Simulating Plug-in Electric Vehicle Charging for AutoML-Based Prediction of Regional Energy Demand

Matthias Schneider and Sören Frey

Daimler TSS GmbH
Wilhelm-Runge-Str. 11, 89081 Ulm, Germany
Email: {matthias.b.schneider, soeren.frey}@daimler.com

***Abstract*—We present a system for simulating home and public charging operations of Plug-in Electric Vehicles (PEVs). We model PEV traffic streams that result in corresponding charging operations. The simulation allows to configure many influential factors, such as the number of PEVs, consumption, charging stations, their locations, charging power, working hour distributions, holiday seasons, and the ratio of regular to irregular rides. In this paper, we demonstrate the applicability of our simulation in the context of predicting the short-term, regional energy demand of PEV charging. The prediction can be used to support energy suppliers and charging infrastructure operation, for instance. We use automated machine learning (AutoML) to train a forecasting model based on the simulation output. This combined workflow, integrating discrete-event simulation and machine learning, allows us to build a prediction pipeline where simulation data can be swapped with real data once available.**

Index Terms*—*Plug-in Electric Vehicle Charging; Simulation; Energy Demand Prediction; Machine Learning; AutoML.

I. Introduction

The transportation sector is facing a massive transformation in the upcoming years as the penetration of Plug-in Electric Vehicles (PEVs) is rapidly increasing [1]. This paper describes our work-in-progress Discrete-Event Simulation (DES) that models PEV traffic and emits corresponding charging operations. The simulation can be used in a variety of ways to advance electric mobility, e.g., for analyzing charging patterns and fostering the understanding of PEV owners' charging behavior and their corresponding needs. The simulation incorporates aspects related to driver behavior, e.g., working hour distributions, holiday seasons, and ratio of regular to irregular rides. It also involves equipment and adoption aspects, such as the number of PEVs, charging stations, charging power, battery capacity, and consumption. We demonstrate the usefulness of our simulation with the following application example.

A vital component for the propagation of electric mobility is a reliable and broadly available charging infrastructure. To enable this, it is important to accurately predict the realistic PEV charging operations while taking into account major influencing factors, such as the driver behavior [2]. Connected cars can fuel the underlying data basis for those predictions by providing data points that cover public and home charging operations. However, until connected PEVs will be the prevalent vehicle class, simulation data can be used as a versatile proxy to build up a resilient prediction pipeline. We use the output of our simulation, i.e., charging data, as the basis for training a prediction model with automated machine learning (AutoML, see Section IV) [3]. The model forecasts the short-term (up to a day) energy demand of PEV charging on a regional level. The paper provides the following main contributions:

1) A simulation of PEV charging behavior
2) An application example using the simulation output to train a model for regional energy demand prediction
3) An evaluation of the simulation and prediction model

The remainder of the paper is structured as follows. Section II examines the related work. The simulation is presented in Section III and employed in the application example in Section IV. Then, the evaluation is described in Section V before the conclusions are drawn in Section VI.

II. Related Work

The simulation of traffic streams was relevant long before electric mobility became more widespread. The traffic simulation SUMO [4] was utilized in traffic management and routing research, for instance. For the analysis of electric vehicle traffic, many different approaches are used, such as Monte Carlo methods [5] or M/M/s queueing theory [6]. In comparison, we use DES and allow for a combination of workday and holiday patterns and incorporate the usually rather slow (private) home charging. Similar to [5] and [6], most other approaches (e.g. [7]) focus only on public charging.

The electric mobility simulators most similar to our system are V2G-Sim [8], ACN-Sim [9], and EVLibSim [10]. V2G-Sim was used to study battery degradation and integration of PEVs in smart grids as power sources, for instance. ACN-Sim and EVLibSim focus on the charging infrastructure perspective and incorporate detailed models for, e.g., pricing or unbalanced three-phase infrastructure. In contrast, we model traffic streams that result in charging operations. There exist several approaches that utilize Machine Learning (ML) for PEV charging load forecasting [11], but our application example is, to the best of our knowledge, the first that uses AutoML.

