

---

# Towards Using Machine Learning to Generatively Simulate EV Charging in Urban Areas

---

**Marek Miltner\***  
FEE CTU  
Prague, Czechia;  
CEE, Stanford University  
Stanford, USA  
marek.miltner@stanford.edu

**Jakub Zíka**  
FEE CTU  
Prague, Czechia  
zikajak3@fel.cvut.cz

**Daniel Vašata**  
FIT CTU  
Prague, Czechia  
daniel.vasata@fit.cvut.cz

**Artem Bryksa**  
FEE CTU  
Prague, Czechia  
bryksart@fel.cvut.cz

**Magda Friedjungová**  
FIT CTU  
Prague, Czechia  
magda.friedjungova@fit.cvut.cz

**Ondřej Štogl**  
FEE CTU  
Prague, Czechia  
stoglond@fel.cvut.cz

**Ram Rajagopal**  
CEE, Stanford University  
Stanford, USA  
ramr@stanford.edu

**Oldřich Starý**  
FEE CTU  
Prague, Czechia  
staryo@fel.cvut.cz

## Abstract

This study addresses the challenge of predicting electric vehicle (EV) charging profiles in urban locations with limited data. Utilizing a neural network architecture, we aim to uncover latent charging profiles influenced by spatio-temporal factors. Our model focuses on peak power demand and daily load shapes, providing insights into charging behavior. Our results indicate significant impacts from the type of Basic Administrative Units on predicted load curves, which contributes to the understanding and optimization of EV charging infrastructure in urban settings and allows Distribution System Operators (DSO) to more efficiently plan EV charging infrastructure expansion.

## 1 Introduction and motivation

One of the major pushes to fight climate change is the decarbonization of energy and mobility, which are closely related and together significantly contribute to global carbon emissions[8]. Within this intersection, the electrification of mobility via large-scale deployment of electric vehicles (EVs) is one of the potential tools to decarbonize mobility in the coming decades[1, 15].

This push, however, requires significant investment in power infrastructure, mainly on the distribution level operated by distribution system operators (DSOs)[5, 4]. This is due to the need to expand the

---

\*Alternative email: marek.miltner@cvut.cz

availability of not only private but also public charging infrastructure, which can help better distribute loads across space and time[6, 13].

Since power engineering infrastructure expansion is an effort requiring significant time and financial resources, a critical challenge in this area is how to optimize grid expansion for efficiency to cover anticipated EV charging demand in various areas while not overloading the current network and not overspending on areas where demand is not as high[11, 14]. This is especially difficult for DSOs since there has been a general lack of studies demonstrating analysis of real-world EV charging data in different geographies, mainly due to the data being vendor-locked and treated as confidential[2, 10, 11].

In this study, we aim to fill this critical gap by collaborating with PREdistribuce, the DSO in Prague, and the largest operator of public EV chargers. We propose a method to generatively create anticipated EV charging load curves based on location characteristics even in places where no chargers are present, in order to allow for DSOs and the wider community to better understand and simulate public EV charging behavior.

## 2 Generative modeling of public charging demand

In order to solve the challenge of predicting charging profiles for selected locations with limited information, we first analyzed the information about the landscape of public chargers in Prague. Our study is based on data sourced from public EV chargers operated by PREdistribuce, the DSO in the Prague area.

We have gotten access to full logs of all charging sessions, including precise start and end times, power consumption, type and power output of each charger, and their location. State of charge or identification of individual vehicles was not available to us. We have paired this information with other geographic and demographic data to complement our understanding of the individual charger locations, based around the basic administrative units the chargers are located in. The public EV charging landscape is further explored in Appendix A.

### 2.1 Model inputs

For the generative predictions of load curves, we have assessed several factors that might come into play to affect charging loads[11]. In connection to the basic administrative units, we were able to include several attributes that, in our opinion, are able to increase the performance of our model, including the area character and the amounts of local and long-haul commuters. Firstly, to reflect the character of the surrounding area, we include the basic administrative unit type, which includes 12 distinct categories ranging from residential to industrial. Secondly, to reflect local commuting, we include population density per the administrative unit and the local number of addresses, including industrial and commercial buildings, per the corresponding basic administrative unit. Thirdly, we have further enhanced this data to include information on vehicles commuting to the wider area of town to reflect long-haul commuting factors. In connection to the charging data, we have performed additional analysis of charging load data, uncovering insights that may be found in Appendix B

### 2.2 Model architecture

There are two distinct, yet related components that DSOs are interested in, the first being peak power demand and the second being the actual load shape during the day [9, 7]. Moreover, our hypothesis is that, in general, there exist  $K$  underlying archetypal charging profiles, which are mixed by factors depending on the spatiotemporal properties of a charging station. Once enough data flows into the model given by a neural network, our aim is to interpret and examine those profiles and how the mixture factors depend on particular spatiotemporal settings.

In order to reflect those criteria and with an emphasis on gaining a deeper understanding of charging behavior, by introducing latent charging profiles, we have chosen a neural network with the architecture shown in Figure 1. As seen in the figure, there are two distinct modules: the  $g$  module is tasked with predicting peak load, and the  $f$  module is used to predict the load shape. As shown in the figure on the right, the  $g$  module is trained to predict the total power scalar  $l$ , and the  $f$  module predicts the probabilities of the  $K$  underlying archetypal latent charging curves  $\mathbf{R}$ , which are shown

on the left. Each of them has 24 values, representing the load curve distribution over 24 hours. The resulting predicted load curves for each charger and corresponding set of inputs described above is a combination of these latent profiles, determined by a ratio output by the  $f$  module and then multiplied by a scalar  $l$  to determine the peak load, which is the output of the  $g$  module.

