

## Forecasting flexibility of charging of electric vehicles: tree and cluster-based methods

Evgenii Genov<sup>1</sup>, Gilles Van Kriekinge<sup>1</sup>, Cedric de Cauwer<sup>1</sup>, Thierry Coosemans<sup>1</sup>, Maarten Messagie<sup>1</sup>

<sup>1</sup>EVERGi Research Group, MOBI Research Centre & ETEC Department, Vrije Universiteit Brussel (VUB),

Pleinlaan 2, Brussel, Belgium, evgenii.genov@vub.be

---

### Executive Summary

Scheduling electric vehicle charging sessions allows to aggregate flexibility in order to minimize energy costs and reduce congestion in the electricity grid. Existing research shows that user input for energy and parking duration does not serve as a reliable prediction. The study presents an evaluation of the forecast error and computational performance of the different models. Two main methods are investigated: tree-based (gradient boosted trees LightGBM) and cluster-based (Gaussian mixture model). We also present a novel cluster-based method, the Similar Sessions method, which employs the similarity between charging sessions based on numerical variables. The results highlight the importance of selecting the forecast model influenced by the availability of training data. The effect of user registration on the accuracy of the forecast is investigated. The tests are run using a *ACN-Data* dataset of Electric Vehicle charging sessions in California, United States. While underperforming on a small dataset with a short look-back period, tree-based methods show superiority while the charging data are accumulating. The Similar Sessions method shows superior accuracy under various data availability conditions. The proposed method requires no prior training, but has slower computational performance in deployment.

*Keywords: Flexibility, Forecasting, GMM, LGBM, User engagement, Smart charging, Regulation market, Grid congestion*

---

## 1 Introduction

Global transport sector relies strongly on fossil fuels and is estimated to be responsible for nearly 37% of CO<sub>2</sub> emissions. This number includes a majority contribution from road transport, which continues to rely largely (97%) on oil products [1]. The electrification of road vehicles is recognized by governments as a promising pathway to reducing greenhouse gas emissions. The car market echoes this ambition, as electric cars made up almost 10% of global car sales in 2021 and continued to grow in 2022 [2]. The rapid expansion of electric vehicle (EV) fleets is accompanied by a corresponding increase in investments in public charging infrastructure. Most cars remain stationary for extended periods; therefore, electric vehicle batteries and associated charging units often remain idle. The resulting discrepancy between the minimal necessary and actual duration of charging sessions creates flexibility, which can be valuable both in energy market arbitrage and in providing services aimed at balancing the electricity grid. EV batteries, managed in sufficiently large aggregations, can provide power flexibility to participate in ancillary service markets, for example, frequency restoration. Such services are referred to in literature as "regulation" or "secondary regulation". [3]

An optimally scheduled charging process for electric vehicles (EVs), often referred to as smart charging, makes use of time and energy flexibility of the vehicles. A controllable charger device can adjust the power and time to match the needs of the grid and the user. Making electric vehicle chargers fully controllable changes the traditional view on load forecasting. The differences between approaches are demonstrated in Figure 1. The principal difference

between the underlying forecasting approaches is the limited predictive power of historical load data. When an EV charger is fully controllable, charging patterns of the past may not be indicative of the charger’s future behavior. Instead, real-time data and a schedule of control signals are more indicative of the charger’s future load profile. However, the forecast of user arrivals, as well as energy demands and session duration of the charging sessions, are essential parameters for scheduling optimization and estimating the potential flexibility resource. Decoupling the modeling of user arrivals and charging session specifications, as shown in 1b, is advantageous in scale-up studies when the known user behavior is interpolated for an anticipated increase in EV drivers. In some pilot sites [4] [5], it is a well-established practice to collect user-specified session requirements directly from the user. These preferences can be taken as parameters to optimize the operation of the electric vehicle charger. However, as noted in Lee et al. [4], users lack the incentive to provide accurate predictions for their parking duration and energy demands. To improve the accuracy of the predictions, data-driven forecasting methods can be used to estimate the expected duration and energy output. Additionally, due to the inherent uncertainty in user behavior, it is beneficial to perform stochastic optimization. Although a plethora of methods have been proposed in the literature, there is no one-size-fits-all approach. In this paper, we demonstrate the relation between the model selection and the availability of training data. We study the effect of user identification on forecast accuracy. Moreover, we present a simple and robust ‘Similar Sessions’ approach, where the predicted value is selected from a sample of numerically close charging sessions. We show its competence in forecasting user behavior in EV charging applications.

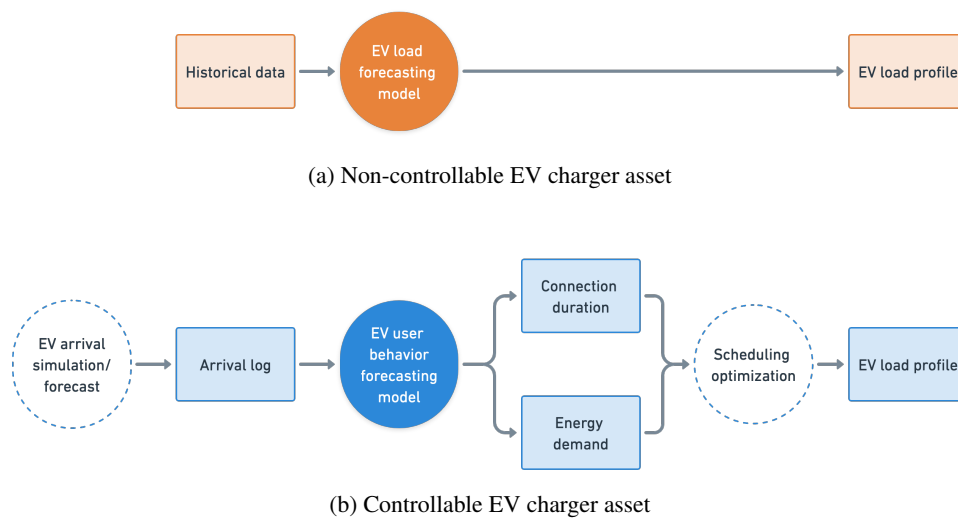


Figure 1: Schematic diagram explaining the difference in modelling approaches used in estimation of aggregate EV charging load profiles