 ISBN: 978-1-61208-898-3

III. PEV Charging Simulation

A. Overview

We utilize the DES framework SimPy [12] for simulating the charging behavior of PEVs. The active components like PEVs or charging stations are modeled as processes, which interact with each other and their environment using events. For example, when a PEV starts to charge, it has to interact with a public or home charging station. This interaction is represented by a triggered charge event, which initiates the recharging of the PEV's battery. The major simulation entities are illustrated in Figure 1. A public or home *ChargingStation* comprises a number of *ChargingPoint*s. At a *ChargingPoint*, only one vehicle can charge at a time. *PEV*s and *ChargingStation*s are always situated at a specific *Location*. Additionally, *PEV*s have a permanently assigned home *Location*. Furthermore, a PEV is also tied to a regular or irregular *Tour* (see Section III-B), which in turn has exactly two *Location*s (start and destination).

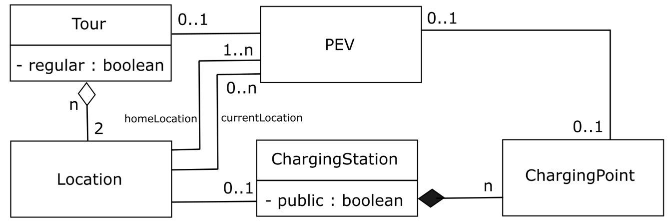

Fig. 1. The most important simulation entities.

B. Assumptions

The charging behavior of the simulated PEVs underlies some basic assumptions.

1) As soon as a PEV reaches its destination, it charges if it is not capable to perform the next tour with the current State of Charge (SoC). In this case it may have to find a suitable charging station nearby.
2) If a PEV arrives at its home location and the home location has an existing charging station, the PEV always starts to charge regardless of its SoC status.
3) If a vehicle needs to charge during a tour, a suitable charging station is searched along the tour.

The consumption is calculated on the basis of the tours' distance and the average consumption of the PEV.

Public and home charging stations are able to charge multiple PEVs at the same time. This behavior is also reproduced in the simulation. The charging power of the charging points can be configured (see Section III-C).

There are two categories of tours a PEV drives: regular and irregular tours. A regular tour is, for example, the commute to the worksite and back. An irregular tour occurs sometime between the regular tours and represents, for example, a spontaneous trip to drive the children to school because the school bus broke down. When an irregular tour overlaps with a regular tour, the regular tour will be canceled. The times when the PEVs start their regular tours are determined by a normal distribution. Each kind of regular tour has its own normal distribution whose parameters can be configured individually. For example, the simulation initiates tours on weekday mornings and late afternoons, representing rush hour traffic. On weekends the distribution changes and the tours are initiated later than on weekdays. Furthermore, the average amount of tours PEVs drive during weekend days is lower than on weekdays.

In addition to the changing driving behavior between weekdays and weekends, vacation periods are also included in the simulation. During vacation periods, significantly fewer trips are simulated because, for example, the commute to work is omitted. Furthermore, it is assumed that some people are not at home during vacation time and therefore fewer trips and charging events take place. Both the daily number of trips during vacation periods and the vacation periods themselves (start and end times and duration) can be configured (see Section III-C).

C. Simulation Input

The simulation is configured via a set of files, which are read at the beginning of the simulation. The files contain information regarding charging stations, worksites, and vacation periods, for instance. Moreover, we account for the inherent probabilistic nature of charging behavior by incorporating several distributions in the configuration files. For example, the duration of a vacation period can be configured via a normal distribution, i.e., through defining its corresponding mean (μ) and standard deviation (σ). An excerpt of the simulation input parameters is shown in Table I. In total, there exist 71 degrees of freedom that can be configured with corresponding simulation parameters.

D. Simulation Design

As mentioned in Section III-A the simulation contains events, which define the simulation flow. The events are triggered by the processes of the simulation. The event flow for each PEV is depicted in Figure 2.

The simulation of each PEV starts with a *WaitEvent*. This event is triggered when a PEV has currently no tour to drive and/or is sufficiently charged. If a *WaitEvent* ends, a new tour begins by triggering a *DrivingEvent*, which simulates a tour of a PEV by reducing the PEV's SoC. In our work-in-progress implementation the amount of SoC reduction solely depends on the distance the PEV traveled and its average speed. It is always decided in advance which tour (regular or irregular) will be run and when it will start. Moreover, it has to be determined if the PEV's battery has to be recharged. If a charging operation has to be initiated, either a *HomeChargeEvent* or a *PublicChargeEvent* is triggered, depending on the PEV's location. If charging was not necessary or the charging operation finished, a new *WaitEvent* is triggered. The termination condition of the simulation is met as soon as the simulation time is greater than the configured maximum simulation time.