As a base scenario, we have run our experiment with the parameters as described in Appendix C, notably making the assumption that there are 4 underlying latent profiles that might be attributed to archetypal behavior.

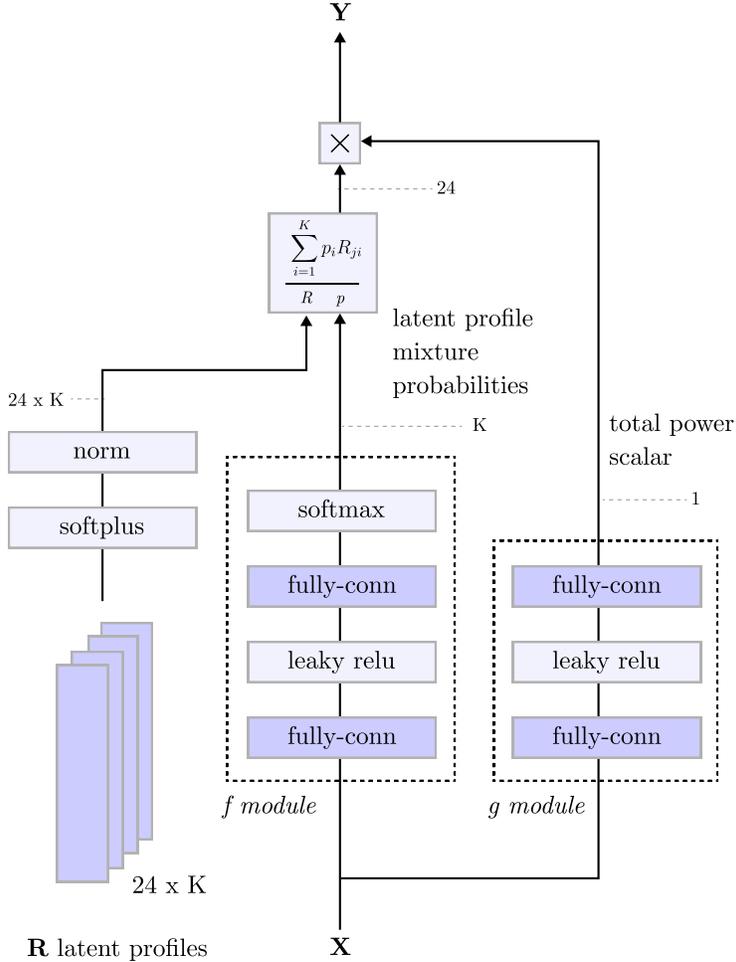


Figure 1: Neural network architecture used in this study.

### 3 Results

Using the described experimental setup, we can examine the generated load profiles in Figure 2. It was our hypothesis that the model would be able to decompose the input parameter influences and assign them to latent load curves representing archetypal behavior. After some experimenting, we have chosen  $K = 4$  latent profiles in this initial study. Since the generated load curves differ from each other starkly, it appears the model works as intended.

A basic interpretation of the generated latent load profiles is that latent curve 1 shows a gradual charging curve with a peak during the day, typical for public charging. The latent curve 2 shows a single morning peak demand, and the latent curve 3 shows a single evening demand, which is generally more typical for private charging. Lastly, latent curve 4 shows a residual demand with multiple peaks. What are the causes of these archetypal behaviors is a topic of further research,

however, if we follow our hypothesis more in-depth, we may also relate the emergent load curves in the context of observed area charging specificity described in Appendix B.1.

When comparing to the generated load profiles observed here, we can see similar load profiles coming out of the original charging data matched to types of basic administrative units. If we were to make this connection speculatively, it creates an implication that the latent curve 1 might represent the local, residential charging archetypal behavior, latent curve 2 might represent recreational commute, latent curve 3 suburban and long-range commute, and latent curve 4 might correspond to irregular industrial and commercial commute. Based on this connection, it appears that out of the assembled input data, the strongest impact on the predicted load curves is held by the type of the local basic administrative unit.

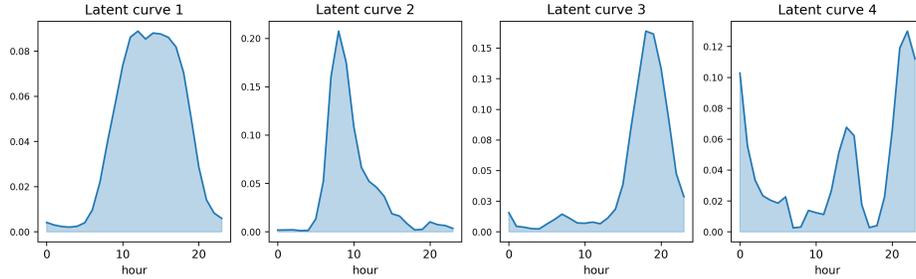


Figure 2: These 4 plots showcase trained latent load curve shapes within the neural network architecture. Note that the latent profiles are probability distributions, and thus the sum of the area they define is equal to 1.

