallardo-Lozano, M. Isabel Milanes-Montero, Miguel A. Guerrero-Martinez and Enrique Romero-Cadaval, *Electric vehicle battery charger for smart grids*, Electric Power Systems Research 90 (2012) 18 - 29

[16] Willett Kempton and Jasna Tomic, *Vehicle-to-grid power fundamentals: Calculating capacity and net revenue*, Journal of Power Sources 144 (2005) 268 - 279

[17] Jukka Lassila, Juha Haakana, Ville Tikka, and Jarmo Partanen, *Methodology to Analyze the Economic Effects of Electric Cars as Energy Storages*, IEEE Transactions on Smart Grid, Vol. 3, No. 1, March 2012

[18] Abouzar Ghavami, Koushik Kar and Aparna Gupta, *Decentralized Charging of Plug-In Electric Vehicles with Distribution Feeder Overload Control*, CoRR 2013

[19] Fjo De Riddera, Luk Knapenb and Davy Janssens, *Applying an Activity Based Model to Explore the Potential of Electrical Vehicles in the Smart*

Grid, The 2nd International Workshop on Agent-Based Mobility, Traffic and Transportation, Procedia Computer Science 19 (2013) 847 - 853

[20] Joosung Kang, Scott J. Duncan and Dimitri N. Mavris, *Real-time Scheduling Techniques for Electric Vehicle Charging in Support of Frequency Regulation*, Procedia Computer Science 16 (2013) 767 - 775

[21] Changhua Zhanga, Qi Huanga, Jiashen Tiana, Lei Chena, Yongxing Caob and Ran Zhangb, *Smart Grid Facing the New Challenge: the Management of Electric Vehicle Charging Loads*, ICSGCE 2011: 2730 September 2011, Chengdu, China, Energy Procedia 12 (2011) 98 - 103

[22] Xiaomin Xi, Ramteen Sioshansi and Vincenzo Marano, *Simulationoptimization model for location of a public electric vehicle charging infrastructure*, Transportation Research Part D 22 (2013) 60 - 69

[23] M.A. Lopez, S. Martin, J.A. Aguado and S. de la Torre, *V2G strategies for congestion management in microgrids with high penetration of electric vehicles*, Electric Power Systems Research 104 (2013) 28 - 34

[24] Jeffrey S. Marshall, Paul D. Hines, Jiarui D. Zhang, Francesca Minervini and Simtha Rinjitham, *Modeling the impact of electric vehicle charging on heat transfer around underground cables*, Electric Power Systems Research 97 (2013) 76 - 83

[25] Sangyoung Park, Younghyun Kim and Naehyuck Chang, *Hybrid Energy Storage Systems and Battery Management for Electric Vehicles*, DAC '13 Proceedings of the 50th Annual Design Automation Conference, Article No. 97, 2013

[26] Attila Gollei, Peter Gorbe and Attila Magyar, *Modeling and optimization of electrical vehicle batteries in complex clean energy systems*, Journal of Cleaner Production 34 (2012) 138 - 145

[27] Fosca Giannotti, Mirco Nanni, Dino Pedreschi, Fabio Pinelli, Chiara Renso, Salvatore Rinzivillo and Roberto Trasarti, *Unveiling the complexity of human mobility by querying and mining massive trajectory data*, The VLDB Journal (2011) 20, 695 - 719

[28] Jesus Fraile-Ardanuy and Carlos Gamallo, ETSI Telecomunicacion, UPM, *Electric Vehicle Range Calculation*, a 2013 Techincal Report DATASIM