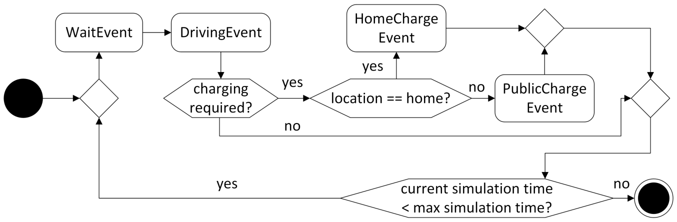

Fig. 2. Event flow for each PEV.

E. Simulation Output

The simulation writes the relevant information of the executed charging operations to a CSV file, such as their start time and duration. The most important charging operation data is shown in Table II.

IV. Application Example

A. Overview

We demonstrate the simulation's applicability with a prediction of regional energy demand on the basis of simulation results. The output of the simulation is utilized to train an AutoML-based model. AutoML creates ML models automatically by using dynamically selected ML techniques. The parameterization of the ML models is also automated. Thus, AutoML only requires a training and test dataset as input, from which an ML model is then generated [3].

In this application example, the presented simulation is used to simulate the charging operations of the city of Stuttgart (Germany) for an entire year. For this purpose, the simulation inputs (e.g., locations of the charging stations or the number of PEVs) are adjusted accordingly. The simulation output is then passed to the AutoML-library auto-sklearn [13], which creates an ML model, predicting the regional energy demands. In order to pass the simulation output to AutoML, the output data must first be transformed. This transformation is also known as Feature Engineering (see Section IV-B) in the context of ML.

B. Feature Engineering

Features are measurable properties of the problem to be solved and are used for training the ML model. We dynamically divide the region Stuttgart into several partitions and create corresponding features by transforming and aggregating the simulation output as follows: (1) All charging operations that occur in a given partition are aggregated. A partition is a rectangle whose height and length are adjustable, i.e., the number of partitions decreases or increases. (2) Charging operations are aggregated by time intervals. (3) All other values describing a charging operation (e.g., charged amount of energy in kW) are aggregated and averaged. In this application example, we use different combinations of one partition size and one time interval at each time to create several AutoML model candidates. The goal is to determine the combination that delivers the AutoML model with the best performance.

TABLE I
Simulation input parameters (Excerpt)

Name	Description
ChargingStation.Location	Location of the charging station
ChargingStation.MaxCPower	Max. charging power of charging station
ChargingStation.NrCP	Nr. of charging points at charging station
Context.NrPEV	Nr. of simulated PEVs
Home.Location	Home location of the PEV's owner
StartOfWork.NDist.Mean	Mean start time of work
StartOfWork.NDist.Sd	Standard deviation start time of work

TABLE II
Simulation output data (Excerpt)

Name	Description
ChargingOperation.Duration	The duration of the charging operation
ChargingOperation.Kw	The charged energy amount in kW
ChargingOperation.Location	Location of corresponding charging point
ChargingOperation.Start	Start time of the charging operation
Vacation.Present	Vacation period present during charging?

Table III shows example input data for a single partition (Id: 6449) and multiple time intervals (with their corresponding start times) and with various derived features. For example, several charging operations might occur in a specific time interval. Hence, we created the feature *MajorityChargingType* that describes, which type of charging operation (public or private) occurred most often. As a consequence, the total charging time (feature *SumChargingTime*) and average charging time (feature *AvgChargingTime*) might also differ if multiple charging operations took place. We used 75% of the data for training and 25% for testing the models.

V. Evaluation

A. Simulation

The simulation has to comply with the underlying assumptions described in Section III-B. In particular, we analyze and compare the results for simulating weekdays with weekend days. Figure 3 contrasts these two cases and shows the corresponding number of private and public charging operations for examplary days. We find that the total number of charging operations on weekends decreases compared to those on weekdays.