## 1.1 Contribution and structure of the paper

This paper aims to contribute to the topic of forecasting the behavior of EV users by examining the crucial issue of selecting an appropriate forecasting model based on the availability of training data. We revisit the dataset presented in [4] and contribute more analysis of the data intensiveness of the state-of-the-art models. We study the effect of user identification data on the accuracy of the forecast. Furthermore, we propose a novel ‘Similar Sessions’ (SimS) approach.

1. We evaluate the performance of cluster-based (GMM) and tree-based (LGBM), as well as the model sensitivity to the data availability and the availability of user identification.
2. We introduce a novel Similar Sessions (SimS) approach that provides an efficient and reliable means of forecasting EV charging data with user logs.
3. We evaluate the effectiveness of the SimS approach in terms of forecast accuracy and computational performance.

The remainder of the paper is structured as follows. In Section 2, we provide an overview of the literature on existing methods for forecasting EV charging behavior. In Section 3, we describe the dataset, analyze the statistical properties of the data, and explain the train-test partition. Section 4 outlines the evaluated forecasting methods, their core principles, and the evaluation methodology. The section includes the description for our proposed ‘Similar Sessions’ approach, including its implementation. In Section 5, the results for the accuracy evaluation on a test set are presented and discussed. Finally, in Section 6, we present our conclusions and suggestions on the directions for future research.

## 2 Literature review

There is a growing body of research on modeling the electric load of electric vehicle charging. Most articles present methods to forecast the aggregated load of the electric vehicle fleet. Article [6] proposes an aggregate EV charging demand forecasting method based on the auto-regressive integrated moving average (ARIMA) model to predict charging demand for electric vehicles. The uncertainty in charging behavior is addressed by solving a chance-constrained scheduling problem using the aggregate forecast as input.

EV charging data commonly shows patterns in user behavior, which can be captured efficiently by clustering methods. Using data from individual meters at charging stations, the authors of the article [7] identify five types of batteries defined by energy and power capacities, and four main clusters of charging habits. This information is used to generate probabilistic forecasts of the aggregated load, as well as day-ahead consumption scenarios for a single EV and the aggregated fleet. The proposed approach is found to be more efficient than the naive and gradient-boosting benchmark models. The improvement is notable in modeling the lower tail of the distribution.

The increasing popularity of machine learning methods has found many applications in research on forecasting EV charging loads. Van Kriekinge et al. [8] study long short-term memory network (LSTM) configurations with a dynamic learning rate to forecast the day-ahead charging demand of electric vehicles with a 15-minute time resolution. The added benefit of exogenous weather variables is also shown. In [9], multiple models were evaluated for one-step ahead load prediction at city aggregation and charger level. The study finds that multivariate models such as ARIMA, ANN, and LSTM show higher accuracy than univariate models. The TBATS model is useful when only univariate values are available, and the LSTM model performs best at the individual charger level. The authors also analyze the data intensiveness of different models, i.e., the sensitivity of the forecast prediction to the size of the training set. The authors observe that the size of the training data does not necessarily improve the accuracy of the predictions, particularly in the LSTM models. However, more historical data make the forecast more robust to long-term perturbations.

An alternative view of modeling the demand for EV charging is to model the demand for individual charging sessions. Therefore, the effective aggregate demand for an EV charging pool can be simulated by adding the charging profiles of individual charging sessions, subject to scheduling optimization. The appropriate target variables are the duration of parking and the energy demands for each charging session. These parameters make useful descriptors of the flexibility potentials, made from the difference in time between minimal time to reach a full state-of-charge (SOC) and the real duration between arrival and departure at the station. There is a wide range of forecasting methods for predicting the behavior of EV users. For a comprehensive review, we recommend to refer to [10]. On this account, select aspects relevant to the discussion and the subject of this research are discussed.

Some researchers take estimates of departure time and energy demands as declared by users [11]. The paper [4] presents ACN-Data, a dynamic dataset of workplace EV charging with over 2 years of data and more than 30,000 sessions at two sites, Caltech and JPL campus. The paper uses the data and presents a method utilizing Gaussian mixture models (GMMs) to predict the charging duration and energy delivered to drivers. The authors explore two approaches, generating a population-level GMM (P-GMM) based on the overall training data and training individual-level GMMs (I-GMMs) for each user by fine-tuning the weights of the components of the P-GMM with sessions linked to a specific user. The derived approaches significantly beat the estimates provided by EV users directly, making the case that data-driven models are a more reliable prediction for these parameters. The authors include an analysis of the prediction accuracy with different training data sizes. The authors demonstrate a trade-off between prediction quality and data size, with the best performance found using a 30-day training set on a Caltech site.

