CTU FEL Prague

Master thesis

EV vehicles

and charging demand

Bc. Jakub Zíka

May 1, 2025

CTU FEL

CTU FEL Prague

Disclaimer
You can edit this page to suit your needs. For instance, here we have a no copyright statement, a colophon and some other information. This page is based on the corresponding page of Ken Arroyo Ohori's thesis, with minimal changes.

Colophon
This document was typeset with the help of KOMA-Script and LaTeX using the kaobook class.
The source code of this book is available at:
https://github.com/fmarotta/kaobook
(You are welcome to contribute!)

Publisher
First printed in May 2019 by CTU FEL

Contents

Introduction

LLM generated slop *give historical context on how automobiles came to be and used and how gas stations infrastacturue had to be created. And how EV chargers are very similar to the spread of ICE vehicles.*

The invention of the automobile in the late 19th century revolutionized human mobility, enabling unprecedented freedom to traverse long distances. However, this breakthrough hinged not only on the internal combustion engine (ICE) itself but also on the parallel development of a critical support system: gasoline stations. Just as early motorists relied on scattered fuel depots to power their journeys, the rise of ICE vehicles necessitated a standardized, accessible network of refueling infrastructure to sustain their adoption. This symbiotic relationship between vehicles and their energy infrastructure became a cornerstone of modern transportation, shaping urban planning, economic systems, and global energy policies.

Today, as societies pivot toward sustainability, electric vehicle (EV) are heralding a similar paradigm shift. Yet their widespread adoption faces a challenge mirroring the early days of automobiles: the need for reliable, equitable, and efficient charging infrastructure. While EVs eliminate tailpipe emissions, their practicality depends on overcoming "range anxiety" and ensuring charging availability aligns with user behavior—issues that gas stations largely resolved for ICE vehicles over a century of iteration. Predicting EV charger usage, therefore, is not merely a technical exercise but a good step in designing infrastructure that mirrors the ubiquity and convenience of gas stations. And helps smoothen transition.

1 Motivation - Electromobility and Climate Change

This chapter explores the fundamental motivations behind the transition to electric mobility, examining the environmental imperatives driving this shift, the evolution of electric vehicles, regulatory frameworks accelerating adoption, and the critical infrastructure challenges that must be addressed to enable widespread electrification of transportation.

1.1 Climate change

Climate change represents one of the most pressing global challenges of our time, with transportation being a significant contributor to greenhouse gas emissions. The Intergovernmental Panel on Climate Change (IPCC) has consistently identified the burning of fossil fuels as the primary driver of anthropogenic climate change, with transportation accounting for approximately 24% of direct CO_2 emissions from fuel combustion globally [**ipcc2022**]. Within this sector, road vehicles—particularly those powered by internal combustion engines (ICEs)—are responsible for nearly three-quarters of transport emissions.

The environmental impact of ICE vehicles extends beyond carbon dioxide emissions. These vehicles also produce nitrogen oxides (NO_x), particulate matter, and other pollutants that contribute to poor air quality, respiratory diseases, and premature deaths in urban areas. The World Health Organization estimates that air pollution causes approximately 7 million premature deaths annually, with vehicle emissions being a significant contributor in urban environments [**who2021**].

Figure 1.1: Global greenhouse gas emissions by sector, highlighting transportation's contribution (Source: IPCC AR6, 2022)

While the scientific consensus on anthropogenic climate change is robust, it is worth acknowledging that the transition to electric mobility is not without controversy. Critics point to the environmental impact of battery production, concerns about electricity generation sources, and the socioeconomic implications of rapid technological change. However, lifecycle analyses consistently demonstrate that even when accounting for battery production and electricity generation, electric vehicles produce significantly lower lifetime emissions than their ICE counterparts in most regions of the world, with this advantage growing as electricity grids incorporate more renewable energy sources [**eea2018**].

The climate imperative is particularly challenging for developing nations, which face the dual pressures of reducing emissions while supporting economic growth and mobility needs. These countries often lack the infrastructure and financial resources to rapidly transition to electric mobility, yet they are frequently among the most vulnerable to climate change impacts. International cooperation, technology transfer, and equitable financing mechanisms will be essential to ensure that the global transition to sustainable transportation does not exacerbate existing inequalities.

1.2 Electric vehicles

The history of electric vehicles (EVs) is marked by a fascinating parallel development alongside internal combustion engine vehicles, rather than being a purely modern innovation. In fact, electric vehicles were among the first automobiles developed in the late 19th century, with inventors like Thomas Parker creating practical electric cars as early as 1884. During the early automotive era, electric vehicles competed directly with steam and gasoline-powered vehicles, and were particularly popular in urban environments due to their quiet operation, absence of exhaust, and ease of use compared to early ICE vehicles that required hand-cranking to start.

However, the limitations of early battery technology—particularly in terms of energy density, range, and recharging infrastructure—combined with the discovery of abundant petroleum reserves and the introduction of the electric starter for gasoline engines, led to the dominance of ICE vehicles throughout most of the 20th century. Electric vehicles remained largely confined to specialized applications such as forklifts, golf carts, and other short-range utility vehicles.

The modern resurgence of electric vehicles began in earnest in the late 1990s and early 2000s, driven by advances in lithium-ion battery technology, growing environmental concerns, and regulatory pressures. The introduction of hybrid vehicles like the Toyota Prius served as a transitional technology, familiarizing consumers with electric drivetrains while alleviating range anxiety through the backup of a gasoline engine. The launch of the Tesla Roadster in 2008 demonstrated that electric vehicles could offer performance comparable to or exceeding that of high-end sports cars, challenging perceptions that EVs were inherently limited in capability.

Today's electric vehicles have largely overcome many of the historical limitations that hindered their adoption. Modern EVs offer ranges exceeding 300-400 kilometers on a single charge, with high-end models approaching 600 kilometers. Fast-charging infrastructure has expanded significantly, enabling long-distance travel with reasonable charging stops. The total cost of ownership for EVs has become increasingly competitive with ICE vehicles due to lower operating and maintenance costs, despite higher initial purchase prices—a gap that continues to narrow as battery costs decline and economies of scale improve.

Figure 1.2: Global electric vehicle sales growth, 2015-2023 (Source: International Energy Agency, Global EV Outlook 2023)

The environmental benefits of electric vehicles are substantial, particularly when powered by low-carbon electricity sources. Even when accounting for the current global electricity mix, which still includes significant fossil fuel generation, EVs typically produce lower lifecycle greenhouse gas emissions than comparable ICE vehicles. As electricity grids continue to decarbonize, this advantage will only increase. Additionally, the shift of emissions from millions of individual tailpipes to centralized power plants offers significant air quality benefits in urban areas and creates opportunities for more efficient pollution control.

However, the transition to electric mobility faces many of the same infrastructure challenges that the early automobile industry encountered. Just as the widespread adoption of ICE vehicles required the development of a comprehensive network of gas stations, repair facilities, and roads, the EV revolution depends on the deployment of charging infrastructure, grid upgrades, and maintenance expertise. These parallels suggest that while the challenges are significant, they are not unprecedented and can be overcome through coordinated investment and policy support.

1.3 EU mandate

The European Union has established one of the world's most ambitious regulatory frameworks to accelerate the transition to electric mobility as part of its broader climate strategy. The cornerstone of this approach is Regulation (EU) 2019/631, which sets CO_2 emission performance standards for new passenger cars and light commercial vehicles. This regulation has been progressively strengthened, culminating in the European Commission's "Fit for 55" package proposed in July 2021 and subsequently adopted, which aims to reduce net greenhouse gas emissions by at least 55% by 2030 compared to 1990 levels.

The most transformative element of this regulatory framework is the mandate that effectively prohibits the sale of new internal combustion engine vehicles in the EU from 2035 onward. Specifically, the regulation requires a 100% reduction in CO_2 emissions from new cars and vans by 2035 compared to 2021 levels, which in practice means that only zero-emission vehicles—battery electric or hydrogen fuel cell—can be sold as new vehicles after this date. This represents a clear and unambiguous signal to the automotive industry, infrastructure developers, and consumers about the direction of transportation policy in Europe.

The EU's approach includes intermediate targets to ensure a gradual transition: a 55% reduction in car emissions and a 50% reduction in van emissions by 2030 compared to 2021 levels. These targets are accompanied by incentive mechanisms for zero- and low-emission vehicles, penalties for manufacturers that exceed fleet-wide emission targets, and provisions for reviewing the effectiveness of the regulation.

This regulatory certainty has already catalyzed significant investment in electric vehicle production and charging infrastructure across Europe. Major automotive manufacturers have announced accelerated timelines for electrifying their fleets, with many planning to phase out ICE vehicle production well before the 2035 deadline. The mandate has also spurred innovation in battery technology, charging solutions, and vehicle design as companies compete to position themselves advantageously in the emerging electric mobility ecosystem.

The EU's approach demonstrates how regulatory frameworks can create the conditions for market transformation by providing clear, long-term signals that enable businesses and consumers to plan and invest with confidence. By establishing a definitive end date for new ICE vehicle sales, the EU has moved beyond incremental improvements to fossil fuel efficiency and committed to a fundamental technological transition in personal transportation.

1.4 Public Electric Charging Locations

Public electric vehicle charging infrastructure represents a critical enabler for widespread EV adoption, particularly for urban residents without access to private charging facilities. Unlike the relatively standardized experience of refueling an ICE vehicle, EV charging encompasses a diverse ecosystem of technologies, power levels, connector types, and usage patterns that reflect the evolving nature of electric mobility.

Modern EV charging infrastructure is typically categorized by power output, which directly affects charging speed:

- **Level 1 (Slow) Charging:** Utilizing standard household outlets (typically 2.3-3.7 kW), these chargers add approximately 10-20 kilometers of range per hour of charging. While inadequate as a primary charging solution for most users, they serve as emergency options or for overnight charging in residential settings.
- **Level 2 (Medium) Charging:** Operating at 7-22 kW, these AC chargers can fully replenish most EV batteries in 4-8 hours, making them suitable for workplace, residential, and destination charging where vehicles are parked for extended periods. They represent the majority of public charging points in most regions.