## 4 Conclusion and further work

In this study, we have introduced a novel approach to creating a machine learning based model which allows for generative creation of public EV charging load profiles in various urban geographies given a wide range of input data. Our method is designed to create internal latent charging profiles, which have the potential to better explain the sources of charging demand if interpreted in connection with archetypal behaviors. In our results, we have offered such an interpretation based on available data analysis.

Our proposed model has the potential to assist DSOs and the wider research community in better modelling public charging demand, model different potential future scenarios, and in the case of DSOs, create more efficient strategies to expand EV charging grid infrastructure and support the EV number increase and its wider accessibility while keeping unexpected grid impacts to a minimum. Our model is still undergoing further refinement and there are some limitations to acknowledge. Firstly, we aim to include more data in our future iteration, as our current data has a cutoff in 2022 as described in Appendix A.2. Connected to this, further investigation is needed to understand what effect the COVID pandemic lockdowns had on our input data, since as shown in Appendix B.5, the decrease in total load was significant. We are also continuing our research in the interpretability of the charging load profile sources, facilitated in our model via the latent archetypal load shapes, with the interpretation being speculative. An alternative approach might be to strictly separate types of inputs to be used for each latent load shape. Lastly, we invite the broader community to apply similar approaches to different geographies to cross-reference our findings.

## Acknowledgments and Disclosure of Funding

The authors would like to thank and acknowledge PREdistribuce, the Prague Distribution System Operator(DSO), for the charging data used in this study, and the Czech Statistical Office for data on demography and geography classification in Czechia. This work was supported by grants number SGS24/093/OHK5/2T/13 and SGS23/117/OHK5/2T/13 provided by CTU Prague, and number TS01020030 provided by the Technology Agency of the Czech Republic. Authors declare no conflict of interest.

## References

- [1] P. Ahmadi. Environmental impacts and behavioral drivers of deep decarbonization for transportation through electric vehicles. *Journal of cleaner production*, 225:1209–1219, 2019.
- [2] B. P. Bhattarai, S. Paudyal, Y. Luo, M. Mohanpurkar, K. Cheung, R. Tonkoski, R. Hovsopian, K. S. Myers, R. Zhang, P. Zhao, et al. Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions. *IET Smart Grid*, 2(2):141–154, 2019.
- [3] Czech Statistical Office. Základní sídelní jednotky - polygony, 2024.
- [4] P. M. de Quevedo, G. Muñoz-Delgado, and J. Contreras. Joint expansion planning of distribution networks, ev charging stations and wind power generation under uncertainty. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2017.
- [5] A. Dias, P. M. Carvalho, P. Almeida, and S. Rapoport. Multi-objective distribution planning approach for optimal network investment with ev charging control. In *2015 IEEE Eindhoven PowerTech*, pages 1–5. IEEE, 2015.
- [6] S. Gnanavendan, S. K. Selvaraj, S. J. Dev, K. K. Mahato, R. S. Swathish, G. Sundaramali, O. Accouche, and M. Azab. Challenges, solutions and future trends in ev-technology: A review. *IEEE Access*, 2024.
- [7] B. Hashemi, M. Shahabi, and P. Teimourzadeh-Baboli. Stochastic-based optimal charging strategy for plug-in electric vehicles aggregator under incentive and regulatory policies of dso. *IEEE Transactions on Vehicular Technology*, 68(4):3234–3245, 2019.
- [8] S. Kais and H. Sami. An econometric study of the impact of economic growth and energy use on carbon emissions: panel data evidence from fifty eight countries. *Renewable and Sustainable Energy Reviews*, 59:1101–1110, 2016.
- [9] V.-L. Nguyen, T. Tran-Quoc, S. Bacha, and B. Nguyen. Charging strategies to minimize the peak load for an electric vehicle fleet. In *IECON 2014-40th annual conference of the IEEE industrial electronics society*, pages 3522–3528. IEEE, 2014.
- [10] V. Potdar, A. Chandan, S. Batool, and N. Patel. Big energy data management for smart grids—issues, challenges and recent developments. *Smart Cities: Development and Governance Frameworks*, pages 177–205, 2018.
- [11] S. Powell, G. V. Cezar, L. Min, I. M. Azevedo, and R. Rajagopal. Charging infrastructure access and operation to reduce the grid impacts of deep electric vehicle adoption. *Nature Energy*, 7(10):932–945, 2022.
- [12] PREdistribuce. Veřejné dobíjení | PRE, 2024.
- [13] F. Un-Noor, S. Padmanaban, L. Mihet-Popa, M. N. Mollah, and E. Hossain. A comprehensive study of key electric vehicle (ev) components, technologies, challenges, impacts, and future direction of development. *Energies*, 10(8):1217, 2017.
- [14] Q. Zhang, H. Li, L. Zhu, P. E. Campana, H. Lu, F. Wallin, and Q. Sun. Factors influencing the economics of public charging infrastructures for EV—A review. *Renewable and Sustainable Energy Reviews*, 94:500–509, 2018. ISBN: 1364-0321 Publisher: Elsevier.
- [15] R. Zhang, J. Zhang, Y. Long, W. Wu, J. Liu, and Y. Jiang. Long-term implications of electric vehicle penetration in urban decarbonization scenarios: An integrated land use–transport–energy model. *Sustainable Cities and Society*, 68:102800, 2021.

