Forecasting Electric Vehicle Charging Behavior in Workplace Charging Infrastructure with Limited Privacy-Restricted Real Data

Alexander Stein*, Sebastian Beichter*, Jacek Hage, Maximilian Beichter, Bernhard Schwarz, Simon Waczowicz, Marc Hiller, Veit Hagenmeyer, Nina Munzke, Ralf Mikut
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany; * equal contribution
{firstname.lastname}@kit.edu

***Abstract*—Workplace Charging Stations (CSs) are well-suited to improve grid stability by scheduling the charging process over the parking duration and thereby reducing the peak load. Therefore, the energy demand and parking duration of single charging sessions must be known as well as the future occupancy of the CS. Since user IDs are often unknown for privacy reasons, this paper investigates how these parameters can be predicted for future charging events. The charging behavior is examined for its characteristic features, such as location, arrival, and departure times. First, calendar, weather, lag and CS-specific features are implemented and used to train nine different machine-learning algorithms. For the observed data, the Random Forest Regressor yields the best results for parking duration and energy demand. For parking duration, a 33.7% improvement in Mean Absolute Percentage Error (MAPE) over the baseline (the mean parking duration) can be achieved. The MAPE of the parking duration forecast is 71.0% and for the energy demand, it is 84.0% which leads to the conclusion that without the knowledge of user IDs predicting the charging behavior of users is possible only to a limited extent.**

ABBREVIATIONS

CS	Charging Station
EMS	Energy Management System
EV	Electric Vehicle
GB	Gradient Boosting
GBR	Gradient Boosting Regressor
LGBM	Light Gradient-Boosting Machine
Lin.-Reg.	Linear Regression
Log.-Reg.	Logarithmic Regression
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NN	Neural Network
Quant.-Reg.	Quantile Regression
RF	Random Forest
RMSE	Root Mean Squared Error
RT	Regression Tree
SVM	Support Vector Machine
XGB	eXtreme Gradient Boosting

I. INTRODUCTION

In recent years, the number of Electric Vehicles (EVs) and the associated charging infrastructure have constantly grown. Many countries are increasingly relying on decentralized renewable energy generation, such as photovoltaic systems. The additional load on the power grid caused by the charging infrastructure and the volatile power generation of renewable energies present challenges for grid stability. With the help of smart Energy Management Systems (EMSs), load peaks of CSs can be avoided by scheduling the charging load over the entire parking duration. CSs at workplaces are very suitable for this problem, as vehicles are usually parked for the entire working day. Therefore, many studies have focused on the optimization of charging infrastructure at workplaces [1]–[5]. For smart charging the energy demand and the parking duration of single charging sessions must be known as well as the future occupancy of the CS. Since user IDs are often unknown for data protection reasons, it is not possible to forecast individual charging profiles per user. The remainder of this paper is organized as follows: Related work is discussed and the contributions of the present paper are pointed out. Afterward, the dataset is analyzed and the method for generating the forecasting models is explained. Finally, the results and conclusions of this study are presented.

II. RELATED WORK

Existing research (cf. Table I) on the forecasting of CSs differs mainly in the forecasted parameters, and the use of user IDs while applying different forecasting models. Using user-specific information, Frendo et al. [6] compare different regression models to forecast the parking duration for employee parking based on historical data and optimize prioritization for smart charging. The dataset contains anonymized user IDs, so that driver-specific behavior can be identified. User-specific data is also used in the approaches of Schwenk et al. [8] and Huber et al. [7], which predict the parking duration per individual driver for different locations. Chung et al. [9] predict both the parking duration and energy demand of charging processes, thereby considering user IDs. Similarly, Almaghrebi et al. [10] predict the energy demand using user-specific information. Most approaches that do not use user-specific information predict the occupancy of CSs [12]–[16]. Only Arias et al. [11] forecast the energy demand based on historic traffic and weather data and the state of charge at the beginning of the charging process.

TABLE I: Literature comparison of forecasted parameters and used forecasting model

Reference	User-ID	Forecasted parameters	Used forecasting model
[6]	✓	Departure time	eXtreme Gradient Boosting (XGB)
[7]	✓	Parking duration, driving distance	Quantile Regression (Quant.-Reg.)
[8]	✓	Parking duration, departure time	Random Forest (RF)
[9]	✓	Parking duration, Energy consumption	Support Vector Machine (SVM), Random Forest (RF)
[10]	✓	Energy consumption	Gradient Boosting Regressor (GBR), Random Forest (RF), Support Vector Machine (SVM)
[11]		Energy consumption	Support Vector Machine (SVM)
[12]		Occupancy	Linear Regression (Lin.-Reg.)
[13]		Occupancy	Logarithmic Regression (Log.-Reg.)
[14]		Occupancy	Logarithmic Regression (Log.-Reg.)
[15]		Occupancy	eXtreme Gradient Boosting (XGB), Random Forest (RF), Gradient Boosting Regressor (GBR), Support Vector Machine (SVM)
[16]		Occupancy	Regression Tree (RT), Neural Network (NN), Support Vector Machine (SVM)