- **DC Fast Charging:** Delivering 50-350+ kW of power directly to the vehicle's battery, these stations can provide an 80% charge in 20-40 minutes for compatible vehicles. They are strategically deployed along major travel corridors and in urban centers to enable long-distance travel and quick top-ups for those without home charging access.

The physical infrastructure of charging stations varies considerably, from simple wall-mounted units to sophisticated multi-port stations with integrated payment systems, user authentication, load management, and network connectivity. Most public charging stations are connected to management platforms that enable remote monitoring, usage tracking, and dynamic pricing, creating a digital layer that enhances the user experience and operational efficiency.

Connector standards have evolved regionally, with the Combined Charging System (CCS) emerging as the dominant standard in Europe and North America, CHAdeMO prevalent in Japanese vehicles, and GB/T standard in China. The industry has been moving toward greater standardization, with many newer vehicles adopting CCS, though legacy systems will remain in operation for years to come.

Despite significant growth in charging infrastructure, access remains unevenly distributed, with substantial gaps in many urban residential areas where residents rely on street parking and lack access to private charging facilities. This "charging desert" phenomenon disproportionately affects apartment dwellers and those in older urban neighborhoods, creating a potential barrier to equitable EV adoption. Studies indicate that approximately 30-40% of European urban residents lack access to private parking where home charging could be installed, highlighting the critical importance of public charging infrastructure for these populations.

The current trajectory of EV adoption, accelerated by regulatory mandates like the EU's 2035 ICE vehicle phase-out, will require a massive expansion of charging infrastructure. The European Commission estimates that up to 3.5 million public charging points will be needed across the EU by 2030 to support the expected growth in electric vehicles—a nearly tenfold increase from current levels. This expansion must be strategically planned to ensure equitable access, grid compatibility, and alignment with mobility patterns.

1.4.1 EV Charging Location Placement problem

The strategic placement of EV charging infrastructure represents a complex optimization challenge with significant economic, technical, and social dimensions. Unlike traditional gasoline stations, which primarily serve vehicles passing through specific corridors, EV charging locations must accommodate diverse charging behaviors that include destination charging, opportunity charging, and en-route fast charging for longer journeys.

From an economic perspective, charging infrastructure deployment requires substantial capital investment—ranging from approximately €2,000-5,000 for a basic AC charging point to €100,000 or more for a high-power DC fast charging station, including installation and grid connection costs. These investments face uncertain utilization rates during the early adoption phase, creating challenging business cases that often require public subsidies or innovative business models to become viable. The long-term profitability of charging operations depends on multiple factors including utilization rates, electricity costs, maintenance requirements, and the ability to capture value through charging fees or complementary services.

The technical challenges of charging infrastructure placement extend beyond the charging equipment itself to include grid integration considerations. High-power charging stations can place significant demands on local distribution networks, potentially requiring costly grid upgrades or reinforcement. Strategic placement that aligns with existing grid capacity can substantially reduce deployment costs and timelines. Additionally, the temporal distribution of charging demand throughout the day creates opportunities for smart charging systems that can help flatten load profiles and reduce peak demand, potentially providing valuable grid services through vehicle-to-grid (V2G) capabilities.

The social dimension of charging infrastructure placement involves ensuring equitable access across different communities and addressing the specific needs of various user groups. Charging deserts in urban residential areas without private parking facilities represent a particular challenge that requires innovative solutions such as curbside charging, integration with street lighting, or community charging hubs. Public charging infrastructure must also be accessible to users with disabilities and designed with safety considerations for all users, particularly in 24-hour self-service environments.

Data-driven approaches to charging infrastructure planning have emerged as essential tools for optimizing placement decisions. These approaches typically incorporate multiple data sources including:

- Traffic flow patterns and vehicle dwell times
- Demographic data and EV adoption projections
- Land use and points of interest
- Existing charging infrastructure distribution and utilization
- Electrical grid capacity and upgrade costs
- Temporal patterns of mobility and energy demand

By integrating these diverse datasets, planners can develop sophisticated models that predict charging demand with high temporal and spatial resolution, enabling more efficient infrastructure deployment that maximizes utilization while minimizing costs. These models become particularly valuable when they can account for the dynamic nature of EV adoption and changing mobility patterns, allowing for adaptive planning that evolves as the market matures.

The ability to accurately forecast charging demand at specific locations and times represents a critical capability for infrastructure planners, grid operators, and charging network developers. High-resolution temporal forecasting enables not only more strategic infrastructure placement but also more efficient operation through predictive maintenance scheduling, dynamic pricing strategies, and load management. This creates a compelling case for advanced modeling approaches that can capture the complex interplay of factors influencing charging behavior and translate them into actionable insights for infrastructure development.

1.5 Goals of the thesis

This thesis aims to address the critical challenge of predicting electric vehicle charging demand with high temporal and spatial resolution, focusing specifically on the context of Prague, Czech Republic. By developing a data-driven approach to forecasting charging patterns, this research seeks to provide valuable insights for infrastructure planners, grid operators, and policymakers navigating the transition to electric mobility.

The specific objectives of this research include:

- Developing a machine learning model capable of predicting average power consumption (APC) at potential charging locations throughout Prague, with hourly temporal resolution
- Identifying and quantifying the influence of various spatial and temporal factors on charging demand, including proximity to points of interest, population density, mobility patterns, and temporal variations by time of day, day of week, and season

- Creating a methodological framework that can be adapted to other urban contexts with similar data availability constraints
- Providing actionable insights for strategic charging infrastructure deployment that maximizes utilization while ensuring equitable access
- Contributing to the broader understanding of electric vehicle charging behavior in Central European urban contexts, where limited research has been conducted compared to Western European and North American settings

This research employs a novel neural network architecture with latent profiles to capture the complex patterns in charging demand, leveraging diverse datasets including charging session records, points of interest, population statistics, and mobility data. The approach balances predictive accuracy with interpretability, allowing for insights into the underlying factors driving charging behavior.

By addressing these objectives, this thesis contributes to the practical challenge of enabling the transition to electric mobility through strategic infrastructure development, while advancing the methodological approaches available for charging demand forecasting in urban environments.

Related research 2

This section mentions relevant literature that focuses either on the very same issue. Or topics close to predicting EV charger demand. There are not that many papers focusing on our specific issue however there is a lot of knowledge ihdden inside of them. The papers analyzed in this chapter provide insight into how similar issues were tackled. And on what does the research focus on.

First, from an outside perspective. Issues and topic of the papers will be explored. What outcome were they focused on. And then an inside look, into what research approaches they took and what methods were used.

Because this is spatial data science. Most of the papers are very practical in a sense that they work with real datasets. And each country and even city has different data gathering culture and data availabilty. The research is tightly connected to what data is available. No relevant paper for our issue was found focusing on Prague.

2.1 Issue addressed

- EV charger demand prediction
- EV charger use anaylsis
- Charging infrastructure planning and optimization
- Digital twin

2.2 Research approaches

Starting from simple statistics. Then menitoning agent simulations. And end with ML.

Research applies itself to all sorts of EV stuff. Starting from understanding coverage of existing EV chargers.

2.2.1 Understanding EV Charger Use by Data

To understand why certain chargers are being utilized the way they are. Research utilizes traditional and Bayesian statistics. As the person that plans charging connector (CP) it is good to have insight into what influences charging demand. That is, why a certain charger is utilized. And what factors contribute to it. So far, we dont care about expansion of the infrastrcture. But could provide insights that allow to place new chargers more strategically.

[1] gathered counts of types of point of interest (CD)[1] near every charger of interest [2] from Open Street maps. Then for each charger computed its utilization. Which is its average daily power consumption. Then used linear regression to test which of the category of point of interest (CD) contributed to the consumption. The study had some statisficaly significant results. They also trained neural network model for capturing non linear relationships. User can use the model to select any point where a charging connector (CP) might be placed and see its estimated utilization and evaluate worthiness of placement. However it does not work with other chargers in the area and does take into account charger density. The paper has identified that certain categories of point of interest (CD)[3] are correlated with charger demand.

[1]: Hecht et al. (2024), 'Global Electric Vehicle Charging Station Site Evaluation and Placement Based on Large-Scale Empirical Data from Germany'

1: types like shops, sport areas, schools, offices

2: Search radius of 2000m around each charger with linear decrease in importance in relation to distance

3: CD data obtained from OpenStreetMaps

[2] uses log-Gaussian Cox process. Which is a statistical model that can handle

[2]: Dong et al. (2019), 'Electric Vehicle Charging Point Placement Optimisation by Exploiting Spatial Statistics and Maximal Coverage Location Models'

dependence between points on a map (EV chargers). It has identified that workplace population and traffic flow are positively related to demand ofcharging connector (CP) while commerce is in a negative relation. [4].

4: <workplace population, traffic flow, commerce description here>

Simulation vs Model

Model is concerned with representing a system of interest. Purpose of it is so it matches its real system in some behaviour while being simpler than reality. There are many ways how to construct models. From simplest small replicas (small car models to use in wind tunnel) to complicated mathematical and software ones.

Simulation is then a use of model to try to replicate some real system of interest to gain insights how a real system might behave .

An example can be a toy car and a ramp which would be our model. Then pushing the car off the ramp to see how far it is able to land. This is much less expensive than pushing real cars of larger ramps.

[3]

2.2.2 Simulations and EV Charger Use

Trafic models are used to estimate charging connector (CP) demand. They are a mathematical model of real-world trafic. And they can estimate charging demand by introducting EV vehicle agents into the simultion. Those agents navigate replica of real road network. While also being able to emulate traffic jams, alternative path finding. Negative of these models is that they are harder to develop, fine tune to match real situation. And performance costly. But once developed and validated with real-world. They can be highly utilized.