## A Prague EV public charging landscape

### A.1 Prague area classification per basic administrative unit

In order to analyse the potential factors affecting charging demand, we have to understand the environment in which individual charging points are located. For this aim, we have utilized the enhanced "multi" layer of Basic Administrative Units (Základní sídelní jednotka, ZSJ), as provided by the Czech Statistical Office[3]. The ZSJ is the smallest administrative unit available, and there are 948 individual units in Prague. Using this data, we have been able to classify individual ZSJ areas in Prague into 12 categories, as seen in Figure 3.

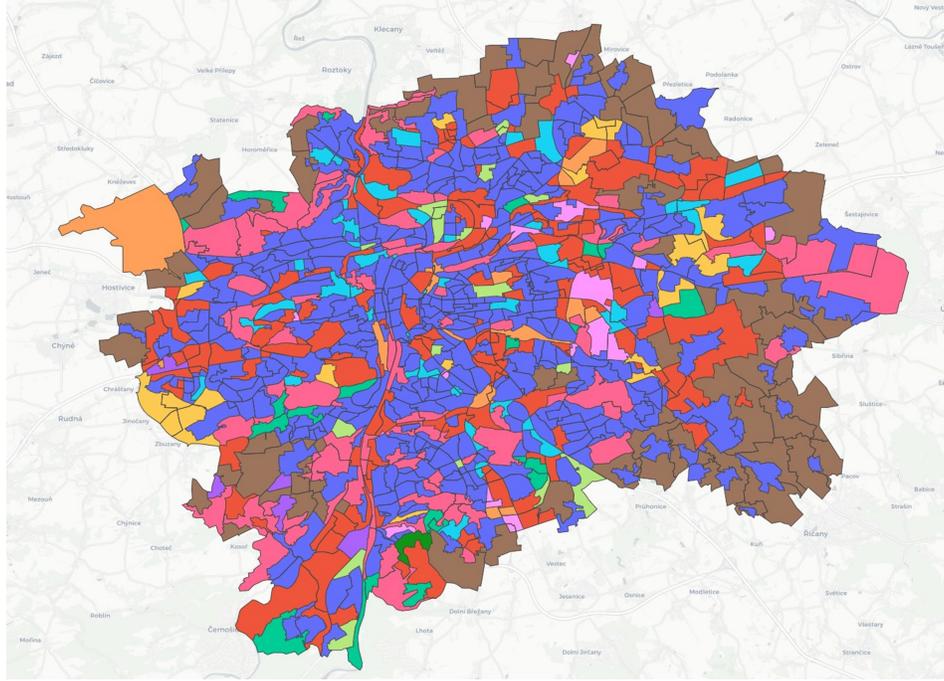


Figure 3: ZSJ categories found in Prague proper

Since the names for ZSJ categories are standardised in Czech, table 1 shows English translations and a colour-coded legend for the following figures.

Basic Administrative Unit (ZSJ) Categories		
Original Czech name	English translation	Colour
Obytná plocha v kompaktní zástavbě	Compact residential area	Purple
Městská a příměstská smíšená plocha	Urban and suburban mixed area	Red
Obytně rekreační plocha	Residential and recreational area	Teal
Odloučená obytná plocha	Separated residential area	Violet
Dopravní areál	Transportation infrastructure area	Orange
Areál občanské vybavenosti	Civic amenities area	Blue
Rekreační plocha	Recreational area	Coral
Ostatní účelová plocha	Urban and suburban mixed area	Lime
Průmyslový areál	Industrial area	Pink
Rezervní plocha	Reserve area	Yellow
Zemědělská plocha	Agricultural area	Brown
Lesní plocha	Forest area	Green

Table 1: Legend for ZSJ categories in following figures.

## A.2 Charging demand localization in Prague

The data on individual charging sessions was kindly provided by Prague's largest provider of public charging points, PREDistribuce[12]. Figure 4 shows a visualisation of charger locations within the ZSJ structure of Prague.

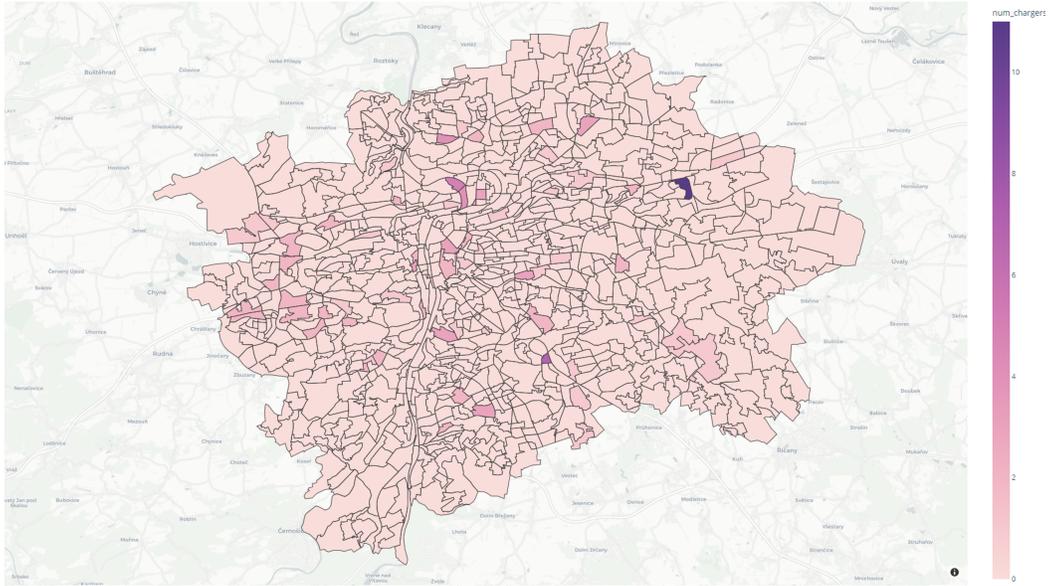


Figure 4: Heatmap of public charging point locations per ZSJ in Prague based on the available data.