Huber et al. [12] develops a forecasting method using German data of travel logs from 6465 car users. The aim is to generate probabilistic forecasts of the duration of parking and the energy requirements. The study employs various machine learning algorithms, including quantile regression, multi-layer perceptrons with tilted loss function, and multivariate conditional kernel density estimators. The results show that the use of probabilistic forecasts is shown to lead to more efficient scheduling compared to point forecasts, resulting in fewer interruptions and increased driver mobility.

Among many publications of machine learning applications in the EV field, the decision tree methods often demonstrate the best performing models. Applied to the charging behavior of electric vehicles, tree-based methods are found in several implementations, including Random Forest [13] and Gradient-boosted Decision Tree (GBDT) implementations, such as XGBoost [14] or LightGBM [15]. The dominance of tree-based methods has also been prominent in other applications dealing with tabular data. The GBDT algorithm is reported [16] to get top places in most of forecasting competitions on Kaggle, a popular online platform for data scientists and machine learning practitioners. GBDT is found to be particularly effective in modeling external information, dealing with in-sample-level shifts, and cross-learning information from similar series. GBDT shows superior performance both in estimating point predictions and in modeling uncertainty. [17] [18] The resilience to adaptation of tree-based methods is often attributed to the inability to extrapolate trends (without appropriate preprocessing), as well as poor performance of the method for small datasets. The relationship between the type of forecasting algorithm and the size of the data is investigated in the following paper. Additional considerations are made towards the application of flexibility forecast in an energy management system.

## 3 Dataset

For the purpose of this research we use an open-source *ACN-Data* dataset [4], a collection of data on charging data, user input, and the respective time series of electric vehicle charging power. The dataset provides details about



demonstrate a discernible daily seasonality, particularly during the weekdays, where morning arrivals typically stay connected within the same day, while evening arrivals exhibit a greater variance in parking time. Similar patterns are also observed in energy delivery data, with vehicles arriving in the evening being supplied with more energy on average. However, the distributions for weekends appear more compact, possibly due to a limited dataset comprising only sessions from workplace charging. The clear differentiation between weekdays and weekends suggests a weekly seasonality.

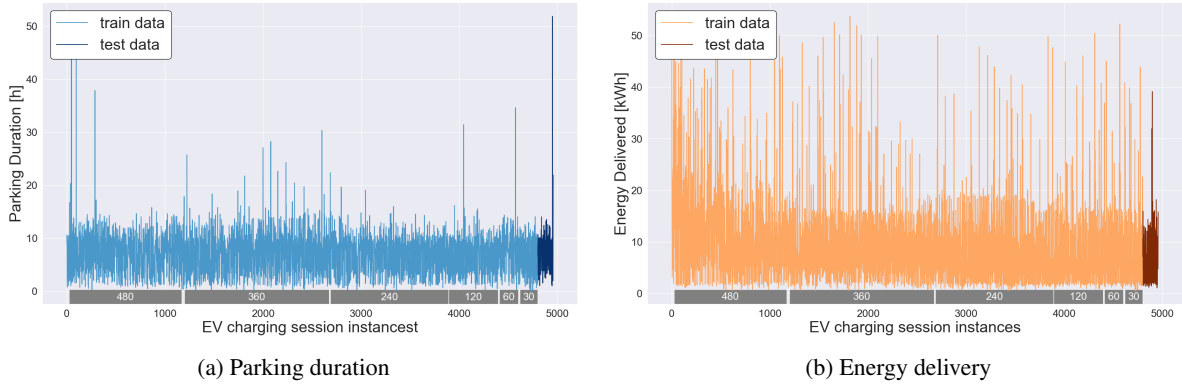


Figure 2: Time-Series of training and test partitions for Caltech dataset

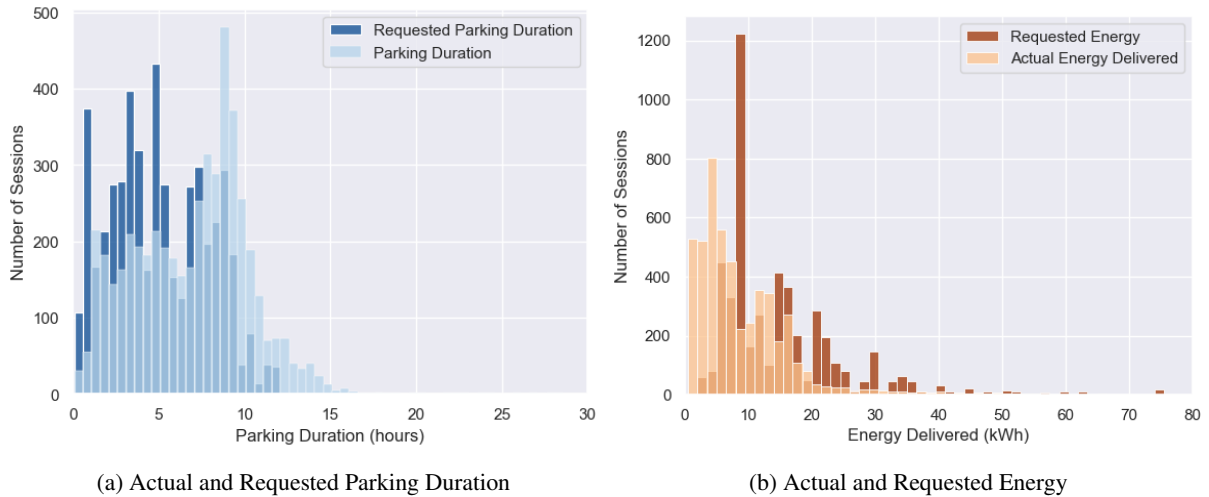


Figure 3: Histogram distributions show the disparity between the requested and actual EV charging specifications.