Micro,Macro,Meno-Scopic

Traffic models can be separated into three groups according to [4][5]([6] also writes about splitting simmulation models into 3 groups with a more theoretical view):

- **Microscopic** - Simulation of individual agents/vehicles. Can be as detailed as emulating accelerating, lane changing, turning. Higher fidelity isuUsable for example when designing intersection with light signals optimizing for traffic or pedestrian flow. But the model can also have less detailed fidelity for vehicles. Which simplifies the simmulation and enables larger area of interest. Like cities or even countries.
- **Mesoscopic** - Provide less fidelity than microscopic. improve
- **Macroscopic** - Work with aggregate information. Work more on analytical model and proven mathematcial relationships between traffic flow. improve

[7]: Community (2025), *MATSim*

Slight foreshadowing but this approach was not taken in this thesis however it was considered.

Matsim [7] is Agent based model (ABM) . It's OpenSource software licensed under GPLv2, implemented in Java. And started in 2004 in Zurich. It falls into category of microscopic models due to it being able to simluate individual agents. The model simmulates simplifed version of vehicle movement and traffic which is needed to correctyl simulate entire city including congestions. The simulation usually concerns only one day. To create such a model the following are needed:

- **Road/communication Network** - Spatial data for area of interest reagarding: roads, intersections, tram or train tracks etc. Obtained from technical data published by goverments or third party map providers. (Mainly OpenStreetMaps)

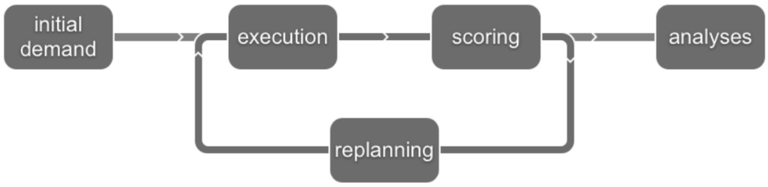


Figure 2.1: Stages of matsim simulation. [8]

- **Population** - Agents with their day plan, where are they going to live and commute to work. Those are generated with help of expert knowledge and travel diaries survey.
- **Facilities** - Locations where activities can be performed. Obtained from third parties like OpenStreetMaps.

The agents plan assigns unique daily schedule for each agent in the population. See 2.1 This estabilished the initial demand. During the execution step, Matsim simulates agents commute and tries to satisft the agents schedule. The agent might fail to be on schedule due to road or public transport congestion from other agents. After finishing the simulation the agents actual daily schedule is scored. Penalizing arriving at work late or being stuck for too long in traffic giving agents lover score. While agents who got where they wanted qucikly and without trafic jams get better score. During replanning agent schedule can be modified so it either adjusts the agents schedule, like less time spent home or taking a different route to work. When the simulation is sufficiently optimized the results of it can be then used for various analysis as the output of the modle is detailed log of agent activities.

To create the agents schedule, day intentions need to be provided to the model. Those can be either manually crafted. With the increased availabiliy of census data, data driven approach can be taken. This trades away the possibility to experiment with different kinds of populations and policies[5] but simplifies the process of creation where only data is needed and not expert knowledge on human behavior. [9] models multiple probability densities for what activity will an agent do, how long it will take and what activities are available in the area. This utilizies several datasets. New schedules can be generated from the model by sampling from the distributions.

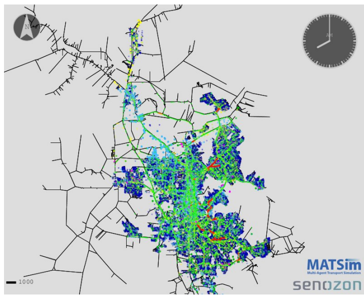


Figure 2.2: Matsim, Joinville example. Model intended to help the city encumbered by high traffic volumes. `https://matsim.org/gallery/joinville/`

5: like exmaining effect of different school time start on traffic

[9]: Drchal et al. (2019), 'Data-Driven Activity Scheduler for Agent-Based Mobility Models'

Finaly, moving to use of Matsim for ev scenarios. [10] studies the impact of EVs on the electric id and electricty production in Croatia. By having a simmulation of power production network for the whole country as well. They first have a simmulation regarding the current state and how that impacts the electrical grid. This sets the baseline and also allows to correctly calibrate the model. Then they increase the ev adoption and see what is the hourly energy consumption of EVs and their impact on the grid.

[10]: Novosel et al. (2015), 'Agent Based Modelling and Energy Planning – Utilization of MATSim for Transport Energy Demand Modelling'

Talk about arup

Stochastic models provide a way for modeling charging behavior of electric vehicles by capturing randomness and uncertainty in travel patterns. They work with more simplified representation fo reality compared traffic models[6]. These models typically employ probability distributions to characterize variables such as departure times, travel distances, parking durations, and charging decisions. By sampling from these distributions using techniques like Monte Carlo simulation, they generate synthetic power demand profiles.

6: Of course trafic models have large part of stochasticity inside themselves

[11] presents an example of stochastic modeling applied to EV charging behavior. Their approach uses a non-parametric copula function to model the dependence structure between six variables: departure time, number of journeys, and total distance traveled across two consecutive days. For this they utilize real-world GPS data collected from electric vehicles. Their model simulates complete journey schedules for individual vehicles and implements a probabilistic charging decision model at each destination, conditioned on the state of charge, parking time, and journey number. The approach captures the variability in charging

[11]: Brady et al. (2016), 'Modelling Charging Profiles of Electric Vehicles Based on Real-World Electric Vehicle Charging Data'

behavior by incorporating factors such as battery characteristics and probabilistic charging point availability.

[12]: Ul-Haq et al. (2018), 'Probabilistic Modeling of Electric Vehicle Charging Pattern in a Residential Distribution Network'

probably skip, the paper is too hard

Paper: [12]

- tests policy changes
- estimate ev charging pattern based on residents activity patterns
- utilizes markov chain to estimate next trip departure time and travel distance
- does not have charging data, so from simmulation of a day it guesses how much the ev needs to be charged
- graphical model with good diagram/description

[13]: Powell et al. (2022), 'Scalable Probabilistic Estimates of Electric Vehicle Charging given Observed Driver Behavior'

Paper: [13]

- groups drivers hiearrchicaly by charging behaviour
- utilization of actual charging data, charging sessions data. In the session, driver id is available. So drivers charging behaviour can be observed.
- utilize monte carlo simulation. Sample from learnt prob distributions.
- validated against real data
- $P(s, z, G) = P(s|z, G)\ P(z|G)\ P(G)$

[12]: Ul-Haq et al. (2018), 'Probabilistic Modeling of Electric Vehicle Charging Pattern in a Residential Distribution Network'

[14]: Zhang et al. (2023), 'Charging Demand Prediction in Beijing Based on Real-World Electric Vehicle Data'

[15]: Zhang et al. (2024), 'An Urban Charging Load Forecasting Model Based on Trip Chain Model for Private Passenger Electric Vehicles'

- [12]
- [14]
- [15]

mention difference forecasting vs modelling

2.3 Forecasting

ML for predictions, not explainable. Just learning on data

2.4 Infrastructure Planning

Charger placement, optimization problem how to cover certain areas. ILP and others used

[16]: Pekárek (2017), 'A Model of Charging Service Demand for the Czech Republic'

[17]: Elomiya et al. (2024), 'An Advanced Spatial Decision Model for Strategic Placement of Off-Site Hydrogen Refueling Stations in Urban Areas'

[18]: Uglickich et al. (2025), 'Poisson-Based Framework for Predicting Count Data'

[19]: Shang et al. (2021), 'Estimating Building-Scale Population Using Multi-Source Spatial Data'

2.5 Czechia and Prague Relevant Research

No research paper has been found with theme of EVs and Czechia/Prague. So only relevant stuff is gonna be mentioned

[16] [17] [18]

2.6 Data sources and transformations

estimate population

Predicts density of each building per $100m^2$ living area. With $R^2 = 94\%$. [19]

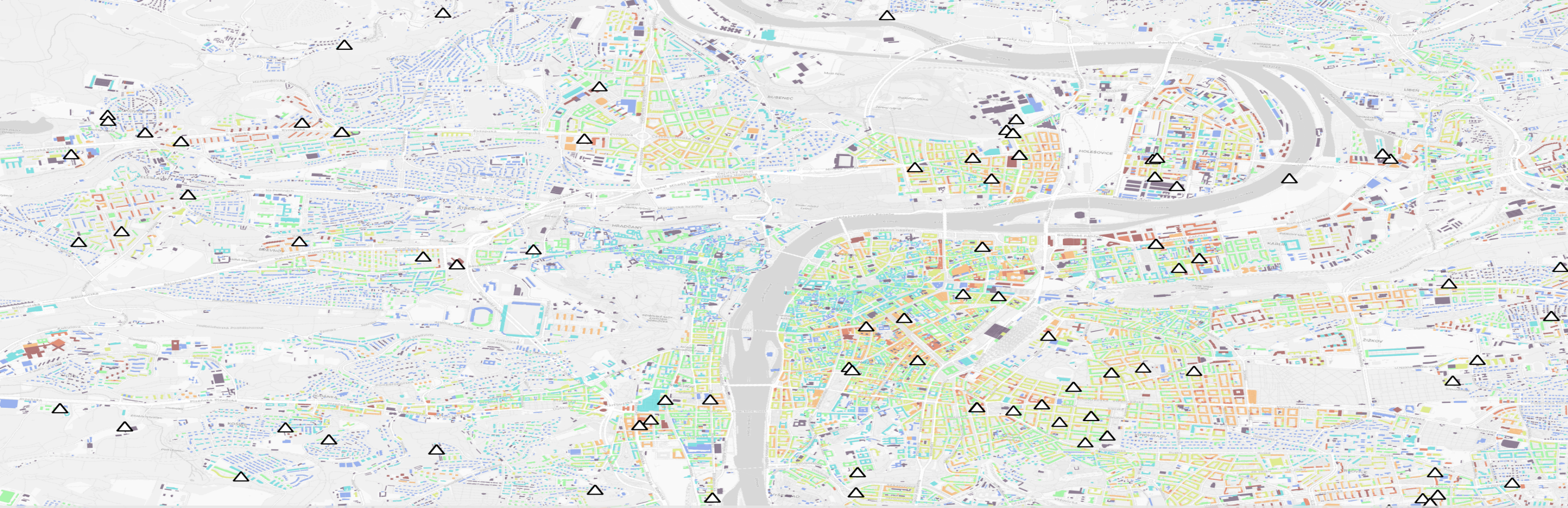
2.7 Discussion

From observed papers regarding simulation to obtain some estimate of charging demand. Either, in absence of data, expert knowledge is utilized to create simmulations. Which lean more on engineering complexity. Due to need to simmulate some percent of city/country population. But they might not always be feasible to develop as well as calibrating. Which is to ensure the simulation matches real world. But then they have large utility and flexibilty. For example examine policy changes.