We can see the broad distribution of public chargers across the city of Prague, with the largest concentration in population centres, most notably the one at Černý most in the top right of the map, with 12 chargers present in one ZSJ. In Figure 5, we can see that most chargers in the dataset are placed in residential areas (70 chargers), followed by civic amenities areas (34 chargers) and urban and suburban mixed areas (16 chargers). The other categories of ZSJ areas have 3 or fewer chargers present.

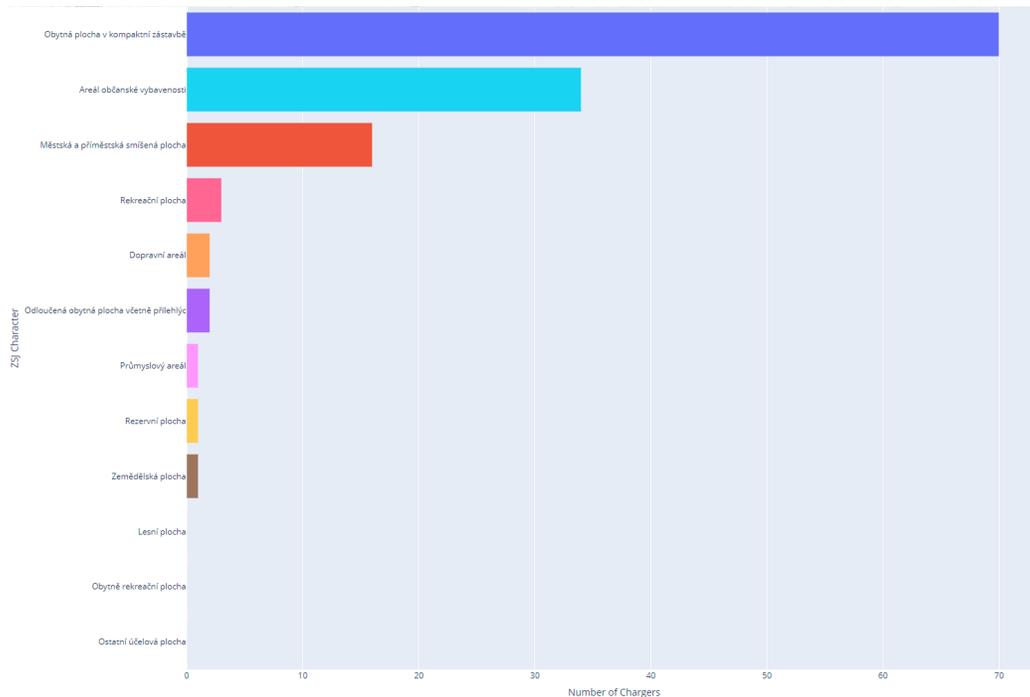


Figure 5: Number of chargers per ZSJ category across Prague

In Figure 6, the monthly relative total shares of each ZSJ area type are visualized for two key elements in our dataset: charging instances and installed charging points. Note that for the charging instances dataset, no datapoints are currently available to us between December 2020 and December 2021, as highlighted in the figure by a red-bordered gray fill-in with interpolated data. We can see that the number of charging sessions in Compact residential areas rose between the years 2019 and 2022. Simultaneously, a modest increase in the relative share of charging instances can be observed in Agricultural and Industrial areas. Conversely, a decline in charging instances share is evident in Civic amenities, Urban and suburban mixed, and Recreational areas. This decrease could be attributed to various factors, such as changes in usage patterns or the availability of charging stations.

However, combined with the analysis of the share of installed charging stations on the lower plot, we can see that the change in charging sessions relative shares for individual ZSJ area types described above are not directly explainable by developments in the number of available chargers in the respective ZSJ types. Indeed, we can see that while the two most dominant ZSJ types (residential areas and civic amenities areas, respectively) are overrepresented in the number of charging instances, the suburban, recreational, transportation, and reserve areas are greatly underrepresented in the charging instances relative to the number of available chargers. This observed behavior may be caused by several characteristics of chargers in the dataset, such as accessibility, geographic location, charging point alternatives, and others. Still, this discrepancy clearly implies a different behavioral pattern for chargers in different ZSJ area types, which is examined next.

