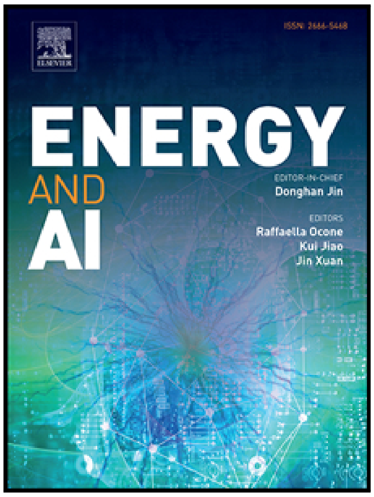
Journal Pre-proof



Data-driven smart charging for heterogeneous electric vehicle fleets

Oliver Frendo, Jérôme Graf, Nadine Gaertner,
Heiner Stuckenschmidt

PII: S2666-5468(20)30007-0
DOI: https://doi.org/10.1016/j.egyai.2020.100007
Reference: EGYAI 100007

To appear in: *Energy and AI*

Received date: 24 March 2020
Revised date: 30 April 2020
Accepted date: 30 April 2020

Please cite this article as: Oliver Frendo, Jérôme Graf, Nadine Gaertner, Heiner Stuckenschmidt, Data-driven smart charging for heterogeneous electric vehicle fleets, *Energy and AI* (2020), doi: https://doi.org/10.1016/j.egyai.2020.100007

Data-driven smart charging for heterogeneous electric vehicle fleets

Oliver Frendo[a,*], Jérôme Graf[a], Nadine Gaertner[a], Heiner Stuckenschmidt[b]

[a] *SAP SE, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany*
[b] *University of Mannheim, Schloss, 68131 Mannheim*

Abstract

The ongoing electrification of mobility comes with the challenge of charging electric vehicles (EVs) sufficiently while charging infrastructure capacities are limited. Smart charging algorithms produce charge plans for individual EVs and aim to assign charging capacities fairly and efficiently between vehicles in a fleet. In practice, EV charging processes follow nonlinear charge profiles such as constant-current, constant-voltage (CCCV). Smart charging must consider charge profiles in order to avoid gaps between charge plans and actual EV power consumption. Generally valid models of charge profiles and their parameters for a diverse set of EVs are not publicly available. In this work we propose a data-driven approach for integrating a machine learning model to predict arbitrary charge profiles into a smart charging algorithm. We train machine learning models with a dataset consisting of charging processes from the workplace gathered in 2016-2018 from a heterogeneous EV fleet of 1001 EVs with 18 unique models. Each charging process includes the time series of charging power. After preprocessing, the dataset contains 10.595 charging processes leading to 1.2 million data points in total. We then compare different machine learning models for charge profile predictions finding that XGBoost yields the most accurate predictions with a mean absolute error (MAE) of 126W and a

⋆This work was supported in part by the German Federal Ministry for Economic Affairs and Energy during the TRADE EVs project.

*Corresponding author
Email address: ofrendo@gmail.com (Oliver Frendo)

relative MAE of 0.06. Simulations show that smart charging with the integrated XGBoost model leads to a more effective infrastructure usage with up to 21% more energy charged compared to smart charging without considering charge profiles. Furthermore, an ablation study on regression model features shows the EV's model is not a necessary attribute for accurate charge profile predictions. However, charging features are required including the number of phases used for charging.

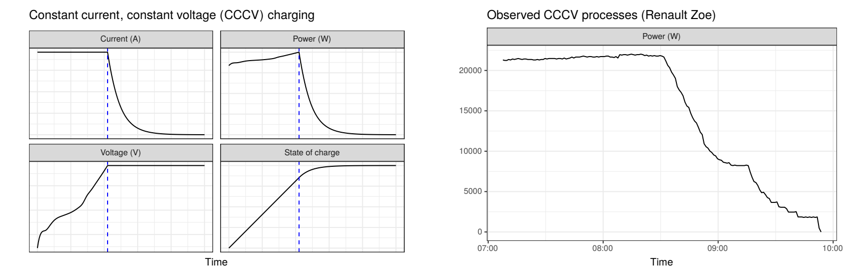
Keywords: Smart charging, electric vehicles, data-driven approach, machine learning

1. Introduction

Mobility is in a phase of transition towards electrification. With electric vehicles (EVs) turning from a niche into a mainstream choice the operation and charging of EVs turns into a widespread challenge. Infrastructure charging capacities are limited by the capacity of grid connection lines. In limited charging infrastructures, charging needs to be organized intelligently to be able to charge fleets of EVs. Charging infrastructures are run by charge point operators (CPOs) who are responsible for implementing smart charging. This work views smart charging from the perspective of the CPO.