When data is available, turn to more data based methods is more common. Instead of microscopic simulations. Stochastic properties of people are obtained from the data or estimated to match the data.

2.8 Research gaps

Mention how the relevant research is helpful for us. Also how noone has studied Prague and that it has different data landscape. And so we have to do a new approach.



3 Data

In this chapter types data will be introduced, how they are classified and data formats in which they were obtained. Chargers, charging session dat and some lightweitght ontology will be described. Then followed with transformations of charging sessions into average power consumption with temporal pattern (APC) which will be target variable in our model.

Then data that are considered and used to predict average power consumption with temporal pattern (APC) will be introduced together with their transformation into a desirable form for our model.

Some of the data have not been used [1] but are part of data landscape considered and hold usefull information on what was thought might have an effect but in fact did not have.

1: The inclusion of these data did not lead to lower validation loss of our model

3.1 Types of data

- **spatial**(geospatial) data are those which have assigned position in real world and are invariant in some timeframe [2]. Such data are locations of charging station (CS), administrative boundaries, road network, buildings.
- **temporal** data are characterized by their variation over time without specific geographical coordinates. In this study, these include charging session durations, energy consumption patterns throughout the day, historical charger utilization rates, and seasonal variations in charging demand.
- **spatio-temporal** data incorporate both location and time elements, providing insights into how phenomena evolve across space and time. Examples relevant to this research include mobility patterns of people, real-time charger availability, and dynamic variations in APC across different city zones during different hours of the day.

2: Nothing is permanent but in the time window of this thesis they have been invarinat

3.2 EV Chargers and Charging Sessions

To introduce our problem domain and a possiblity to connect the data and how they relate. An simple charger ontology is introduced [3]. Inspired by AURORAL EV-charger Ontology [20] an ontolgoy of EV charging is introduced. The scope of it is to aide in understanding of the domain in this thesis.

3: Ontology describe subjects of some system and a way how they are related together.

[20]: (), *Ontology Documentation Generated by WIDOCO*

The model matches data obtained from PRE.

Below is description of individual parts of the charging ontology as visible in 3.1.

Title image is a map o f Prague with all chargers denoted as triangles in available datasets. The layer below displays all buildings in Prague with color being the number of floors

charging station (as visible in 3.2) is equipment that connects an EV to a source of electricity to recharge them. A charging station typically consists of physical infrastructure including power conversion hardware, connectivity modules, authentication systems, and user interfaces. Charging stations vary in their power delivery capabilities, ranging from slow AC chargers (3.7-22 kW) commonly found in residential and workplace settings to fast DC chargers (50-350+ kW) deployed in public corridors and commercial hubs. Within our dataset, charging stations from PRE's network predominantly consist of public AC and DC installations distributed throughout Prague's urban and suburban areas.

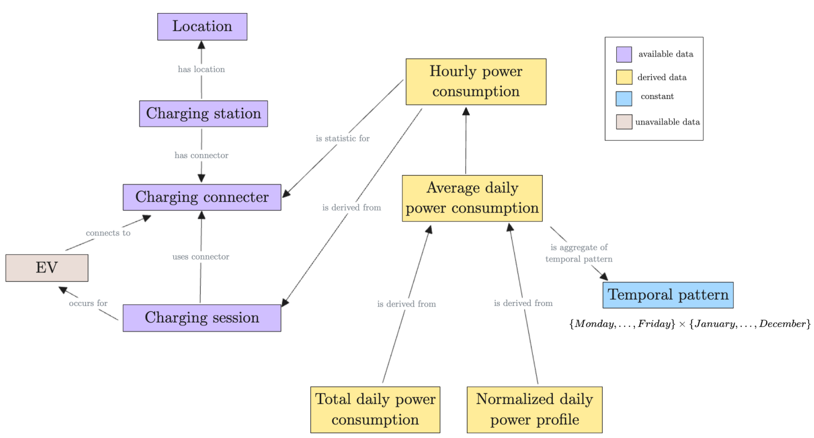


Figure 3.1: A charging ontology

charging connector (as visible in 3.3) one or many are part of a CS. These physical interfaces allow for the actual connection between the vehicle and the charging infrastructure. Connectors follow different standards depending on region and charging speeds. Each connector type supports specific charging protocols and power levels. In our studied network, the majority of charging stations feature multiple connectors (typically two to four), enabling simultaneous charging of different vehicles and supporting various connector standards to accommodate the heterogeneous EV market.

Figure 3.2: Picture of charging station. It has one connector on each of its sides. One of which has charging cable attached.

charging session occurs when an EV arrives at a CS and connects to a CP. This interaction initiates a session that is logged by the CS together with various parameters including connection time, disconnection time and total power consumed. The charging session represents the fundamental unit of analysis in our study, as it captures both temporal patterns (duration, time of day, day of week) and energy consumption behaviors. Our dataset contains records of these sessions across PRe chargers Prague's charging network over a multi-year period.

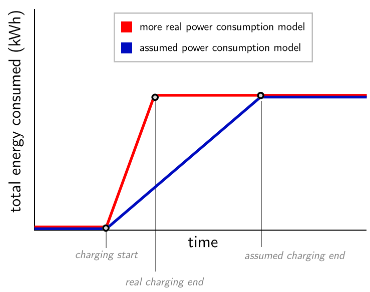
Figure 3.3: View of 1 of the 2 charging connectors the CS has

location denotes the geographical position where the charger is installed. This spatial attribute is helpfull to our analysis framework as it allows for correlation between charging demand and various features of the surrounding environment.

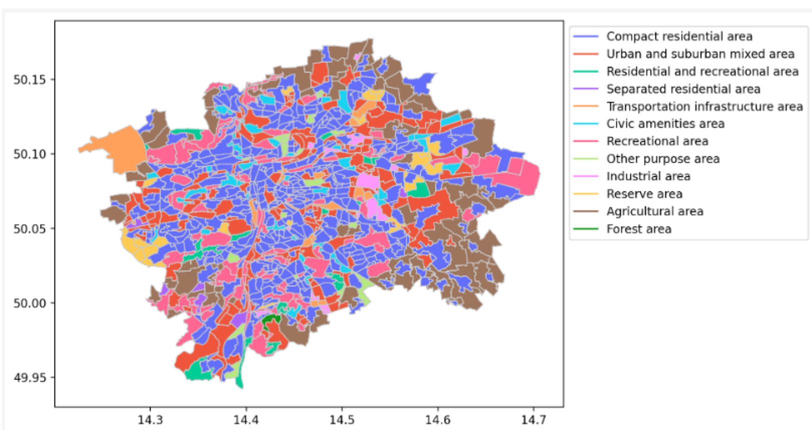
3.2.1 Power consumption assumption

Before we describe how charging session (CSS) have been processed into hourly power consumption and average power consumption with temporal pattern. In our model, we take simplyfing assumption on how actual power consumed during chargins session is consumed. Since actual power consumed can vary during charging of the electric vehicle. Most important is, the vehicle from connection time is being charged with the maximum power the electric vehicle can handle and charging connector can provide. This leads to charging session being split into two parts. First is charging part, that is when power is being delivered to the vehicle. And idle period where the vehicle has been fully charged and no power is being consumed[4]

4: Some charging station providers financially penalize this period of time, as an other EV could have been charging. This can lead to improved availabilty of charging station.



3.3 Population numbers (ZSJ)



3.4 Points of Interest

3.5 People Mobility

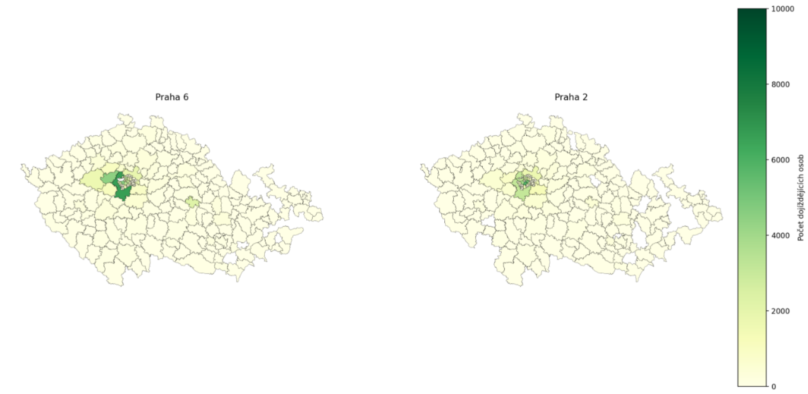


Figure 3.4: Chapter content overview.