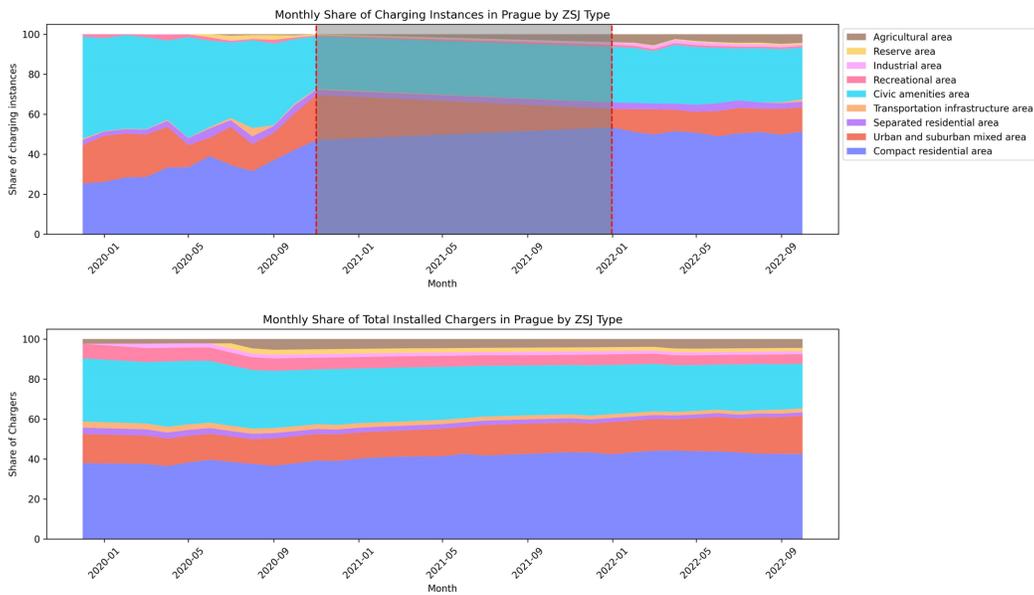


Figure 6: Temporal relative share development analysis for charging instances and installed chargers in Prague, classified per ZSJ type. Note that the red-bordered gray fill area in the upper chart of charging instances corresponds to the interpolated region of unavailable data, and the lower chart of the number of installed chargers includes chargers put into operation before the span of the chart timeline, December 2019.

## B Public charging load curve analysis

### B.1 Load-curve analysis based on basic administrative unit type

We have investigated the specificity of ZSJ area types by plotting demand curves of EV charging split per individual type. Figure 7 provides a visualization of the normalized average number of charging instances per hour of day, which approximates load curve behavior patterns across different ZSJ area types. The analysis clearly shows that demand patterns between different ZSJ types are heterogeneous. We see four distinct groups emerge:

- **Group 1: Sustained single peak areas** demonstrate a sustained demand in charging instances with a gradual single peak throughout the day, suggesting a consistent demand for charging facilities in these areas. The following ZSJ types fall within this category:
  - Compact residential area
  - Urban and suburban mixed area
  - Civic amenities area
- **Group 2: Morning single peak areas** exhibit a clear peak in the morning hours, around 8:00. The following ZSJ types fall within this category:
  - Transportation infrastructure area
  - Recreational area
- **Group 3: Evening single peak areas** demonstrate a single peak in the evening, around 17:00. The following ZSJ types fall within this category:
  - Separated residential area
  - Agricultural area
- **Group 4: Double peak areas** demonstrate both the morning peak and the evening peak, though potentially shifted by an hour. The following ZSJ types fall within this category:
  - Reserve area
  - Industrial area

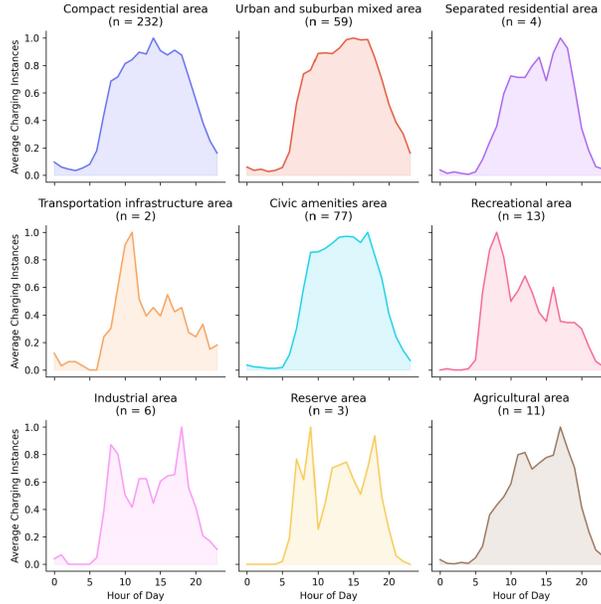


Figure 7: Normalized demand curves for ZSJ types with  $n$  defined as number of charge points present in the dataset for each ZSJ type

In the figure we compare data within time series or categories of inherently different volume, in categories of locations with varying numbers of charging points, we have therefore utilized max normalization feature scaling. It rescales all values of a non-negative data series to a maximal value of one, while retaining the relative distance from zero. Normalized series and their characteristics are then directly comparable. The  $i$ -th element of the series is rescaled as follows:

$$x'_i = \frac{x_i}{\max(x)}$$

## B.2 Average load curves per month and weekday

We have also observed trends in seasonality between weekdays and months as shown in Fig. 8. While no significant load shape differences are found, we can observe marked lower demand on the weekends and in summer months.