## 4 Method

Based on the arrival time and user identification, we make predictions of waiting time and energy demands of the electric vehicle. Cluster-based Gaussian mixture models and a tree-based gradient-boosted model are tested and compared for different look-back periods, indicative of data size.

### 4.1 Gaussian Mixture Models

Gaussian mixture models are used to model the behavior of EV drivers based on their charging patterns. The complete description of the method is provided in [4]. The data is represented by a set of vectors  $X_i = (a_i; d_i; e_i)$  with three values: arrival time  $a_i$ , duration  $d_i$ , and total energy delivered  $e_i$ . The GMM model assumes that the data points are generated from a finite number of typical profiles, and the noise is Gaussian, which allows the model to cluster data points based on their profiles. The GMM is fit by estimating the parameters  $\theta = (\pi_k; \mu_k; \Sigma_k)_{k=1}^K$  where  $\pi_k$  is a probability that the EV user has the profile  $k$ ,  $\mu_k$  is a means vector of a Gaussian random variable, and  $\Sigma_k$  is a covariance matrix. The probability density of observing a data point  $x$  can then be approximated using the GMM. Assuming that the arrival time  $a_j$  of user  $j$  is known a priori, the model can predict the duration  $d_j$

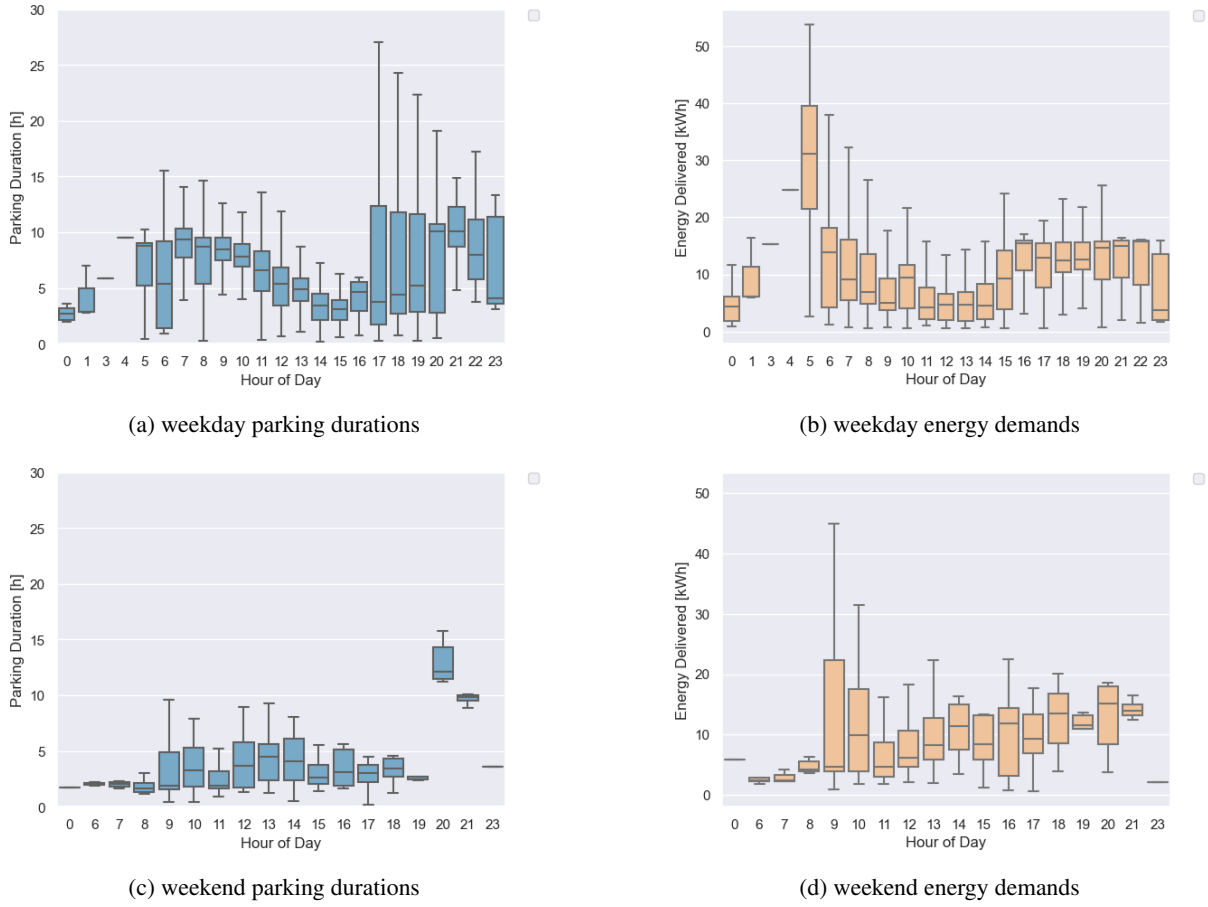


Figure 4: Box plots depicting the distribution of energy delivery and parking duration of EV charging sessions by hour of arrival. Box plots summarize the distribution of data by displaying the median and quantile ranges. Outliers are dismissed for clarity

and energy  $(j)$  to be delivered by user  $j$  as conditional Gaussians. These predictions are made using formulas that depend on the estimated parameters  $(j) = ( (j); (j); (j) )_K$  for the user  $j$ , where  $(j)$  is a vector that contains the estimated arrival time, duration, and energy for the user  $j$ .