On a weekday, around 5am, private charging operations decrease noticeably, while charging operations at public charging stations increase. This is due to the fact that PEVs are being driven to work sites or other locations. In the evening, the private charging operations then rise considerably, which means that the PEVs are coming home again. This is also supported by the fact that public charging initially decreases during this period. These observations are consistent with findings from [14] and [15]. Chowell et al. and Mucelli Rezende Oliveira et al. report that (1) trips to the worksite usually start around 5am, (2) reach their maximum at 8am, and (3) the workday ends on average between 3-4pm. The renewed increase in public charging operations in the evening suggests that tours take place again (to locations of public life and leisure). This behavior is not seen as strongly on

TABLE III
ML INPUT DATA (EXAMPLE)

StartTime	DayOfWeek	Partition Id	ChargedKwh	MajorityChargingType	IsVacation	SumChargingTime	AvgChargingTime
1590364800	monday	6449	39.88	public	true	36723.46	36723.46
1590451200	tuesday	6449	55.18	private	true	50814.11	25407.05
1591660800	tuesday	6449	41.54	public	false	38255.0	12751.66

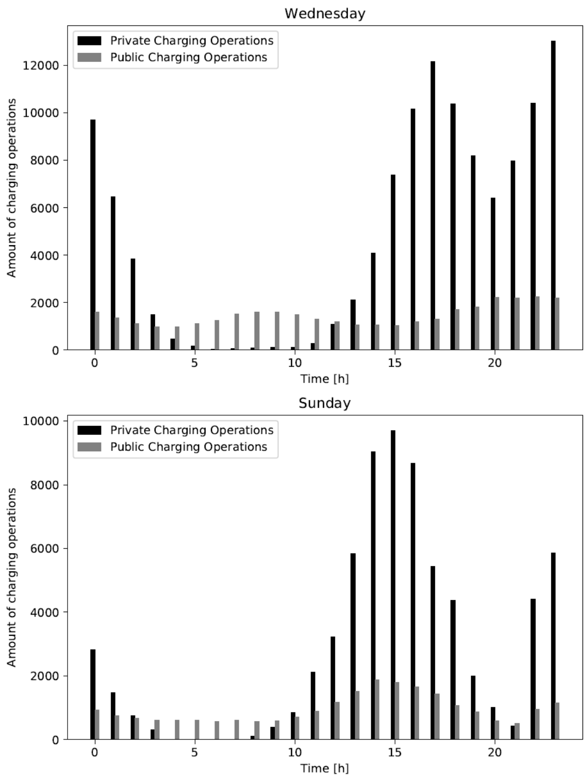

Fig. 3. Amount of charging operations for an exemplary weekday (Wednesday) and weekend day (Sunday).

weekend days. On weekends, it can be observed that there is a strong increase in home charging at around 3pm. This suggests that PEVs are returning from leisure trips (started earlier) at these times. The fact that charging does not increase as much thereafter suggests that many PEVs do not make any more trips on a weekend day. Thus, on average, a PEV drives less tours on weekend days than on weekdays.

The overall behavior of the simulation is plausible and corresponds to the assumptions made in Section III-B.

B. AutoML

Figure 4 compares the AutoML model candidates created for the combinations of partition sizes (height = length) and time intervals using the metric R^2 score (coefficient of determination [16], best 1.0, worst 0.0). This metric is defined as follows:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

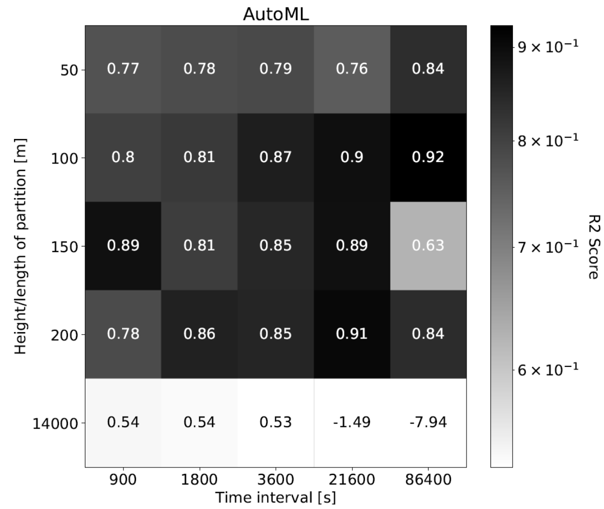

Fig. 4. R^2 scores for various AutoML model candidates. Each model was trained with a specific pair of partition- and time interval configuration.

Here, y_i is the true value of the i-th sample, $\hat{y}_i$ is the predicted value for the i-th sample, and $\bar{y}$ is the arithmetic mean [16].