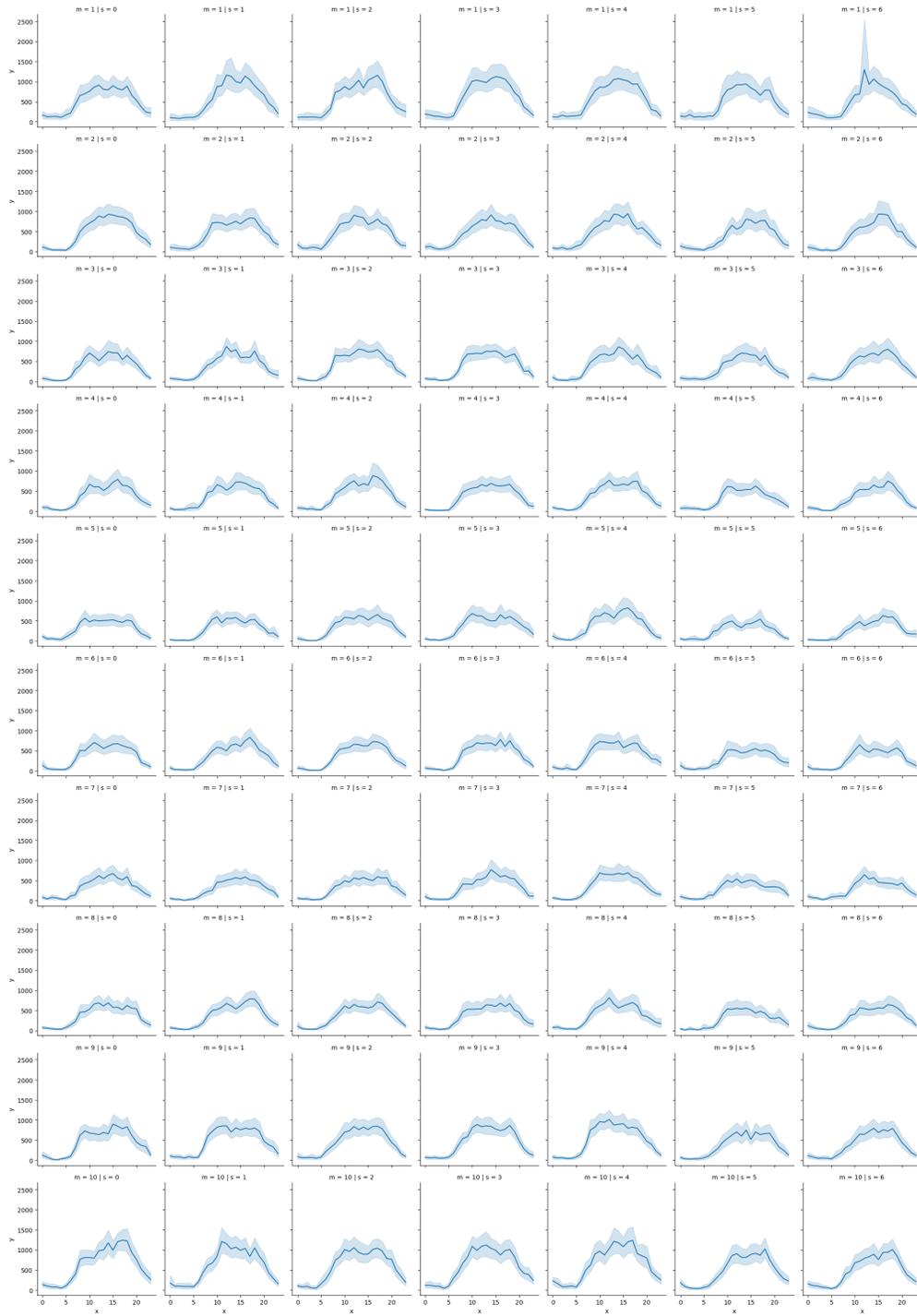


Figure 8: Comparing the average load curves and their confidence intervals for each weekday (columns Monday to Friday) and month (rows January to December)

### B.3 Weekday analysis

Further delving into the differences between individual weekdays, in Fig. 9 we have shown individual load-curve shapes between weekdays and weekends for a subset of chargers in the dataset. While weekends show generally lower demand, high variability is also observed.

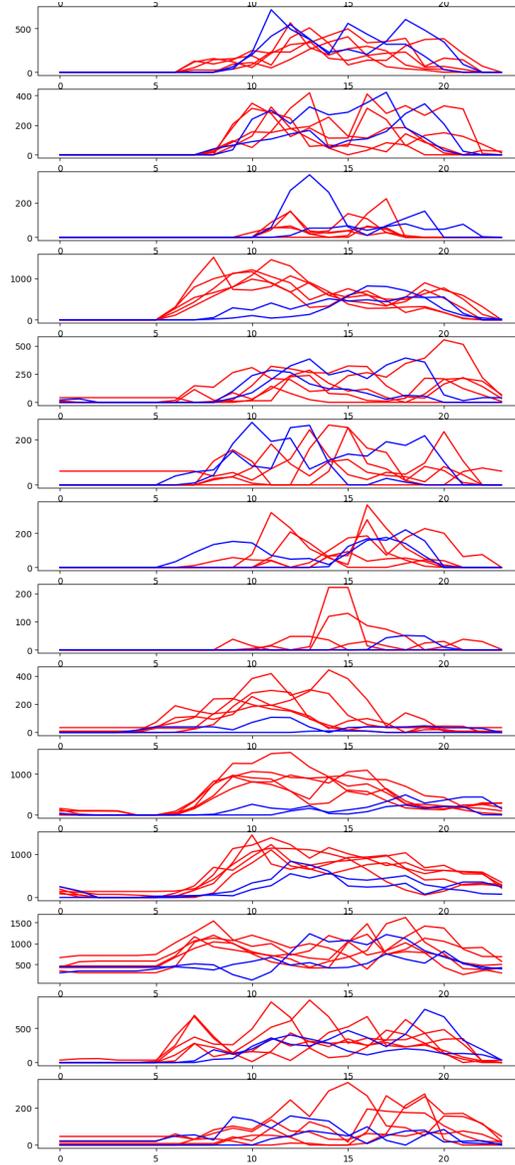


Figure 9: Comparing load curves for weekends (marked in blue) and weekdays (marked with red) for a subsample of individual chargers