3.6 Mobility Survey - cesko v pohybu

Figure 3.5: Chapter content overview.

—

Talks about what was extracted from datasets in 3

3.7 Spatial data

3.7.1 ZSJ Type and Population

3.7.2 Points of Interest

Has statistically significant results for certain PoI [1][2] ([2] uses Gaussian cox model, [1] uses Neural networks and linear regression)

[1] states that radius of relevant PoIs is 2000metres. The research stated that the distance is sensible. One solution is to just count all the PoIs in the radius but as they are more far away from the CP their relevance might decrease. The [1] thus uses importance factor for pair of PoI and CP.

$IF(PoI_i, CP_k) = max(r - d_{\text{sphere}}(PoI_i, CP_k), 0)$

Buildings (OsmPoisPbf)

Public Ammenities (OSMOX)

how osmox works

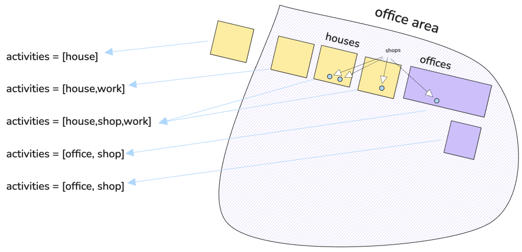
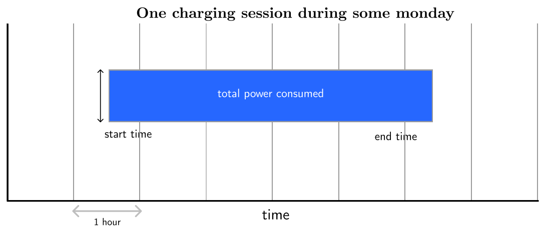
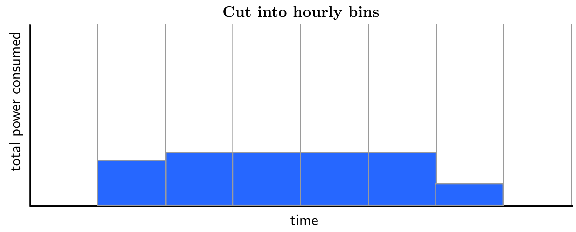


Figure 3.6: osmox

3.8 Charging profiles





3.9

Research and implementation 4

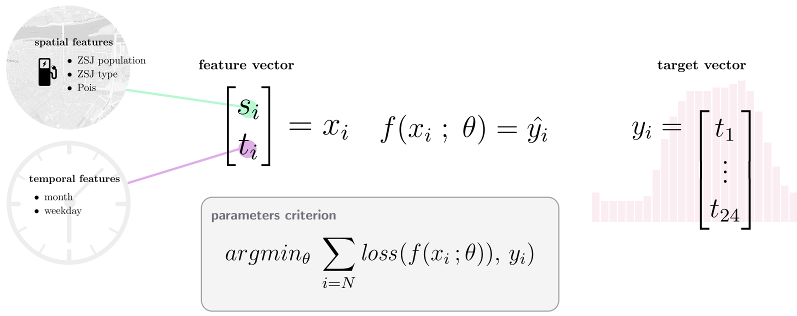


Figure 4.1: Problem approach overview

In this chapter, the problem will be formulated, and then an approach of solving it will be presented. The choice for the ML model will be discussed as well as its interpretable structure. A way of processing the data, splitting it into training and validation sets will be presented. And then a way to evaluate the model in comparison to other ML models will be described.

4.1 Problem statement

We will be interested in estimating average power consumption with temporal pattern (APC) from some set of features. And will like to know if our data-driven approach is viable.

To formulate how the model will look like, the model f_θ will execute the following mapping:

$$f_\theta : X \to Y$$

Where X will be the set of all feature vectors and Y will be the set of all targets. The available data are split into *training test* and *validation* sets. The first available at training time with goal of the training algorithm to minimize the empirical risk. How those sets are constructed is explained in 4.4

$X \in \mathbb{R}^M$ and $Y \in \mathbb{R}^{24}$. And $|A| = |B| = P$.

The model function f will have trainable parameters θ. And we will be interested in finding such θ that minimizes the empirical risk:

$$loss_{total}(f, \theta) = \frac{1}{P} \sum_{(x,y)\in\mathcal{T}} \alpha \cdot a_x + \beta \cdot b_x \tag{4.1}$$

Where:

$$\begin{aligned} a_x &= loss_{power}(\left\|f(x_i;\theta)\right\|_1, \left\|y_i\right\|_1) \\ b_y &= loss_{norm}(\frac{f(x_i;\theta)}{\left\|f(x_i;\theta)\right\|_1}, \frac{y_i}{\left\|y_i\right\|_1}) \end{aligned} \tag{4.2}$$

where $loss_{power}$ and $loss_{norm}$ will be individual loss functions. Because we will be interested in model performance in estimating total daily power consumption (TDPC) and normalized daily power consumption (NDPC). This will match with the chosen NN model architecture that will be discussed further in this chapter.

The use case of the model is to answer the EV charger planners' question: what will be the charger's power consumption if he decides to place a new charging station (CS) at a new place in Prague? And he is interested in its behavior given the temporal pattern[1].

1: e.g., how would a new charger at place x perform on Thursdays of February

4.2 Model features and feature engineering

The feature vector of the model consist of **temporal** and **spatial** parts. Features regarding the charger's capabilities should be present as well, but that will be a limitation of the current chargers dataset that we will work with. The most useful will be the charger's maximum power output. Since that will most certainly influence the charger's average consumption. From this absence, the normalized daily power consumption (NDPC) will also be in our interest to estimate. Since the total power consumed might not have that big of an influence on it.

The feature items will fall into two categories divided by the data type. Either they will be categorical or numerical. If they will be categorical, they will be transformed with one-hot encoding. That is, given a category with n items to transform the feature into n binary features, where each binary feature will correspond to one of the possible values. For each observation, exactly one of these binary features will have the value 1, indicating the presence of that categorical value, while all others will be 0.

For example, if we have a categorical feature "day of week" with 7 possible values (Monday through Sunday), one-hot encoding transforms this into 7 binary features: "is_Monday", "is_Tuesday", etc. If an observation occurs on Wednesday, then the "is_Wednesday" feature would be 1, while all other day features would be 0. This transformation allows the model to properly handle categorical variables without imposing an arbitrary ordinal relationship between category values.

In our case, categorical features like the day of the week, month, and location characteristics will be one-hot encoded before being fed into the model. This will ensure that the model can effectively learn from these categorical variables without the constraints of numerical ordering.

Numerical features, on the other hand, will be standardized by subtracting the mean and dividing by the standard deviation to ensure all features will be on a comparable scale. This normalization process will prevent features with larger scales from dominating the learning process and will help achieve faster convergence during model training.

The feature vector will be constructed in the following way:

$$x_i = \begin{bmatrix} s_i^T \\ t_i^T \end{bmatrix} \tag{4.3}$$

spatial features s_i will be a vector of spatial features which contents will be described in 4.1 together with how this feature will be encoded.

$$s_i = \begin{bmatrix} s_i^1 \\ \vdots \\ s_i^R \end{bmatrix} \tag{4.4}$$

temporal features t_i will be a vector of temporal features which contents will be described in 4.2 together with how this feature will be encoded.

$$t_i = \begin{bmatrix} t_i^1 \\ \vdots \\ t_i^P \end{bmatrix} \tag{4.5}$$

- introduce latent profiles neural network model
- mention training procedure in all the detail, because everyone does this
- train test data split (based on location, to avoid double positions)
- loss function

Table 4.1: Overview of spatial features used in the feature vector.

Index	Name	Type	Value from	Additional processing
s_1	ZSJ population	numeric	charger in ZSJ polygon	normalization by the polygon area
$s_{2:10}$	ZSJ type	categorical	charger in ZSJ polygon	one-hot encoding
s_{11}	ZSJ number of addresses	numeric	charger in ZSJ polygon	normalization by the polygon area
s_{12}	Number of people commuting into the district from inside Prague	numeric	charger in the district polygon	normalization by the polygon area
s_{13}	Number of people commuting into the district from outside of Prague	numeric	charger in the district polygon	normalization by the polygon area
$s_{14:162}$	Points of Interest	numeric	number of PoIs by euclidean distance	importance calculation (value of single PoI is 1 if its distance from charger is 0, 0 if it is of distance 2km or further)

Table 4.2: Overview of temporal pattern used in feature vector.

Index	Name	Type	Value from	Additional processing
$t_{1:7}$	day of the week	categorical	average power consumption with temporal pattern (APC)	one-hot encoding
$t_{8:19}$	month	categorical	average power consumption with temporal pattern (APC)	one-hot encoding

- parameter tuning

4.3 Architecture of the Latent Neural Network

The formulation of the machine learning problem provides us with ability of many solutions. Mainly from the class of nerual networks.

Neural networks are computational models. They consist of layers of interconnected nodes or "neurons" that process information. A typical neural network contains an input layer that receives data, one or more hidden layers that perform computations, and an output layer that produces the final result. Each connection between neurons has an associated weight that is adjusted during the training process. Information flows through the network via activation functions, which introduce non-linearity and allow the network to learn non-linear patterns. The training process involves feeding the network with labeled examples and using optimization algorithms, typically variants of gradient descent, to minimize a loss function by adjusting the weights. Backpropagation is the primary algorithm used to calculate gradients and update weights efficiently.

This leads us to the proposed neural network with latent profiles. Diagram of the network is visible at 4.2. Before diving into the detailed network architecture a high level overview. The goal of the network is to construct for its internal use K [2] latent profiles. Which is a matrix R where $R \in \mathbb{R}^{24 \times K}$. And then estimate with use of f module their mixture into the resulting L1 normalized profile. Then there is the g module. Whose purpose is to estimate the overall day power consumption. Where the output of it is multiplied with the already mixed latent profiles.