The contributions of the present paper are the analysis of typical workplace charging behavior, the comparison of CS forecasting models, and the examination of the limitations of predicting charging behavior when user IDs are not available. For this purpose, the amount of energy charged, parking duration, and occupancy are considered simultaneously in comparison to the comparable literature. In addition, we investigated how in the present use case clustering can improve the forecasts.

III. Characteristics of Workplace Charging

The forecasting models are based on charging data from the corporate parking lot of a German automotive supplier. The data, which covers approximately 16,000 charging sessions, was collected during the year 2022 from five different sites and twelve CSs, some of which are only accessible to employees (non-public). The information available for each charging session includes the site, CS-ID, start time, end time, and energy demand.

Table II lists the properties and statistics of the charging data including the median, the mean value, and the standard deviation for each CS. The individual CSs are divided into non-public and public. It should be mentioned that the standard deviation of the parking duration for some CSs is very high compared with the corresponding median and mean values.

Fig. 1 shows the proportion of start and end times (in percent) over the entire week (Monday - Sunday) of the charging sessions of the individual CSs. On the one hand, it can be seen that the start times are around 8:00-9:00 am local time, and the end times are very pronounced at workplaces and the curves show similarities.

The start and end times result in the parking durations, whose weekly distribution is shown as heat maps for two typical CSs in Fig. 2 according to the start time. As expected, there are no charging processes at the non-public charging station (Fig. 2a) on the weekend. A large proportion of the charging sessions last approximately 300 minutes (part-time employees) and 600 minutes (full-time employees) with an arrival time between 8:00-9:00 local time (typical for employees). With CS 4 (Fig. 2b), on the other hand, the charging sessions are more widespread over the day, considerably shorter, and also occur on weekends.

CSs can be grouped by similar parking duration distributions which leads to three clusters:

Parking duration cluster 1:

- Mostly charging sessions shorter than 300 min (Fig. 3a, bottom)
- Public CSs
- CS-IDs: 3, 4, 5, 12

Parking duration cluster 2:

- The majority of charging sessions last between 500 and 650 min. (Fig. 3a, center). This suggests that there were many full-time workers.
- Public and non-public CSs
- CS-IDs: 1, 7, 8, 9, 10, 11

Parking duration cluster 3:

- A distribution with two peaks, one at approximately 200 min and one around 600 min, suggests a mixture of part-time and full-time workers (Fig. 3a, top).
- Non-public CSs
- CS-IDs: 2, 6

In comparison, it can be seen that the distribution of the charged energy (Fig. 3b) varies slightly among the same CSs. The most common values are between 10-11 kWh.

IV. Methodology

The aim is to create forecasting models for the parking duration and energy demand of individual charging sessions.