$$(j) = \sum_{k=1}^K (j)_k a_k^{(j)} + (j) a_k^{(j)} \frac{(j)(1;2)}{(j)(1;1)} \quad (1)$$

$$e^{(j)} = \sum_{k=1}^K (j)_k e_k^{(j)} + (j) a_k^{(j)} \frac{(j)(1;3)}{(j)(1;1)} \quad (2)$$

where  $(j)(1;1)$ ,  $(j)(1;2)$  and  $(j)(1;3)$  are the first, second, and third entries of the covariance matrix.  $(j)_k$  are the modified weights conditioned on arrival time.

$$k := \frac{p^{(j)} j a_k^{(j)}; (j)(1;1)}{\sum_{k=1}^K p^{(j)} j a_k^{(j)}; (j)(1;1)} \quad (3)$$

Two training approaches are tested for GMMs: population and individual level. On the individual level, the model is tuned on the data set of sessions with the same user.

## 4.2 Gradient boosted trees

Light Gradient Boosted Machine (LGBM) is a high performance gradient boosting framework that uses tree-based learning algorithms. It is designed to be efficient and scalable for large datasets and can handle sparse data. In

contrast to traditional gradient boosted algorithms that grow trees depth-wise, LGBM grows trees leaf-wise, which reduces the number of instances traversed during tree building and improves training efficiency. Additionally, LGBM supports categorical features, which can be converted into numerical features using one-hot encoding. The default configuration is used while fitting an LGBM model. While training, an objective function is set to minimize the mean squared error of the predictions. No extra hyperparameter search was performed. However, some additional feature engineering is performed. The complete sequence of steps in preprocessing and modeling is shown in the diagram in Figure 5. LGBM is a univariate model, so the training is run twice for each of the target variables: the duration of the charging session and the amount of electricity delivered. During the initial data exploration of Figure 2, we found that the time series for parking duration exhibits signs of non-stationarities in data distribution. The variance of the data is not constant over time. To address this issue, we performed a log transformation on the parking duration target variable, a measure widely used in time-series analysis. Taking the natural logarithm compresses the range of values while reducing the influence of extreme values.

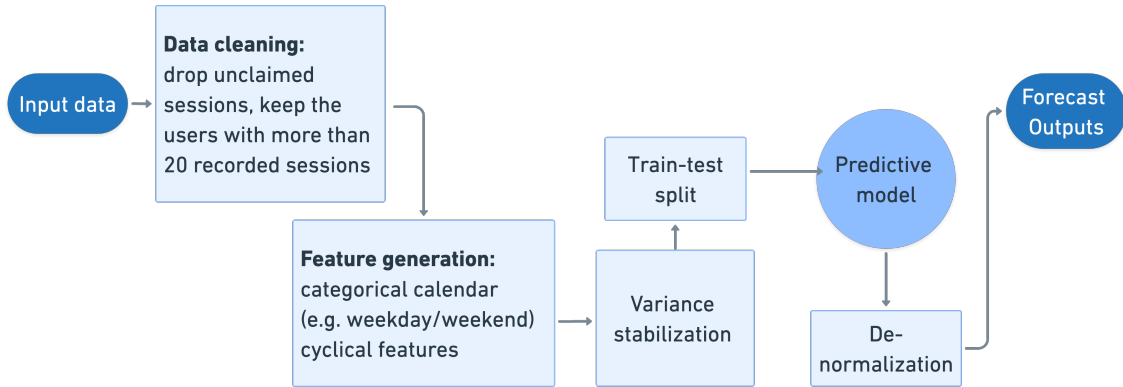


Figure 5: Workflow and dataflow diagram

Table 2: Hyperparameters of LGBM

Hyperparameter	Description	Range
num_leaves	Maximum number of leaves in one tree	[2, 200]
max_depth	Maximum depth of tree	[2, 5]
learning_rate	Learning rate	[0.001, 0.1]
min_data	Minimum number of data in one leaf	[1, 100]
lambda_1	L1 regularization	[0.001, 1000]
lambda_2	L2 regularization	[0.001, 1000]

### 4.3 Similar sessions

In load forecasting, the similar day method has long been a reliable technique for predicting energy consumption. By analogy, we propose a novel method, called the similar session method (SimS), to address the challenge of predicting the charging behavior of EV users. However, similarity-based methods deal with a common challenge in identifying similarity patterns. This approach draws inspiration from collaborative filtering, a technique commonly used in movie recommendation systems. Although clustering techniques are often used for this purpose, they come with certain limitations, such as the need for model selection, hyperparameter search, and the treatment of categorical variables.

To overcome these issues, we propose measuring similarity between charging sessions based on the cosine distance between numerical variables, including sine and cosine cyclical features of hour, month, and year, as well as a dummy variable for the user identifier. The cosine distance can be calculated using the formula:

$$\text{COS}(\mathbf{s}_1; \mathbf{s}_2) = \frac{\mathbf{s}_1 \cdot \mathbf{s}_2}{\|\mathbf{s}_1\| \|\mathbf{s}_2\|} \quad (4)$$

where  $\mathbf{s}_1$  and  $\mathbf{s}_2$  are the numerical feature vectors of two charging sessions. The cosine distance is calculated between the feature vector of the given session and the feature vectors of all other sessions in our dataset. Sessions with cosine distances close to 1 are considered the most similar, while those with distances close to -1 are considered the least similar. Once the most similar charging sessions have been identified, a selected number of corresponding target values (e.g. parking times) are used to calculate a mean value, which is then used as the predicted target value for the given session. This approach is based on the assumption that the similar session method

takes advantage of the patterns observed in both same-user charging sessions and sessions with matching arrival times. The design of the method entails one hyperparameter that denotes the number of similar sessions used for calculations.