2: K is hyperparameter

The network utilizes the following layers:

Trainable layers:

- **Fully connected (Linear transformation)**
  $\mathrm{Linear}^n_m(x) = Wx + b$
  $\mathrm{Linear}^n_m : \mathbb{R}^n \to \mathbb{R}^m \ , \ W \in \mathbb{R}^{m \times n} \ , \ b \in \mathbb{R}^m$
  W, b are learnable

  ...

- **Latent vectors (Embedding)**
  $\mathrm{LatentVec}_K = R$
  $\mathrm{LatentVec}_K : \emptyset \to \mathbb{R}^{24 \times K}$
  R is learnable

  ...

Non-parametric operations:

- **Softplus (Smooth activation)**
  $f(x) = \ln(1 + e^x)$
  $f : \mathbb{R} \to \mathbb{R}^+$

  ...

- **Normalization**
  $\mathrm{Norm}(x) = \frac{x}{\|x\|_2}$
  $\mathrm{Norm} : \mathbb{R}^n \to \{y \in \mathbb{R}^n : \|y\|_2 = 1\}$
  input dimension matches output dimension

  ...

- **Leaky-relu**
  $$\mathrm{LReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases}$$
  $\mathrm{LReLU} : \mathbb{R} \to \mathbb{R} \ , \ \alpha$ is a hyperparameter

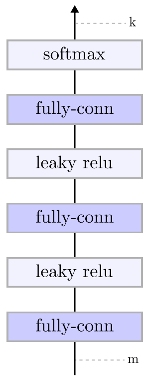
  Extension of ReLU. In this work there is no clear motivation for its use over tanh or ReLU.

Those layers joined together form 3 modules. Each of which has assigned purpose by the way the are constructed and what is their possible output value range.

Non-parametric operations:

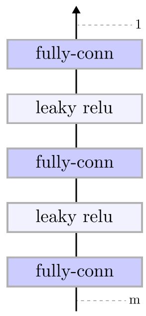
- **f module (Latent profile probabilities)**
  $f : \mathbb{R}^d \rightarrow \mathbb{R}^K$
  $f = \text{Softmax} \circ \text{Linear}_K \circ \text{LeakyReLU} \circ \text{Linear}_{64} \circ \text{LeakyReLU} \circ \text{Linear}_h$
  Where d is feature size, h is hidden size, and K is latent profiles count
  Outputs normalized weights for latent profiles

  Purpose of this module is to predict the contribution of individual profiles into the resulting output normalized profile. In other words, this module is tasked with estimating the day rhythm of the charger without the actual total power. The input is feature vector, and it is transformed by two linear layers and ReLUs. The output is vector of K values and is transformed by softmax to ensure the sum of its values equals 1.



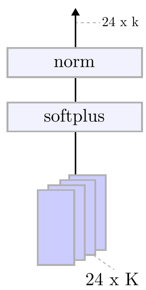
- **g module (Total power)**
  $g : \mathbb{R}^d \rightarrow \mathbb{R}$
  $g = \text{Linear}_1 \circ \text{LeakyReLU} \circ \text{Linear}_{32} \circ \text{LeakyReLU} \circ \text{Linear}_{h_g}$
  Where d is feature size and h_g is hidden size for g module

  This module predicts just the total power for the given temporal pattern and location. It consists of 3 linear layers joined with LReLU. Its input is a feature vector and outputs just one scalar. With which combined output of h and f is multiplied at the end to obtain the prediction.



- **h module (Latent profiles)**
  $h(\mathbf{x}) = f(\mathbf{x}) \cdot R^T$
  $p : \mathbb{R}^d \rightarrow \mathbb{R}^T$
  Where $R \in \mathbb{R}^{T \times K}$ is the normalized latent profiles matrix
  T is time granularity (24), K is latent profiles count

  Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.



Outputs of h and f modules are then combined like so:

$$\begin{aligned} \text{Combine}(R, p) &= \sum_{i=1}^{K} p_i \, R_i \\ \text{Combine} &: \mathbb{R}^{24 \times K} \times \mathbb{R}^{K} \to \mathbb{R}^{24} \end{aligned} \tag{4.6}$$

This is then multiplied by the scalar from g module which finally gives us the resulting prediction.

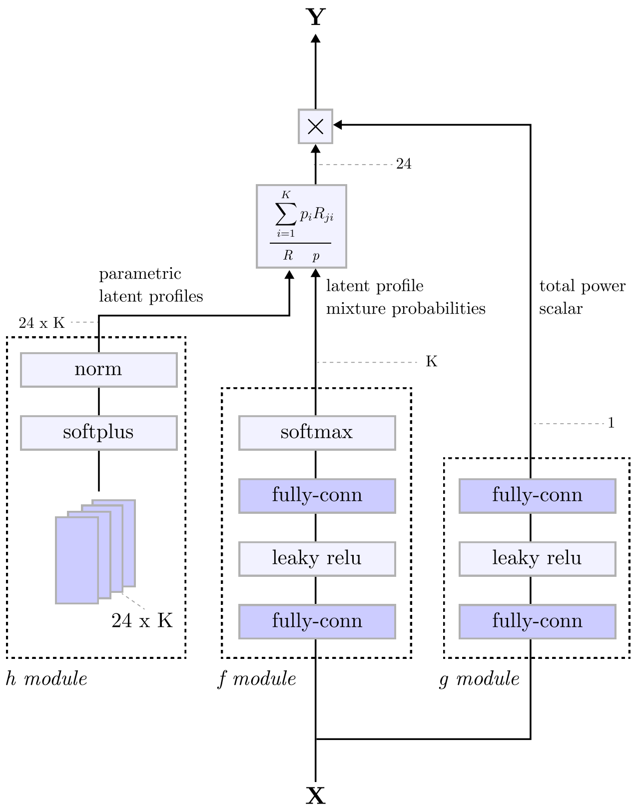


Figure 4.2: Latent neural network architecture. Light blue rectangles denote NN layers without trainable parameters. While blue denotes layers with parameters learned by SGD. Notation borrowed from Fleurets book "Little book of deep learning"

[21]: Paszke et al. (2019), *PyTorch: An Imperative Style, High-Performance Deep Learning Library*

The model is implemented in Python with use of Pytorch library [21].

4.4 Dataset splitting

To elaborate more on the way the datasets $\mathcal{T}, \mathcal{S}, \mathcal{V}$ are split. Since we are interested in modelling demand for location. We cannot just take all the average power consumption with temporal pattern. Because of an issue, where one location has multiple ABM for various temporal pattern and also because the CS has multiple CP. And there might be a high correlation between the multiple CP and our goal is for the model to be able to estimate demand for new locations. So the split is done based on location. Where features for each location can be only present in one of the sets.

$$\textit{train}: \ \mathcal{T} = ((x_i, y_i) \in X \times Y |\ i = 1, \ldots, P) \tag{4.7}$$

This dataset is used for training the model with use of stochastic gradient descent for empirical risk minimization.

$$\textit{test}: \ \mathcal{S} = ((x_i, y_i) \in X \times Y |\ i = 1, \ldots, R) \tag{4.8}$$

Is used for inspecting results of the model to see its performance on never seen data and tuning hyperparameters. And is used as an estimate of true model risk.

$$\textit{validation}: \ \mathcal{V} = ((x_i, y_i) \in X \times Y |\ i = 1, \ldots, O) \tag{4.9}$$

This dataset is used to obtain the final model risk on model with already trained parameters and chosen hyperparameters.

4.5 Training procedure

- SGD
- batches
- early stopping to avoid overfitting
-

4.6 Other models for quantitative comparison

The model is quantitatively compared with other models from the filed of machine learning. Results from the comparsion can motivate further research with the existing model or enable to quickly reject it.

One issue arrises from our custom loss, which is a mix with of two $L1$ losses. The ML models assume they want to minimize L1 loss. This might seem naive and not as a direct comparison hovewer more info will be provided in 5.

The models for comparsion are the following:

- **Average** - The most simple baseline. That is taking average over the whole datasets feature vectors and using that value as a prediction. Both the train and test datasets $\mathcal{T}, \mathcal{S}$, are utilized for this "model". It does not take into account the input and utilizes one value for any further predictions.
  We provide the implementation.

$$\begin{aligned} &f(z)^{\mathcal{D}} = \frac{1}{|\mathcal{D}|} \sum_{(x,\_) \in \mathcal{D}} x \\ &f : \mathbb{R}^m \to \mathbb{R}^n \\ &n, m \in \mathbb{N}, \mathcal{D} \in (X \times Y)^p, p \in \mathbb{N} \end{aligned} \tag{4.10}$$

- **Linear regression** - Linear model
  Implementation is taken from Python library Scikit Learn [22]

[22]: Pedregosa et al. (2011), 'Scikit-Learn: Machine Learning in Python'

$$
\begin{aligned}
&f(z) = \alpha + \beta z \\
&f : \mathbb{R}^m \longrightarrow \mathbb{R}^n \\
&\alpha \in \mathbb{R}^{n \times m}, \beta \in \mathbb{R}^n
\end{aligned}
\tag{4.11}
$$

- **XGBoost** - A scalable end-to-end tree boosting system [23]

[23]: Chen et al. (2016), 'XGBoost: A Scalable Tree Boosting System'

The perofrmacne of these models is compared with our model based on mean absolute error, mean square error on the following: average power consumption with temporal pattern, total daily power consumption and normalized daily power consumption.

Results 5

Say that the other ml models do not need more detailed loss functions since our model was not convincingly better.

In this chapter. Prediction model results will be discussed. An quantitative comparison to other models is provided.

To evaulate the quality of the model and its potential benefits.

5.1 Quantitative comparison to other models

Table 5.1: Comparison of model performance metrics including Mean Absolute Error (MAE) and Mean Squared Error (MSE) with their normalized and power variants.