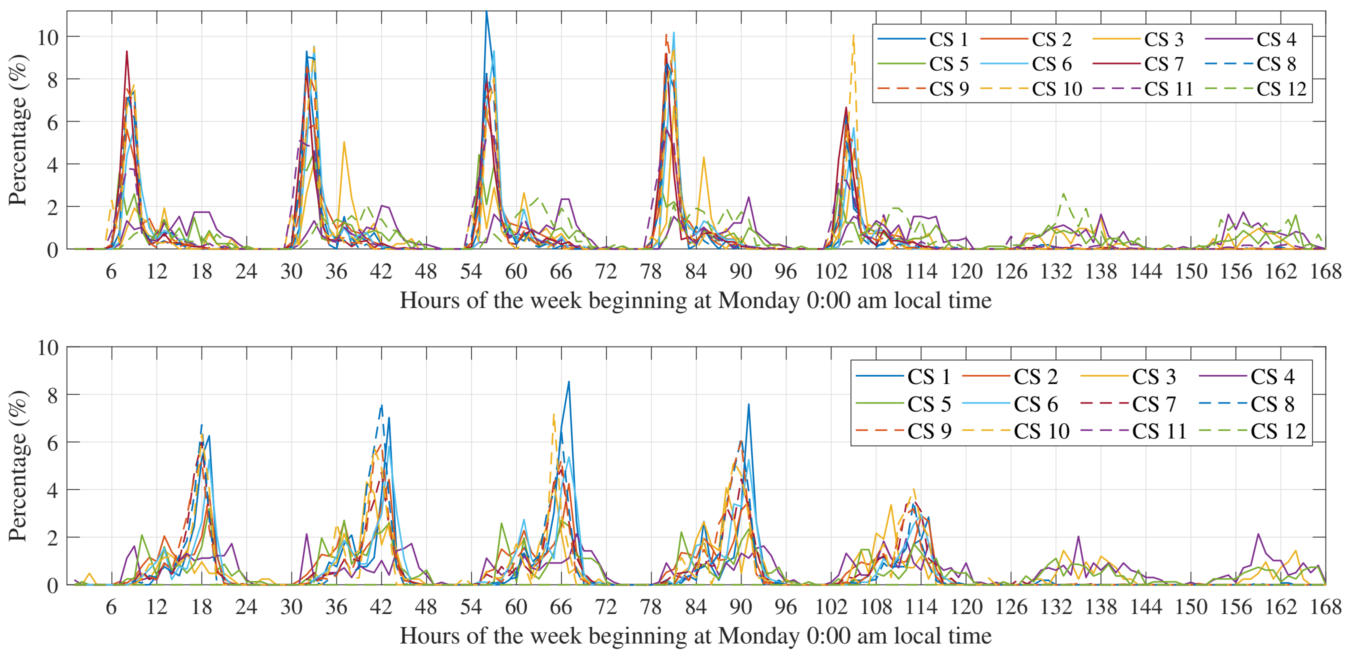
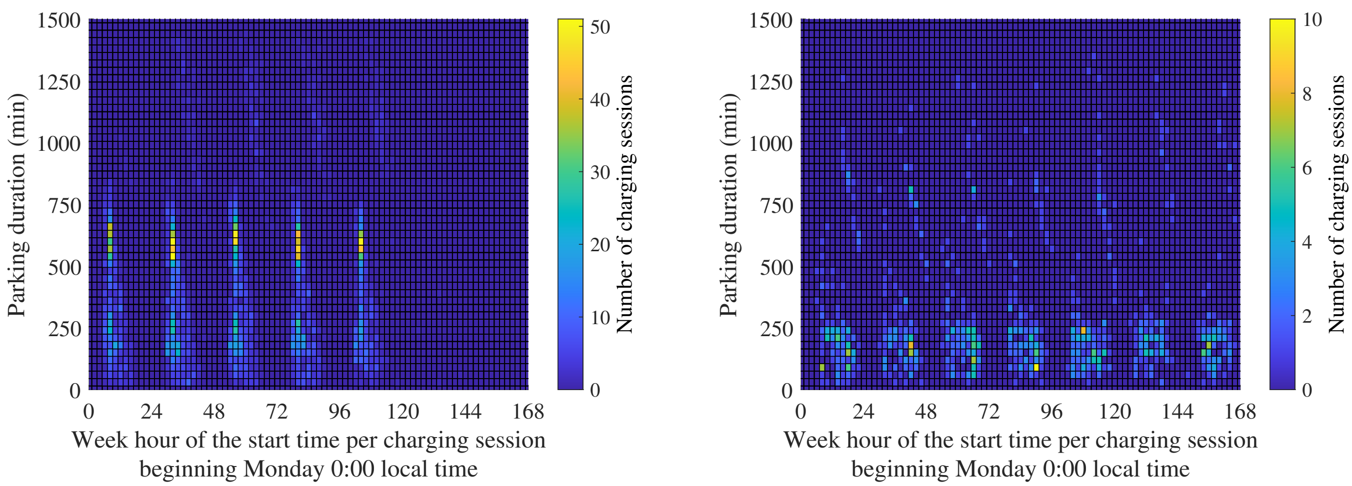


Fig. 1: Proportion of start (top) and end (bottom) times (in percent) over the entire week (Monday - Sunday) of the charging sessions of the individual CSs



(a) CS-ID 2 of parking duration cluster 3

(b) CS-ID 4 of parking duration cluster 1

Fig. 2: Distribution of the parking duration depending of the start time of the charging session per weekday