This approach is expected to be particularly efficient in the context of modeling the behavior of EV users. The method is meant to be very efficient in terms of data intensiveness. Compared to smart-meter data, charging session data is typically sparse and scattered over time. Furthermore, the limited data pool enables the calculation of the cosine distance to take minimal calculation time and memory.

#### 4.4 Evaluation method

The performance of four models, LGBM, P-GMM, I-GMM, and SimS, in predicting charging sessions is evaluated. The error of the models is measured using the mean absolute error (MAE), which is defined as the average of absolute differences between predicted and actual values. The formula for MAE is the following:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

where  $n$  is the number of observations,  $y_i$  is the actual value and  $\hat{y}_i$  is the predicted value. One advantage of MAE is that it is measured in data-native units, which makes it easier to interpret and compare the error between different models. The computation times for training and deployment of the forecasting models are measured using the same hardware setup - a 2020 M1 Macbook Pro.

## 5 Results and Discussion

Table 3: Mean Average Error of parking duration forecasts [h]

Look-back period	Logged User						Benchmark	
	LGBM	Similar sessions	LGBM-L	SimS-L	P-GMM	I-GMM	Mean	User Input
30	2.690	2.354	2.690	1.947	4.159	2.329	<b>1.764</b>	2.534
60	2.565	2.151	2.147	1.774	4.477	1.711	<b>1.534</b>	2.534
120	2.539	2.174	2.159	<b>1.767</b>	6.217	3.242	1.812	2.534
240	2.443	2.219	1.918	<b>1.636</b>	5.752	3.640	1.694	2.534
360	2.460	2.154	1.827	<b>1.597</b>	6.896	3.971	1.997	2.534
480	2.463	2.176	1.845	<b>1.578</b>	6.595	3.996	2.029	2.534

Table 4: Mean Average Error of energy delivery forecasts [kWh]

Look-back period	Logged User						Benchmark	
	LGBM	Similar sessions	LGBM-L	SimS-L	P-GMM	I-GMM	Mean	User Input
30	5.002	3.907	5.002	<b>1.597</b>	2.066	1.638	1.823	7.174
60	4.063	4.211	1.982	<b>1.413</b>	1.890	1.491	1.809	7.174
120	3.540	4.031	1.554	<b>1.102</b>	1.996	1.644	3.383	7.174
240	3.410	3.850	1.509	<b>1.072</b>	1.873	1.519	3.696	7.174
360	3.693	3.807	1.446	<b>1.195</b>	2.006	1.745	4.138	7.174
480	3.460	3.862	1.435	<b>1.177</b>	1.926	1.770	4.093	7.174

Tables 3 and 4 report measurements of mean absolute error (MAE) of predictions produced by models trained with datasets of different look-back periods. The visualization of the errors is given in Figure 6.

We observe that when compared to the benchmarks, most of the forecasting methods succeed in beating the user input estimations. The other benchmark, the mean of the variable from the same user, provides estimates with better precision. To assess the influence of user registration, we evaluate the performance of the model with and without this input. The LGBM and SimS configurations without user logs perform significantly worse. However, the results indicate that it is still possible to make predictions with higher accuracy than the estimates derived from user requests.



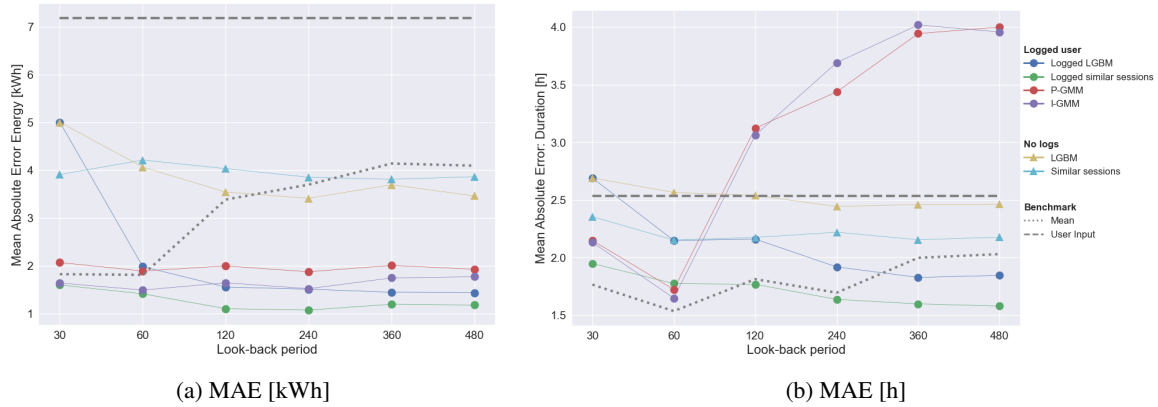


Figure 6: Exploring the Impact of Varying Training Set Size on Mean Absolute Error Metrics for Model Predictions

While the look-back period increases, the error of the historical mean increases in both energy and connection time estimations. This is expected since over longer periods the distributions of the variable become more diverse, and taking the mean value does not incorporate the correlations between arrival time and the target variables. The fact that the historical mean of the same user benchmark hits the lowest error among all models when predicting the connection time of EVs, particularly using the data of the last 60 days, suggests that there is a high degree of consistency in the data. This may be due to the fact that the data describe charging sessions of EVs in a workspace environment, where people tend to have a regular schedule. However, this is not true for the prediction of the energy demands, where the I-GMM and Logged Similar Sessions (L-SimS) models outperform the mean benchmark.