A common example for fleet charging is the scenario of employee charging at the workplace [1]. Workplace charging is characterized by an accumulation of many EVs in one location, long parking durations and comparatively low charging needs due to limited commute distances. In this work we consider heterogeneous EV fleets which consist of different EV models. The challenge in workplace charging is to assign the available charging capacity fairly between EVs. A real-time application of smart charging assigns charging resources by continuously prioritizing between EVs. For this purpose, smart charging relies on information about each EV, including its state of charge (SoC) and the remaining parking time.

Existing approaches to smart charging [2, 3, 4] have achieved real-time charge



(a) Simulated CCCV process consisting of two distinct stages

(b) Observed CCCV process from the perspective of the charge point operator (CPO)

Figure 1: CCCV: Simulation vs observed process

scheduling with increasing efficiency in the assignment of scarce charging resources. However, in practice the execution of the charge schedules deviates from the theoretically planned schedules when the complex charging behavior of the EVs themselves are not reflected in the scheduling decisions [5]. Battery management systems (BMS) in EVs control the charging behaviour of the battery [6]. In particular, BMS limit the power drawn during charging to protect battery health and safety. Consequently, reserved infrastructure charging capacity is wasted when batteries draw less power than planned by the central smart charging system.

Battery behavior during charging is determined via the BMS' implementation of *charging profiles*. Charging profiles express charging currents drawn over time. Common charging profiles include constant-current, constant-voltage (CCCV) [7]. Charging profiles represent patterns that can be observed in every charging process. Figure 1 shows a simulation of a CCCV process compared to a real process.

One approach to modelling battery behavior are equivalent circuit models (ECMs) [8]. ECMs are a standard modelling technique in electrical engineering which can be used to approximate battery behavior [9]. ECMs conceptually enable close approximations of the battery behavior. However, detailed parameters of an ECM as well as the battery pack composition generally equate to trade secrets of the manufacturers and are cumbersome to determine via experimental battery measurements and reverse engineering. Modelling battery behavior with factors such as temperature and battery state of health (SoH) requires more complex ECMs and more data.

In this work, we use an alternative approach to modelling battery behavior. Instead of approximating battery behavior by adding more and more refining elements in an ECM we infer charge profiles from a machine learning model that has been trained on real historical data.

The ECM model requires battery parameters (resistance and capacitance) and battery inputs and outputs (voltage and current). In contrast, the model learned via machine learning only requires the car model and its state of charge. We compare different machine learning models, namely linear regression, neural networks and XGBoost. The learned battery behavior is directly incorporated in smart charging.

We thus propose an integrated approach which produces charge plans for a heterogeneous set of EVs while considering their battery behavior. A charge plan is a time series of discrete timeslots and a prescribed charging power per timeslot. To summarize, we propose a practical, data-driven approach to smart charging reflecting not only the fair assignment of scarce charging resources but also incorporating the effects of battery behavior.

We address the following research question: How do integrated predictions of battery charge profiles affect smart charging? Our main contributions are:

- A methodology for preprocessing a dataset and training a regression model to predict battery charge profiles,
- the integration of the regression model into a smart charging algorithm and
- a quantification of the impact of integrated charge profile predictions on smart charging in simulations of historical data

The remainder of the paper is organized as follows. Section 2 discusses the related work of smart charging and prediction in the context of electric mobility.

Section 3 introduces the data-driven methods, including the methods for data preparation, the machine learning models and the embedding smart charging heuristic. Sections 4 and 5 describe the experimental setup and the experimental results for smart charging with an integrated prediction mechanism for battery behavior. Finally, section 6 discusses the applicability and limitations of the results while section 7 presents the conclusions of this work.

2. Related work

Smart charging. There is a large body of related work on smart charging for EV fleets. We refer to recent surveys such as [2, 3, 4] for an overview. Common goals include driver satisfaction [10, 11, 12, 13], energy cost minimization [14, 15, 16] and mitigating battery aging [17]. In this work, we focus on driver satisfaction which in our context corresponds to maximizing the average SoC across all EVs. Furthermore, we consider a real-time smart charging approach.

Battery models in smart charging. Smart charging approaches implicitly (or explicitly) contain a model of EV batteries and their charging profiles. The simplest model and charging profile is represented by the assumption that the EV can always charge at maximum power independent of SoC [12, 17, 18]. Traditional approaches to modelling batteries include ECMs [6, 19, 20] and physical models [9]. Physical models are more accurate compared to ECMs but are also more expensive computationally and thus unsuitable to apply in real-time [8]. Additionally, machine learning (ML) models present a recent development to predicting battery behavior [21, 22]. We refer to a recent survey [8] for an overview of the state of the art for predicting battery behaviour based on internal EV battery data such as voltage, current and battery pack configuration.

Smart charging approaches using ECMs for single battery cells or packs have been proposed to maximize battery life [23, 7, 24]. Regarding smart charging for EV fleets, examples of related work with integrated ECMs [25, 26] use model predictive control (MPC) to schedule EV charging while minimizing energy costs.