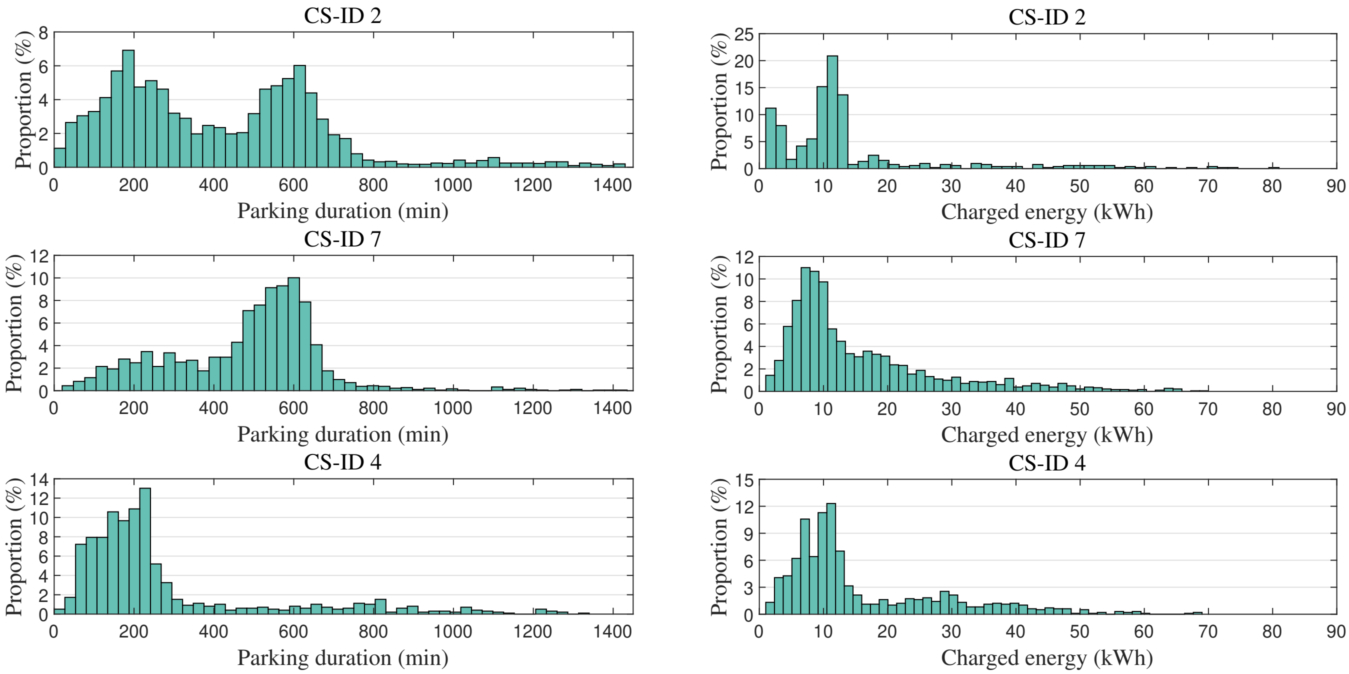
In addition, a forecasting model will be developed for the occupancy of a selected CS with eight charging points. After data cleaning and feature engineering, the best-performing models are selected. Subsequently, the influences of the features are analyzed and cluster-specific models for optimized forecasting are investigated. These steps are explained below.

A. Data Cleaning

During the cleaning process, charging sessions shorter than 15 min and sessions longer than 24 hours are removed because both are not relevant for load shifting. In addition, sessions with unrealistically high charged energy and charging power as well as charging sessions with less than 1 kWh of charged energy were removed because of implausibility. After cleaning the dataset approximately 15,800 charging sessions remain.

B. Feature Engineering

The features can be categorized into date-time features (time of day, weekday, month, calendar week), calendar features (holidays, school holidays, bridge days, boolean if the next day is a holiday), weather features (temperature, humidity, and wind speed), and CS specific features (site ID, CS ID, charge point ID, nominal power, publicly accessible). For the occupancy model, additional lagged features of up to one week from the past are used.



(a) Charging duration distributions.

(b) Charged energy distributions.

Fig. 3: Distributions of representative CSs from the different parking duration clusters: Charging duration cluster 1 (bottom), charging duration cluster 2 (center), charging duration cluster 3 (top)

TABLE II: Statistical data of the charging characteristics per CS based on the whole dataset

			Parking Duration			Charged Energy		
CS-ID	Public / Non-Public	Number Charging Processes	Median Parking Duration (min)	Mean Parking Duration (min)	Standard Deviation Parking Duration (min)	Median Charged Energy (kWh)	Mean Charged Energy (kWh)	Standard Deviation Charged Energy (kWh)
1	Non-Public	527	558	471.71	198.11	11.85	13.82	13.24
2	Non-Public	4004	381	414.11	267.24	10.43	13.86	11.15
3	Public	417	200	233.53	167.47	11.34	15.78	11.32
4	Public	983	200	278.99	251.70	10.94	16.06	12.51
5	Public	813	232	295.81	201.10	10.70	14.85	12.78
6	Non-Public	913	455	426.17	197.03	10.80	14.45	13.53
7	Non-Public	1817	522	479.83	190.94	10.75	15.57	12.19
8	Public	1306	523	479.91	141.78	9.78	14.62	10.64
9	Public	1752	496.5	450.70	155.22	11.89	18.52	14.43
10	Non-Public	347	476	438.36	166.79	11.36	19.81	16.99
11	Public	2354	498	437.07	177.29	11.05	17.36	14.36
12	Public	575	197	437.10	448.76	9.53	11.59	7.82
Total		15808	452.5	419.31	231.13	10.72	15.48	12.71