It can be seen that the AutoML models with the greatest height/length of 14,000 meters produce inferior results. All other combinations of edge length and time interval give acceptable results. Apart from the bottom line (partition height/length 14,000), the quality of the models also tend to increase with larger time intervals. In addition, it can be seen that the best models have neither very small nor very large partition sizes. The best result is obtained by the model with partition height/length of 100 meters and a time interval of 86,400 seconds. These observations are also consistent when applying other metrics, such as Mean Squared Error (MSE). Here, the best combination found with R^2 also shows a low MSE (799.53) in contrast to the other combinations that exhibit poor R^2 results. For example, for the combination of 14,000 meters and 86,000 seconds, the MSE also reaches the overall worst result of 11,320,389.63. In summary, depending on the partition size and time interval, AutoML has created well-performing ML models based on the simulation output.

VI. CONCLUSION

Simulating PEV charging behavior can support numerous use cases in the context of electric mobility. We presented our corresponding simulation and showed that its results can

 ISBN: 978-1-61208-898-3

be applied to train useful ML models. Employing AutoML in our application example yields suitable ML models that can provide very good predictions for the short-term, regional energy demand that is induced by PEV charging operations.

We will further improve the simulation's underlying physical model to take into account additional relevant parameters, such as the outside temperature, battery temperature, and PEVs' charging curves.

REFERENCES

[1] G. Crabtree, "The coming electric vehicle transformation," *Science*, vol. 366, no. 6464, pp. 422–424, 2019.

[2] K. Chaudhari, N. K. Kandasamy, A. Krishnan, A. Ukil, and H. B. Gooi, "Agent-Based Aggregated Behavior Modeling for Electric Vehicle Charging Load," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 856–868, 2019.

[3] F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., *Automated Machine Learning: Methods, Systems, Challenges*. Springer, 2019.

[4] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent Development and Applications of SUMO - Simulation of Urban MObility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 128 – 138, 2012.

[5] Y. Xiang, Z. Jiang, C. Gu, F. Teng, X. Wei, and Y. Wang, "Electric vehicle charging in smart grid: A spatial-temporal simulation method," *Energy*, vol. 189, p. 116221, 2019.

[6] S. Bae and A. Kwasinski, "Spatial and Temporal Model of Electric Vehicle Charging Demand," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 394–403, 2012.

[7] Q. Yan, C. Qian, B. Zhang, and M. Kezunovic, "Statistical analysis and modeling of plug-in electric vehicle charging demand in distribution systems," in *2017 19th International Conference on Intelligent System Application to Power Systems (ISAP)*, 2017, pp. 1–6.

[8] "V2G-Sim," http://v2gsim.lbl.gov/, 2021, [Online; accessed 2021-08-24].

[9] Z. J. Lee, D. Johansson, and S. H. Low, "ACN-Sim: An Open-Source Simulator for Data-Driven Electric Vehicle Charging Research," in *2019 IEEE Int'l. Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2019, pp. 1–6.

[10] E. S. Rigas, S. Karapostolakis, N. Bassiliades, and S. D. Ramchurn, "EVLibSim: A tool for the simulation of electric vehicles' charging stations using the EVLib library," *Simulation Modelling Practice and Theory*, vol. 87, pp. 99–119, 2018.

[11] J. Zhu et al., "Electric Vehicle Charging Load Forecasting: A Comparative Study of Deep Learning Approaches," *Energies*, vol. 12, no. 14, 2019.

[12] M. Norman, "Introduction to Discrete-Event Simulation and the SimPy Language," *Davis, CA. Dept of Computer Science*, vol. 2, 01 2008.

[13] "auto-sklearn," https://automl.github.io/auto-sklearn/, 2021, [Online; accessed 2021-08-24].

[14] G. Chowell, J. M. Hyman, S. Eubank, and C. Castillo-Chavez, "Scaling laws for the movement of people between locations in a large city," *Phys. Rev. E*, vol. 68, pp. 661 021–661 027, Dec 2003.

[15] E. Mucelli Rezende Oliveira, A. Carneiro Viana, C. Sarraute, J. Brea, and I. Alvarez-Hamelin, "On the regularity of human mobility," *Pervasive and Mobile Computing*, vol. 33, pp. 73–90, 2016.

[16] J. Devore, *Probability and Statistics for Engineering and the Sciences*. Cengage Learning, 2015.

 ISBN: 978-1-61208-898-3