It is observed that the prediction error generated with a gradient-boosted model generally decreases as the look-back period increases. The model is able to capture more of the underlying patterns and trends in the data as more historical data is incorporated into the model. This is a common observation with machine learning models. The LGBM algorithm may not be able to learn the underlying patterns in the data, as well as other algorithms designed for small datasets. The reason lies in the leaf structure of the decision trees that grow [19]. The other model designs, GMM and SimS, are less data intensive. The Gaussian mixture model is a probabilistic model that assumes that the data are generated from a mixture of Gaussian distributions. Therefore, GMMs can make use of a prior assumption about the Gaussian nature of the variable distribution to make better predictions in data-scarce environments. Sim-S shows stable error performance across different look-back periods, as Sim-S selects a set of similar sessions that are likely to be close on a temporal scale. The method defines similarity based on a vector that includes temporal features.

The quality of the predictions produced with GMM-based algorithms varies between the variables of energy demand and parking duration. The error is less aligned with the size of the data. When the look-back period of the training set is increased, the quality of the prediction for the duration of parking improves initially. However, the error increases to a higher level with datasets that exceed 120 days in length. This occurrence is indicative of an existing trade-off within GMM models between model performance and data size, also observed in [4]. Analyzing the plots for time-series in Figure 2 and the techniques employed, we hypothesize that the main cause for the error swing is the lack of models' ability to adapt to data drifts and non-stationarities in parking duration time series. In turn, LGBM employs variance stabilization, while the SimS model is designed to select a sample of recent historical sessions.

Our assessment of model performance also includes an evaluation of the computational speed taken in training and deployment of forecasting models. The results are displayed in Tables 5 and 6. SimS model is not included in the evaluation of training time, since the similarity scores are calculated at each instance the model is deployed. From the table 5, it can be observed that the I-GMM model is significantly faster in terms of training time compared to the LGBM and P-GMM models. Although P-GMM and I-GMM employ the same principle, population-level GMM takes the extra step of tuning the weights of the Gaussian components based on individual user data. The weight-tuning step slows down the training of a P-GMM model, making it close to the training time of LGBM. The latter uses more time due to hyperparameter optimization. Small datasets are particularly susceptible to overfitting, and hyperparameter optimization can help mitigate this risk. Furthermore, as the look-back period increases, the training time for all models tends to increase as well, which is also expected, since more data is used for training. Once the model is trained on the available data, it can be used to make predictions about the behavior of the EV user using the user identification and the arrival time. This step, known as model deployment, involves applying the trained model to new, unseen data. Our results, presented in Table 6, demonstrate that the LGBM model requires significantly less execution time compared to the GMM-based models and SimS, with a difference of up to an order of magnitude and two orders of magnitude, respectively. LGBM models use decision trees to partition the feature space into regions, and the prediction is based on the majority response. Decision trees' inference can be calculated independently and largely parallelized. When deploying the GMM-based model, the probability density of observing a data point is calculated by summing over  $K$  components, representing the number of typical profiles in the data. This requires calculating the probability density for each component, which involves calculating the Mahalanobis distance between the data point and the mean of each component, as well as the determinant of the

covariance matrix. Evidently, this process is more computationally expensive. Despite taking longer to deploy, the SimS model is capable of achieving a reasonable prediction time of approximately 1 *ms* per prediction. The extra time is required for the model to calculate cosine similarity metrics and identify the sessions that are most similar to a given session in the testing data. Note that SimS does not require prior training, which makes it more robust to data drift. Considering data scaling, LGBM, I-GMM and P-GMM models show no clear correlation between deployment time and the look-back period. SimS deployment time is more aligned with the size of the data. However, it remains within a feasible range since the frequency of the predictions is limited by the frequency of the arriving vehicles, while the historical data used in similarity calculations are sparse time-series data.

Table 5: Training time [*s*] for varying model configurations and look-back periods

Look-back period	LGBM	P-GMM	I-GMM
30	23.734	17.598	<b>0.605</b>
60	25.660	19.867	<b>0.901</b>
120	25.475	41.270	<b>2.231</b>
240	34.857	55.285	<b>3.373</b>
360	32.255	66.570	<b>4.191</b>
480	50.035	64.673	<b>4.279</b>

Table 6: Deployment time [*ms*] per univariate prediction of an EV charging session parameter for varying model configurations and look-back periods

Look-back period	LGBM	IGMM	PGMM	SimS
30	<b>0.010</b>	0.179	0.185	0.889
60	<b>0.021</b>	0.178	0.180	0.802
120	<b>0.011</b>	0.161	0.170	0.834
240	<b>0.021</b>	0.162	0.159	0.937
360	<b>0.010</b>	0.180	0.164	1.009
480	<b>0.008</b>	0.152	0.157	1.031

## 6 Conclusions

The study aims to evaluate the performance of four models in forecasting the behavior of the EV user in terms of energy demand and parking duration. The models used in the study are LGBM, P-GMM, I-GMM, and SimS. The study investigates the influence of the look-back period on model performance and evaluates the impact of the user predictor variable. The results show that the models performed better than the user input requests, with the LGBM and SimS models outperforming the other models. The historical mean of the same user benchmark hit the lowest error when predicting the connection time of EVs, while the I-GMM and SimS-L models outperform the mean benchmark when predicting energy demands. The quality of the predictions produced with GMM-based algorithms varied between the variables of energy demand and parking duration. Although the error is less aligned with the size of the data, it is found that the error in the estimate of parking duration increased to a higher level with datasets exceeding 120 days in length. The study also evaluates the computational speed of training and the deployment of forecasting models. The results show that the LGBM model requires significantly less execution time compared to the GMM-based models and SimS.