C. Training

For training a train-test-split of 80/20 is applied and the loss function is the Root Mean Squared Error (RMSE). In the case of time series data for the occupancy forecast, the last 20% of the time series is used for validation. The model-specific hyperparameters are optimized with the training data using Bayesian Optimization with a 5-fold-cross-validation.

D. Model Selection

In total, nine different machine learning models and two baseline models, within three different scope variables (parking duration, energy, and occupancy) are compared. Each model is optimized using hyperparameter optimization. For the parking duration and energy demand, the mean and median values are used as baseline models. The baseline model for the occupancy is the occupancy of the previous week. The model that performed best on the RMSE is selected for further analysis.

E. Analysis 1: Detailed feature analysis

In this step, the selected model is trained with combinations of feature sets. For each feature set, the model parameters are optimized using hyperparameter optimization.

F. Analysis 2: Evaluation of data clustering for forecast optimization

As mentioned previously, the CSs have different charging patterns. It is tested whether creating cluster-specific models can improve forecasting results. Therefore, each cluster was grouped according to whether the stations are private or public. The other two clustering categories are the parking duration clusters and CS specific models.

V. Results

First, the results for the parking duration and the charged energy are presented. The best model is further used to investigate the influence of features and clustering. Afterward, as the only time-series forecast, the results of the occupancy are shown.

A. *Model comparison for parking duration and energy demand*

The model comparison (Table III) shows that the Random Forest model performs best for the parking duration and the energy demand forecast in terms of the RMSE, followed by the Light Gradient-Boosting Machine (LGBM) model. For the parking duration, a MAPE of 72.6% is achieved, which corresponds to a Mean Absolute Error (MAE) of 135.3 min. The rather simple Linear Regression and Bayesian Ridge models achieve the worst results but still achieve better predictions than the base models.

For the energy demand, the Random Forest model achieved the best RMSE of 12.4 kWh. It must be emphasized, that the median baseline model is the best in terms of MAPE and MAE. This shows that the models were optimized for the RMSE.

In Fig. 4 the forecasted parking duration is compared to the mean and the actual observed parking duration. It can be observed that the algorithms can follow the curve with certain deviations for short and long parking durations.

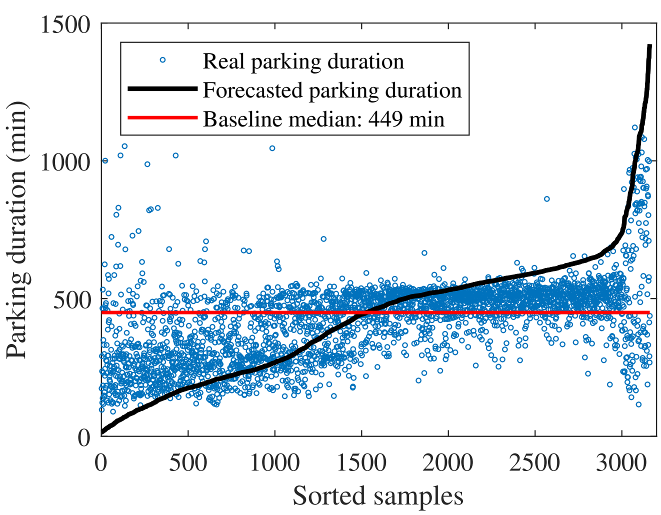


Fig. 4: Forecasting results for the parking duration of the test dataset with the Random Forest model

B. *Feature comparison*

By subsequently adding feature sets, it is possible to examine how the forecasting results change depending on the features. Because date-time features are cyclic, they can be converted into a cyclic representation using sine-cosine encoding. The results of the parking duration forecast (Table IV), which can also be transferred to the energy demand forecast, show that the algorithm works slightly better with the non-cyclic features. Furthermore, it can be seen that the weather features have no positive effect on the forecast and are therefore not used for further considerations.