Model	MAE	MAE norm	MAE power	MSE	MSE norm	MSE power
ChargingProfileModel	86.3404	0.0447	1553.9763	33729.4410	0.0057	7748441.2767
LinearRegression	107.9335	0.0962	2181.7242	41590.8896	2.9124	10658337.9148
TrainAverageModel	87.9659	0.0435	1629.9461	33973.4273	0.0057	7853195.2634
ValidationAverageModel	87.9659	0.0435	1629.9461	33973.4273	0.0057	7853195.2634
XGBoost	92.4132	0.0456	1639.3703	42282.8242	0.0065	7993035.6818

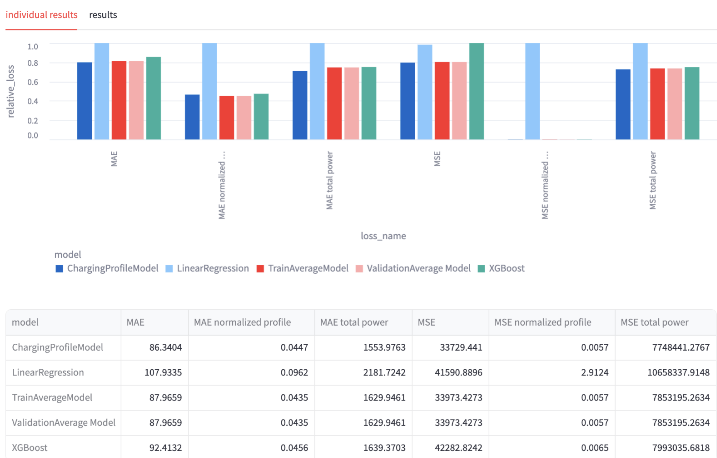
5.2 Qualitative latent profiles analysis

Inspection of the outputed latent profiles in prague together with hypotethising if that at least makes sense. Show the latent profile contribution on the training dataset because it interesting to see what it learned from the data and not necesarily rating its prediction power.

5.3 Tool for visualisation of prediction results

To be able to inspect the prediction results in spatial context a tool was also developed with use of Streamlit[24]. The dashboard is in a form of a webapplication. It consists of two screens. First one allows inspection of overall model losses as visible in 5.1. In bar charts and tabular formats. Where the bar chart shows all the types of losses relatively to each other.

[24]: (2021), *Streamlit • A Faster Way to Build and Share Data Apps*



model	MAE	MAE normalized profile	MAE total power	MSE	MSE normalized profile	MSE total power
ChargingProfileModel	86.3404	0.0447	1553.9763	33729.441	0.0057	7748441.2767
LinearRegression	107.9335	0.0962	2181.7242	41590.8896	2.9124	10658337.9148
TrainAverageModel	87.9659	0.0435	1629.9461	33973.4273	0.0057	7853195.2634
ValidationAverage Model	87.9659	0.0435	1629.9461	33973.4273	0.0057	7853195.2634
XGBoost	92.4132	0.0456	1639.3703	42282.8242	0.0065	7993035.6818

Figure 5.1: Data dashboard website showing bar chart with relative losses of the models

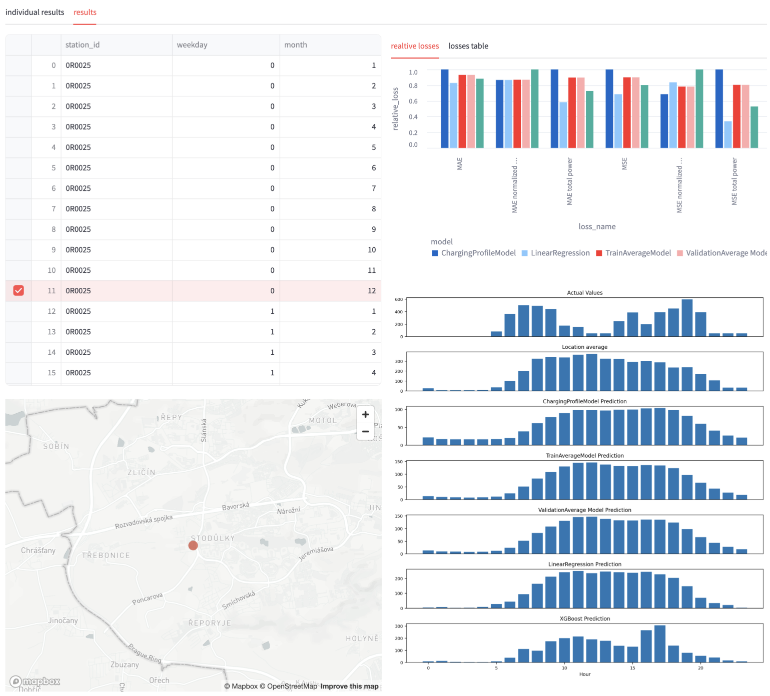


Figure 5.2: Data dashboard website showing table with data from validation data set with one selected row. Of which the prediction is computed from the main model and the other models for comparison.

The second screen displays table of all charging location together with day of the week and month. The individual rows of the table are selectable. When selection haapens location of the charger in map is shown and prediction from the model together with other models used for comparison is computed and its results are shown in a bar chart graph. Together with losses agains the original label y value.

Conclusion 6

This thesis has addressed the critical challenge of predicting electric vehicle charging demand with high temporal and spatial resolution in the context of Prague, Czech Republic. As the transition to electric mobility accelerates—driven by environmental imperatives, technological advancements, and regulatory frameworks like the EU's 2035 ICE vehicle phase-out—the strategic deployment of charging infrastructure becomes increasingly crucial. This research contributes to this challenge by developing a data-driven approach to forecasting charging patterns that can inform infrastructure planning decisions.

6.1 Research achievements

The primary contribution of this research is the development and validation of a neural network model with latent profiles capable of predicting average power consumption (APC) at potential charging locations throughout Prague with hourly temporal resolution. This model offers several advantages over existing approaches:

- **Interpretable architecture:** The latent profile structure of the neural network provides insights into underlying charging patterns, allowing for the identification of distinct temporal profiles that characterize different types of charging behavior. This interpretability enhances the model's utility for infrastructure planners by enabling them to understand not just where and when charging demand will occur, but also the nature of that demand.
- **Dual optimization:** By simultaneously optimizing for both total daily power consumption (TDPC) and normalized daily power consumption (NDPC), the model captures both the magnitude and temporal distribution of charging demand. This dual focus is particularly valuable for grid integration planning, as it provides insights into both overall energy requirements and potential peak demand challenges.
- **Feature importance insights:** The model's structure allows for the analysis of feature importance, revealing the relative influence of various spatial and temporal factors on charging demand. This analysis has identified significant correlations between charging patterns and specific points of interest, population characteristics, and temporal variables, providing actionable insights for infrastructure planning.
- **Competitive performance:** As demonstrated in the quantitative comparison presented in Chapter 6, the model achieves performance comparable to or exceeding that of alternative approaches including linear regression, XGBoost, and baseline average models across multiple evaluation metrics. This validates the effectiveness of the latent profile approach for this application domain.

Beyond the model itself, this research has made several additional contributions:

- **Methodological framework:** The research establishes a comprehensive methodological framework for charging demand prediction that integrates diverse data sources including charging session records, points of interest, population statistics, and mobility data. This framework can be adapted to other urban contexts with similar data availability constraints.
- **Data processing pipeline:** The development of a robust data processing pipeline for transforming raw charging session data into average power consumption profiles with temporal patterns provides a valuable foundation for future research in this domain.

- **Visualization tool:** The interactive dashboard developed for visualizing prediction results enables stakeholders to explore model outputs in a spatial context, enhancing the practical utility of the research for infrastructure planning applications.

6.2 Evaluation against research objectives

Reflecting on the research objectives outlined in Chapter 2, this thesis has successfully:

- Developed a machine learning model capable of predicting average power consumption at potential charging locations throughout Prague with hourly temporal resolution, as evidenced by the performance metrics presented in Chapter 6.
- Identified and quantified the influence of various spatial and temporal factors on charging demand, including proximity to points of interest, population density, mobility patterns, and temporal variations by time of day, day of week, and season. The feature importance analysis and latent profile examination provide insights into these relationships.
- Created a methodological framework that can be adapted to other urban contexts with similar data availability constraints, documented throughout Chapters 4 and 5.
- Provided actionable insights for strategic charging infrastructure deployment that maximizes utilization while ensuring equitable access, particularly through the analysis of spatial determinants of charging demand and the development of the visualization dashboard.
- Contributed to the broader understanding of electric vehicle charging behavior in Central European urban contexts, where limited research has been conducted compared to Western European and North American settings.

However, it is important to acknowledge certain limitations in the current approach:

- The model's predictive accuracy, while competitive with alternative approaches, still exhibits significant error margins that could impact infrastructure planning decisions. This reflects the inherent complexity of charging behavior and the influence of factors not captured in the available data.
- The reliance on historical charging data introduces potential biases related to the current distribution of charging infrastructure and EV adoption patterns, which may not reflect future conditions as the market matures.
- The spatial resolution of certain data sources, particularly population and mobility data, limits the model's ability to capture micro-level variations in charging demand that may be significant for precise infrastructure placement.

Despite these limitations, the research represents a significant advancement in data-driven approaches to charging infrastructure planning in the Prague context, providing valuable insights for stakeholders navigating the transition to electric mobility.

6.3 Future work

Building on the foundation established by this research, several promising directions for future work emerge:

6.3.1 Model enhancements

The current neural network architecture could be extended in several ways to improve predictive performance and expand its capabilities:

- **Temporal dynamics:** Incorporating recurrent neural network components to capture temporal dependencies and seasonal patterns more effectively could enhance the model's ability to predict charging demand variations over longer time horizons.
- **Spatial relationships:** Integrating graph neural network elements to model the spatial relationships between charging locations could improve predictions by accounting for network effects and competition between nearby charging stations.
- **Transfer learning:** Exploring transfer learning approaches that leverage models trained on data-rich regions to improve predictions in areas with limited historical charging data could expand the model's applicability to emerging markets.
- **Uncertainty quantification:** Extending the model to provide probabilistic forecasts with confidence intervals would enhance its utility for risk-aware infrastructure planning decisions.