#### B.4 Easter holiday analysis

Following our analysis in B.2 and B.3, we have examined what effect public holidays might have on charging, by comparing the Easter holiday week average load curve with other weeks shown in Fig. 10, however no major differences or anomalies were found.

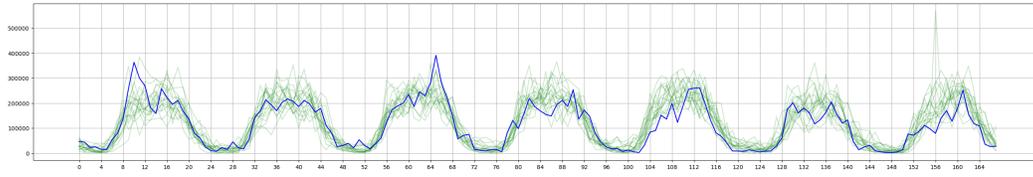


Figure 10: Easter week average charging load curve in blue compared to average loads of standard workweeks in green

#### B.5 COVID lockdown total load impact analysis

As shown in Fig. 11, COVID lockdowns in Prague after May 2020 have had a stark impact on general public charging loads. The impact on our dataset is a topic of further research.

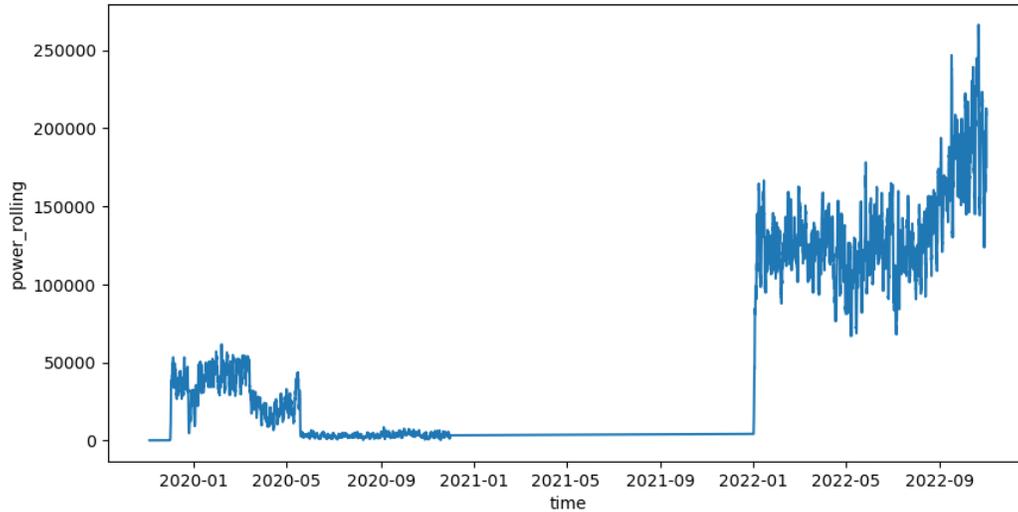


Figure 11: Total load observed during the dataset

## C Hyperparameters used for running the Experiment

The following table lists the hyperparameters used to configure and train the `ChargingProfileModel` in the experiment described in this paper:

Hyperparameter	Value
Charging Profile Granularity	24
Hidden Size	128
Hidden Size (g)	64
Latent Profiles Count	4
Loss Function	<code>nn.MSELoss(reduction="mean")</code>
Optimizer	<code>torch.optim.Adam</code>
Learning Rate	0.0004
Number of Epochs	150

Table 2: Hyperparameters used in the `ChargingProfileModel` experiment.

In our experiment, we used the `ChargingProfileModel` with the following hyperparameters:

- **Charging Profile Granularity:** Set to 24, this parameter defines the granularity of the charging profile, corresponding to the daily granularity of a 24 hour cycle in our case.
- **Hidden Size:** The size of the hidden layers in the *f* module, set to 128.
- **Hidden Size (g):** The size of the hidden layers in the *g* module, set to 64.
- **Latent Profiles Count:** The number of latent profiles, set to 4 in our experiment run.
- **Loss Function:** Mean Squared Error (MSE) loss function, specified as `nn.MSELoss(reduction="mean")`.
- **Optimizer:** Adam optimizer, used for updating the model parameters.
- **Learning Rate:** Set to 0.0004, this controls the step size during optimization.
- **Number of Epochs:** The model is trained for 150 epochs.

These hyperparameters were chosen to optimize the performance of the `ChargingProfileModel` on the given dataset. The model was trained using the Adam optimizer with a learning rate of 0.0004 and the MSE loss function. The training process was carried out for 150 epochs to ensure adequate learning.