C. *Clustering*

The aim is to investigate which forecasts can be improved if specific models are developed for similar CS. For this purpose, the Random Forest model with the continuous, calendar, and CS features is trained for the following specific grouped CSs and compared to the model that was trained on all station data:

- One model for each CS.
- One model for each parking duration cluster from Chapter III.
- Separate models for public and non-public CSs.

It can be observed that clustering generally only has a minor influence on the predictions (Table V). This means that the Random Forest model can learn relevant clusters from the entire dataset.

If a separate model is created for each CS, no improvement can be achieved compared to the universal model for all CS or the other clusters (Fig. 5). In the case of CS 1 and 8 the CS-specific model performs worse than the other models.

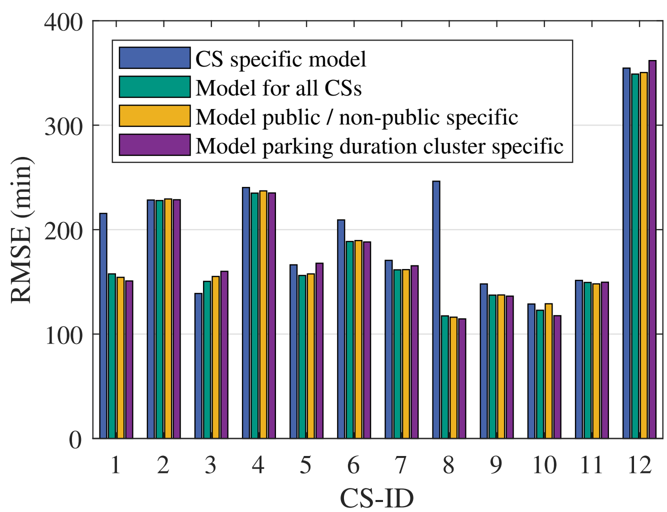


Fig. 5: Forecasting results for the parking duration of the test data set applying the described cluster variants using the Random Forest models evaluated for the individual CSs

D. *Occupancy forecasts*

Based on the session data, the occupancy of CS 2 with eight charge points is calculated and converted to a time-series with

TABLE III: Model comparison of the forecasting results of the test data for parking duration, energy consumption and occupancy

Modell	Parking duration MAPE (%)	Parking duration MAE (min)	Parking duration RMSE (min)	Charged energy MAPE (%)	Charged energy MAE (kWh)	Charged energy RMSE (kWh)	Occupancy MAE (%)	Occupancy RMSE (%)
Linear Regression	85.3	162.7	223.3	90.3	9.34	12.93	10.41	15.96
Decision Tree	73.4	142.0	204.7	88.1	9.16	12.81	8.46	16.21
Gradient Boosting Regressor	73.8	137.4	194.1	84.3	8.81	12.42	7.54	14.20
LGBM Regressor	73.7	137.6	193.9	84.2	8.78	12.41	**7.25**	13.98
XGBoost	73.9	137.9	197.0	84.4	8.83	12.45	7.45	**13.68**
Random Forest	**72.6**	**135.3**	**193.4**	**84.0**	**8.78**	**12.40**	7.25	14.20
Bayesian Ridge	85.4	162.9	223.4	90.4	9.35	12.93	10.41	15.97
Support Vector Machine	73.6	149.3	208.6	90.8	9.39	12.80	9.62	17.94
MLP Regressor	83.1	156.0	214.6	87.3	9.19	12.89	9.17	17.48
Baseline Mean	106.3	192.3	235.6	93.5	9.51	13.09		
Baseline Median	113.9	190.1	236.4	57.5	8.27	13.98		
Baseline Last week							9.18	17.46

TABLE IV: Comparison of the forecasting results of the test dataset depending on different feature sets for the parking duration with Random Forest

Features				MAPE	MAE	RMSE
Date-time	Calendar	CS	Weather	(%)	(min)	(min)
Continuous				84.9	159.7	219.6
Continuous	✓			84.1	158.9	219.1
Continuous	✓	✓		71.0	130.9	190.3
Continuous	✓	✓	✓	72.2	134.8	193.3
Cyclic				84.5	159.4	219.4
Cyclic	✓			84.4	159.4	219.3
Cyclic	✓	✓		71.9	132.3	190.8
Cyclic	✓	✓	✓	72.6	135.3	193.4

15 min intervals. In addition, lagged features are added that contain the values of the same time interval of the past days, as well as their mean values. Since the charging behavior during Christmas time differs a lot from normal charging behavior and there is only one year available the last week of the year 2022 was excluded. After the forecast, in post-processing, the forecasts are rounded to multiples of 12.5%. Comparing the forecasting models, XGBoost performed best in terms of the RMSE (Table III) with a 2% better MAE than the baseline model - the occupancy of the week before. On average, the forecast is off by less than one occupied charger, with an MAE of 7.25%. But this analysis also includes weekends and night times, which are easier to predict. The plot of the forecasts (Fig. 6) demonstrates, that the algorithm was able to learn that there are no charging events on weekends. However, it failed to learn the characteristics of holidays (e.g. November 1st, 2022).