Overall, the study demonstrates the effectiveness of data-driven models in forecasting the behavior of EV users. The proposed SimS method provides a promising solution to predict the parameters of a charging session. By avoiding the usual limitations of traditional clustering techniques, we offer a robust and efficient approach that is tailored to the characteristics of the charging session data. The method shows superior accuracy under various data availability conditions. We show that potential concerns about slower computational performance are valid. However, the method remains feasible in forecasting applications dealing with sparse time-series data, such as those related to EV charging sessions.

For further research, it is worth investigating the synergies between elements within the optimal scheduling framework of the electric vehicle charging process pictured in Figure 1b. The Similar Sessions has the capacity to represent uncertainty by generating scenarios drawn from similar sessions. Therefore, the advantages of stochastic optimization are subject to further study. Furthermore, the user arrival generation is a challenging problem in itself, as the subject time-series are irregular event-based, meaning it is binary, sparse, and non-equidistant. Specialized techniques are typically required to treat and model such data.

## Acknowledgements

This work has been supported by the BD4OPEM H2020 project, which has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 872525.

## References

- [1] Transport – Analysis - IEA, March 2023. URL <https://www.iea.org/reports/transport>. [Online; accessed 20. Mar. 2023].
- [2] Global EV Outlook 2022 – Analysis - IEA, March 2023. URL <https://www.iea.org/reports/global-ev-outlook-2022>. [Online; accessed 20. Mar. 2023].
- [3] S de la Torre, JA Aguado, and E Sauma. Optimal scheduling of ancillary services provided by an electric vehicle aggregator. *Energy*, 265:126147, 2023.
- [4] Zachary J Lee, Tongxin Li, and Steven H Low. Acn-data: Analysis and applications of an open ev charging dataset. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, pages 139–149, 2019.
- [5] Chris Develder, Nasrin Sadeghianpourhamami, Matthias Strobbe, and Nazir Refa. Quantifying flexibility in ev charging as dr potential: Analysis of two real-world data sets. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 600–605. IEEE, 2016.
- [6] M. Hadi Amini, Amin Kargarian, and Orkun Karabasoglu. Arima-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation. *Electric Power Systems Research*, 2016. doi: 10.1016/j.epr.2016.06.003.
- [7] Alexis Gerossier, Robin Girard, and George Kariniotakis. Modeling and forecasting electric vehicle consumption profiles. *Energies*, 12(7), 2019. ISSN 1996-1073. doi: 10.3390/en12071341. URL <https://www.mdpi.com/1996-1073/12/7/1341>.
- [8] Gilles Van Krieking, Cedric De Cauwer, Nikolaos Sapountzoglou, Thierry Coosemans, and Maarten Mesagie. Day-ahead forecast of electric vehicle charging demand with deep neural networks. *World Electric Vehicle Journal*, 12(4):178, 2021.
- [9] Yunsun Kim, Yunsun Kim, and Sahm Kim. Forecasting charging demand of electric vehicles using time-series models. *Energies*, 2021. doi: 10.3390/en14051487.
- [10] Sakib Shahriar, Abdul-Rahman Al-Ali, Ahmed H Osman, Salam Dhou, and Mais Nijim. Machine learning approaches for ev charging behavior: A review. *IEEE Access*, 8:168980–168993, 2020.
- [11] Tristan Rigaut, Alejandro Yousef, Maria Andreeva, and Vanya Ignatova. Scalable forecasting and model predictive control for electric vehicles smart charging. In *CIREP Porto Workshop 2022: E-mobility and power distribution systems*, volume 2022, pages 893–897. IET, 2022.
- [12] Julian Huber, David Dann, and Christof Weinhardt. Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging. *Applied energy*, 262:114525, 2020.
- [13] Yiqi Lu, Yongpan Li, Da Xie, Enwei Wei, Xianlu Bao, Huafeng Chen, and Xiancheng Zhong. The application of improved random forest algorithm on the prediction of electric vehicle charging load. *Energies*, 11(11): 3207, 2018.
- [14] Oliver Frendo, Nadine Gaertner, and Heiner Stuckenschmidt. Improving smart charging prioritization by predicting electric vehicle departure time. *IEEE Transactions on Intelligent Transportation Systems*, 22(10): 6646–6653, 2020.
- [15] Yukai Chen, Khaled Sidahmed Sidahmed Alamin, Daniele Jahier Pagliari, Sara Vinco, Enrico Macii, and Massimo Poncino. Electric vehicles plug-in duration forecasting using machine learning for battery optimization. *Energies*, 13(16):4208, 2020.
- [16] Casper Solheim Bojer and Jens Peder Meldgaard. Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting*, 37(2):587–603, 2021.
- [17] Spyros Makridakis, Evangelos Spiliotis, and Vassilios Assimakopoulos. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 2022.
- [18] Clayton Miller, Pandarasamy Arjunan, Anjukan Kathirgamanathan, Chun Fu, Jonathan Roth, June Young Park, Chris Balbach, Krishnan Gowri, Zoltan Nagy, Anthony D Fontanini, et al. The ashrae great energy predictor iii competition: Overview and results. *Science and Technology for the Built Environment*, 26(10): 1427–1447, 2020.