6.3.2 Additional data sources

The integration of additional data sources could address some of the limitations identified in the current approach:

- **Vehicle registration data:** Incorporating detailed EV registration statistics at a fine spatial resolution would provide insights into the distribution of potential users and their vehicle characteristics, which influence charging requirements.
- **Grid capacity data:** Integrating information on electrical grid capacity and constraints would enable the model to account for infrastructure limitations in its predictions, supporting more holistic planning approaches.
- **Real-time traffic data:** Leveraging real-time or historical traffic flow data could improve the model's ability to capture the relationship between mobility patterns and charging demand.
- **Weather data:** Including weather variables such as temperature, precipitation, and wind speed could enhance predictions by accounting for their influence on both vehicle energy consumption and travel behavior.
- **Socioeconomic indicators:** Incorporating more detailed socioeconomic data could improve the model's ability to capture variations in EV adoption and charging behavior across different demographic groups.

6.3.3 Application extensions

Beyond improvements to the core prediction model, several application extensions could enhance the practical impact of this research:

- **Optimization framework:** Developing an optimization framework that leverages the prediction model to identify optimal charging infrastructure deployment strategies under various constraints and objectives would provide direct decision support for infrastructure planners.
- **Scenario analysis:** Extending the model to support scenario analysis for different EV adoption trajectories, policy interventions, and technological developments would enhance its utility for long-term planning.
- **Integration with grid planning:** Coupling the charging demand prediction model with electrical grid simulation tools would enable integrated planning that accounts for both mobility needs and grid constraints.

- **Equity analysis:** Developing methods to evaluate the equity implications of different infrastructure deployment strategies would support more inclusive planning approaches that ensure access across diverse communities.

6.3.4 Validation and deployment

Finally, several activities could enhance the validation and practical deployment of the research:

- **Longitudinal validation:** Conducting longitudinal validation studies that compare model predictions with actual charging demand as new infrastructure is deployed would provide valuable insights into the model's real-world performance and opportunities for improvement.
- **Stakeholder engagement:** Engaging with infrastructure planners, grid operators, and policymakers to refine the model and visualization tools based on their practical needs and feedback would enhance the research's impact.
- **Cross-city comparison:** Applying the methodological framework to other cities and comparing results would provide insights into the generalizability of the approach and the transferability of findings across different urban contexts.

In conclusion, while this research has made significant contributions to the challenge of predicting electric vehicle charging demand in urban environments, it represents just one step in an ongoing journey toward enabling the sustainable transition to electric mobility. The future work outlined above offers promising pathways to build on these foundations and address the evolving needs of infrastructure planners, grid operators, and policymakers navigating this critical transformation.

List of Figures

List of Tables

Bibliography

Here are the references in citation order.

[1] Christopher Hecht et al. 'Global Electric Vehicle Charging Station Site Evaluation and Placement Based on Large-Scale Empirical Data from Germany'. In: *eTransportation* 22 (Dec. 1, 2024), p. 100358. DOI: 10.1016/j.etran.2024.100358. (Visited on 02/28/2025) (cited on pages 9, 18).

[2] Guanpeng Dong et al. 'Electric Vehicle Charging Point Placement Optimisation by Exploiting Spatial Statistics and Maximal Coverage Location Models'. In: *Transportation Research Part D: Transport and Environment* 67 (Feb. 2019), pp. 77–88. DOI: 10.1016/j.trd.2018.11.005. (Visited on 01/28/2024) (cited on pages 9, 18).

[3] Anu Maria. 'Introduction to Modeling and Simulation'. In: *Proceedings of the 29th Conference on Winter Simulation*. 1997, pp. 7–13 (cited on page 10).

[4] *Traffic Analysis Tools: Types of Traffic Analysis Tools - FHWA Operations*. URL: https://ops.fhwa.dot.gov/trafficanalysistools/type_tools.htm (visited on 04/13/2025) (cited on page 10).

[5] *Home | Austroads*. URL: https://austroads.gov.au/ (visited on 04/13/2025) (cited on page 10).

[6] Mirosław Lachowicz. 'Microscopic, Mesoscopic and Macroscopic Descriptions of Complex Systems'. In: *Probabilistic Engineering Mechanics*. Special Issue: Stochastic Methods in Mechanics — Status and Challenges 26.1 (Jan. 1, 2011), pp. 54–60. DOI: 10.1016/j.probengmech.2010.06.007. (Visited on 04/13/2025) (cited on page 10).

[7] Matsim Community. *MATSim*. MATSim. Mar. 31, 2025. URL: https://matsim.org/ (visited on 04/13/2025) (cited on page 10).

[8] Andreas Horni, Kai Nagel, and Kay Axhausen, eds. *Multi-Agent Transport Simulation Matsim*. London: Ubiquity Press, Aug. 2016, p. 618 (cited on page 11).

[9] Jan Drchal, Michal Čertický, and Michal Jakob. 'Data-Driven Activity Scheduler for Agent-Based Mobility Models'. In: *Transportation Research Part C: Emerging Technologies* 98 (Jan. 1, 2019), pp. 370–390. DOI: 10.1016/j.trc.2018.12.002. (Visited on 04/15/2025) (cited on page 11).

[10] T. Novosel et al. 'Agent Based Modelling and Energy Planning – Utilization of MATSim for Transport Energy Demand Modelling'. In: *Energy*. Sustainable Development of Energy, Water and Environment Systems 92 (Dec. 1, 2015), pp. 466–475. DOI: 10.1016/j.energy.2015.05.091. (Visited on 04/13/2025) (cited on page 11).

[11] John Brady and Margaret O'Mahony. 'Modelling Charging Profiles of Electric Vehicles Based on Real-World Electric Vehicle Charging Data'. In: *Sustainable Cities and Society* 26 (Oct. 2016), pp. 203–216. DOI: 10.1016/j.scs.2016.06.014. (Visited on 01/13/2024) (cited on page 11).

[12] Azhar Ul-Haq, Carlo Cecati, and Ehab El-Saadany. 'Probabilistic Modeling of Electric Vehicle Charging Pattern in a Residential Distribution Network'. In: *Electric Power Systems Research* 157 (Apr. 2018), pp. 126–133. DOI: 10.1016/j.epsr.2017.12.005. (Visited on 01/15/2024) (cited on page 12).

[13] Siobhan Powell, Gustavo Vianna Cezar, and Ram Rajagopal. 'Scalable Probabilistic Estimates of Electric Vehicle Charging given Observed Driver Behavior'. In: *Applied Energy* 309 (Mar. 2022), p. 118382. DOI: 10.1016/j.apenergy.2021.118382. (Visited on 12/29/2023) (cited on page 12).

[14] Jin Zhang et al. 'Charging Demand Prediction in Beijing Based on Real-World Electric Vehicle Data'. In: *Journal of Energy Storage* 57 (Jan. 2023), p. 106294. DOI: 10.1016/j.est.2022.106294. (Visited on 01/16/2024) (cited on page 12).

[15] Lei Zhang et al. 'An Urban Charging Load Forecasting Model Based on Trip Chain Model for Private Passenger Electric Vehicles: A Case Study in Beijing'. In: *Energy* 299 (July 15, 2024), p. 130844. DOI: 10.1016/j.energy.2024.130844. (Visited on 03/28/2025) (cited on page 12).

[16] Jan Pekárek. 'A Model of Charging Service Demand for the Czech Republic'. In: *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* 65.5 (Oct. 2017), pp. 1741–1750. DOI: 10.11118/actaun201765051741. (Visited on 01/19/2024) (cited on page 12).

[17] Akram Elomiya et al. 'An Advanced Spatial Decision Model for Strategic Placement of Off-Site Hydrogen Refueling Stations in Urban Areas'. In: *eTransportation* 22 (Dec. 1, 2024), p. 100375. DOI: 10.1016/j.etran.2024.100375. (Visited on 02/09/2025) (cited on page 12).

[18] Evženie Uglickich and Ivan Nagy. 'Poisson-Based Framework for Predicting Count Data: Application to Traffic Counts in Prague Areas'. In: *Journal of Computational Science* 85 (Feb. 1, 2025), p. 102534. DOI: 10.1016/j.jocs.2025.102534. (Visited on 02/09/2025) (cited on page 12).

[19] Shuoshuo Shang et al. 'Estimating Building-Scale Population Using Multi-Source Spatial Data'. In: *Cities* 111 (Apr. 1, 2021), p. 103002. DOI: 10.1016/j.cities.2020.103002. (Visited on 04/09/2025) (cited on page 12).

[20] *Ontology Documentation Generated by WIDOCO*. URL: https://auroral.iot.linkeddata.es/def/electrical-vehicle/#https://auroral.iot.linkeddata.es/def/adapters#EnergyProperty (visited on 05/01/2025) (cited on page 15).

[21] Adam Paszke et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019. URL: https://arxiv.org/abs/1912.01703 (cited on page 26).

[22] F. Pedregosa et al. 'Scikit-Learn: Machine Learning in Python'. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cited on page 28).

[23] Tianqi Chen and Carlos Guestrin. 'XGBoost: A Scalable Tree Boosting System'. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Kdd '16. ACM, Aug. 2016, pp. 785–794. DOI: 10.1145/2939672.2939785 (cited on page 28).

[24] *Streamlit • A Faster Way to Build and Share Data Apps*. Jan. 14, 2021. URL: https://streamlit.io/ (visited on 04/17/2025) (cited on page 29).