VI. Conclusion and Outlook

This paper deals with the feasibility of predicting the charging behavior of EVs at workplace CSs which includes the parking duration and required energy per EV and the occupancy of the entire CS. These findings are important for the design and operation of CSs. The data set consists of anonymous charging sessions of one year providing the arrival and departure times and the amount of energy charged.

First, the charging data from different locations is processed and their characteristics are examined. Furthermore, various common machine learning algorithms are systematically applied to the dataset and their suitability is assessed. Our findings indicate that meteorological features have minimal impact, even resulting in a decline in the results. The lack of user-specific data posed a major challenge for the forecast, as it is not possible to differentiate between part-time and full-time employees.

Therefore, further studies should investigate how the results can be improved using user-specific data. Furthermore, it should be investigated how the importance of features, such as public holidays, changes when the dataset covers a longer period. We recommend investigating the influence of incentivizing users to specify the departure time to improve the precision of the forecast.

Acknowledgment

S. Beichter and A. Stein contributed equally to this work. This work contributes to the research performed at the Institute of Electrical Engineering (ETI) and the Institute for Automation and Applied Informatics (IAI) of KIT. The results were generated within the “SKALE” project (funding code: 01MV19004D), funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) and by the Helmholtz Association under the Program “Energy System Design”, and the German Research Foundation (DFG) as part of the Research Training Group 2153 “Energy Status Data: Informatics Methods for its Collection, Analysis and Exploitation”.

Authors Contributions

Author contributions according to Contributor Roles Taxonomy (CRediT, https://www.casrai.org/credit.html): Conceptualization: S.B., A.S., J.H., M.B.; Software: S.B., A.S., J.H.; Validation: A.S., S.B.; Visualization & Writing - original draft: S.B., A.S.; Writing - review and editing: S.B., A.S., R.M., M.B., S.W., N.M., B.S.; Supervision: R.M., B.S., N.M.; Funding acquisition: V.H, M.H., N.M.

TABLE V: Results of cluster-specific models and model trained by all data (in brackets) of the test data with Random Forest

	Parking duration			Energy Demand		
Cluster	MAPE (%)	RMSE (min)	MAE (min)	MAPE (%)	RMSE (kWh)	MAE (kWh)
Public	62.5 (64.4)	172.1 (172.6)	118.3 (118.5)	85.1 (85.8)	12.07 (12.11)	8.97 (9.0)
Non-public	67.8 (66.3)	187.2 (185.0)	138.6 (137.2)	73.4 (74.0)	11.34 (11.35)	7.44 (7.49)
Parking duration cluster 1	133.5 (131.2)	254.4 (252.5)	169.5/(166.6)	90.1 (89.4)	11.08 (11.02)	8.18 (8.14)
Parking duration cluster 2	42.8 (44.5)	145.3 (145.8)	102.6 (104.3)	86.0 (87.0)	12.17 (12.16)	9.01 (9.10)
Parking duration cluster 3	88.2 (89.5)	217.0 (214.5)	157.0 (156.4)	65.3 (66.5)	10.65 (10.65)	6.56 (6.66)

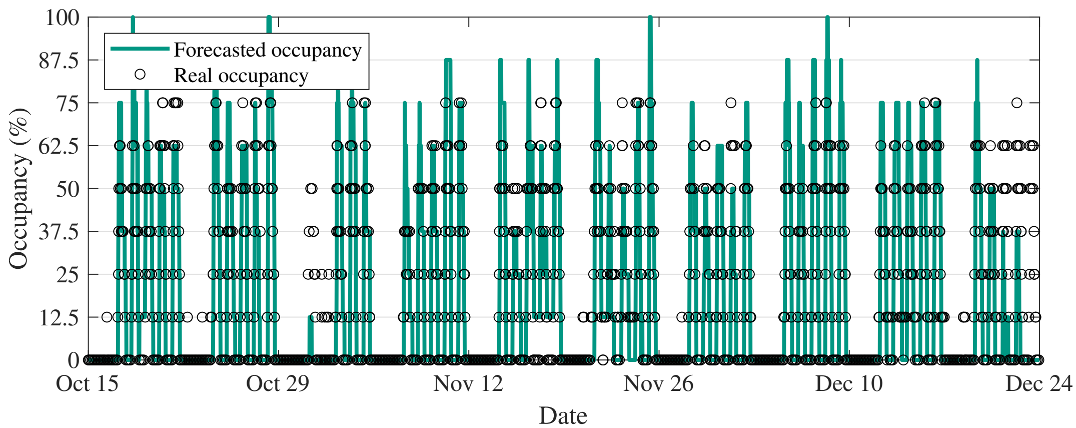


Fig. 6: Forecasted occupancy for the test period with XGBoost

References

[1] N. Erdogan, D. Pamucar, S. Kucuksari, and M. Deveci, "A Hybrid Power Heronian Function-Based Multicriteria Decision-Making Model for Workplace Charging Scheduling Algorithms," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 1, pp. 1564–1578, Mar. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9807339/

[2] S. Poniris and A. I. Dounis, "Electric Vehicle Charging Schedules in Workplace Parking Lots Based on Evolutionary Optimization Algorithm," *Energies*, vol. 16, no. 1, p. 221, Dec. 2022. [Online]. Available: https://www.mdpi.com/1996-1073/16/1/221

[3] C. B. Saner, A. Trivedi, and D. Srinivasan, "An EV Charging Scheduling Methodology to Reduce Demand and Energy Charges in Industrial and Commercial Sites," in *2022 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*. Singapore, Singapore: IEEE, Nov. 2022, pp. 285–289. [Online]. Available: https://ieeexplore.ieee.org/document/10003640/

[4] P. J. Tulpule, V. Marano, S. Yurkovich, and G. Rizzoni, "Economic and environmental impacts of a PV powered workplace parking garage charging station," *Applied Energy*, vol. 108, pp. 323–332, Aug. 2013. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0306261913001876

[5] F. Welzel, C.-F. Klinck, Y. Pohlmann, and M. Bednarczyk, "Grid and user-optimized planning of charging processes of an electric vehicle fleet using a quantitative optimization model," *Applied Energy*, vol. 290, p. 116717, May 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S030626192100235X

[6] O. Frendo, N. Gaertner, and H. Stuckenschmidt, "Improving Smart Charging Prioritization by Predicting Electric Vehicle Departure Time," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6646–6653, Oct. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9082829/

[7] J. Huber, D. Dann, and C. Weinhardt, "Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging," *Applied Energy*, vol. 262, p. 114525, Mar. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0306261920300374

[8] K. Schwenk, K. Phipps, B. Briegel, V. Hagenmeyer, and R. Mikut, "A Benchmark for Parking Duration Prediction of Electric Vehicles for Smart Charging Applications," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. Orlando, FL, USA: IEEE, Dec. 2021, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/9660063/

[9] Y.-W. Chung, B. Khaki, T. Li, C. Chu, and R. Gadh, "Ensemble machine learning-based algorithm for electric vehicle user behavior prediction," *Applied Energy*, vol. 254, p. 113732, Nov. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0306261919314199

[10] A. Almaghrebi, F. Aljuheshi, M. Rafaie, K. James, and M. Alahmad, "Data-Driven Charging Demand Prediction at Public Charging Stations Using Supervised Machine Learning Regression Methods," *Energies*, vol. 13, no. 16, p. 4231, Aug. 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/16/4231

[11] M. B. Arias and S. Bae, "Electric vehicle charging demand forecasting model based on big data technologies," *Applied Energy*, vol. 183, pp. 327–339, Dec. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0306261916311667

[12] C. Bikcora, N. Refa, L. Verheijen, and S. Weiland, "Prediction of availability and charging rate at charging stations for electric vehicles," in *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. Beijing: IEEE, Oct. 2016, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7764216/

[13] M. Motz, J. Huber, and C. Weinhardt, "Forecasting BEV charging station occupancy at work places," 2021, iSBN: 9783885797012 Publisher: Gesellschaft für Informatik, Bonn. [Online]. Available: http://dl.gi.de/handle/20.500.12116/34779

[14] F. Soldan, E. Bionda, G. Mauri, and S. Celaschi, "Short-term forecast of EV charging stations occupancy probability using big data streaming analysis," Apr. 2021, arXiv:2104.12503 [cs]. [Online]. Available: http://arxiv.org/abs/2104.12503

[15] C. Hecht, J. Figgener, and D. U. Sauer, "Predicting Electric Vehicle Charging Station Availability Using Ensemble Machine Learning," *Energies*, vol. 14, no. 23, p. 7834, Nov. 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/23/7834

[16] Yanxu Zheng, S. Rajasegarar, and C. Leckie, "Parking availability prediction for sensor-enabled car parks in smart cities," in *2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. Singapore: IEEE, Apr. 2015, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7106902/