However, the parameterization for ECMs, physical models and ML models is limited in practice by data availability on operational conditions such as temperature and state of health. Additionally, the EV's BMS controls the selection of the charging profile [7]. From the perspective of the charge point operator such battery models are even more difficult to apply for heterogeneous fleets. Different EVs each require one set of model parameters. In this work we thus consider an approach without knowledge of detailed battery model parameters.

Predicting EV charging demand. Predicting the charging demand of a fleet of EVs without detailed battery knowledge is a well studied problem. For example, a fleet of EVs is simulated based on three different static charging profiles in [27]. EVs are assumed to arrive according to a probability distribution of arrival and departure times. An artificial neural network (ANN) is used in [28] to predict daily arrival time and travel distance to allow forecasting load. Data on traffic patterns is used to forecast power demand of a fleet of EVs [29]. Lastly, [30] uses a support vector machine (SVM) to predict building power load while taking into account an EV fleet.

More granular approaches predict the charging demand of individual EVs. GPS data from 76 (conventional) vehicles is used in [31] to simulate PHEV charging processes. A constant battery size (24kWh) and charging rate (3.7kW) is assumed. Similarly, an ANN is used in [32] to predict individual charging profiles. The ANN is trained on data generated by a stochastic random process. In this work we use regression models trained on real data to predict individual charging profiles while taking into account EV specific characteristics.

In contrast to considering charge profiles, there is also work on predicting individual EV energy consumption during driving [33, 34]. However, in this work we focus on EV charging profiles.

An obvious approach to predicting EV charge profiles would be to use traditional time series approaches for forecasting. Traditional time series models for forecasting univariate time series include *Auto Regressive Moving Average* (ARMA) and *Auto Regressive Integrated Moving Average* (ARIMA) [35]. For example, [36] uses an ARIMA model for forecasting the load demand of a parking lot. A fleet of heterogeneous EVs is considered with a discrete set of constant charging rates. Forecasts are then used in day-ahead scheduling. Furthermore, [37] uses an ARMA model to predict PV generation. The ARMA model is combined with an EV scheduling approach. Lastly, charging process time series from 20 real charging stations are divided into fragments and fragments are clustered using the Euclidean distance measure in [38]. Future energy consumption for a given fragment is predicted using the most similar existing fragment via k-nearest neighbours.

In this work, time series forecasting would be of interest for predicting individual EV charge profiles. However, the integration of time series forecasting with real-time smart charging is conceptually problematic because charge plans influence the course of each time series. Instead of time series forecasting we predict individual EV power in relation to SoC.

Data-driven approaches. Previous work on data-driven approaches in the context of electric mobility typically makes use of private datasets consisting of aggregated information on charging processes. For example, the statistics of 400.000 charging processes in the Netherlands are described in [39]. A dataset consisting of 21.918 charging processes in 2012-2013 from 255 different charging stations in the UK is analyzed in [40]. The dataset is combined with weather data to characterize the load demand of a fleet of EVs. Similarly, 8.929 processes from 2014-2016 are analyzed in [41]. Compared to other work in this category, the timestamp of reaching full SoC is included. The popularity of EV charging infrastructure is predicted in [42] based on geographic information system (GIS) data. A large dataset of 500.000 charging processes and 2.000 charging stations in the USA from 2013 is used to estimate the benefits of smart charging in [43]. Charging profiles are considered as static profiles depending on temperature in [34]. The profiles are computed based on 8.300 EVs in 2011-2013. Lastly, [44] uses GPS and trip meter data of 490 PEV taxis in 2013 to simulate charging profiles. In [44], a single EV model (75kWh) is modelled to charge using constant-power at 22kW up to 80% SoC even though a two-stage charging profile is mentioned.

In this work, in contrast to the references above we use a historical dataset which includes charging process data over time. The data-driven approaches above each use a dataset which includes only aggregated attributes such as energy (Wh), average power (W) or maximum power (W) and lack the accompanying time-series of power over time. Due to the more granular dataset used in this work we are able to take into account arbitrary charging profiles.

Public datasets. Public historical datasets in the domain of electric mobility lack the required granularity and diversity required for training machine learning models to predict charging profiles. Basic datasets include charging station locations for use in navigation such as [45].

Charging process data is available in datasets such as [46] and typically includes information such as the total energy charged (Wh) and the duration of the process. Time series data (power over time) is not included.

More granular charging process data includes the accompanying time series. For example, [47] provides a dataset which contains simulated time series of charging currents for homogeneous EVs based on household energy consumption from 2009. A large dataset containing 4 million charging processes from 8.300 EVs in 2011-2013 is analyzed in [34, 48]. The data includes 6 different car models such as the Nissan Leaf and the Toyota Prius. However, the dataset includes a limited number of sample charging process time series (19) without SoC and is therefore not suitable for training machine learning models. A more comprehensive dataset is discussed in [1]. The data is a result of applying a smart charging approach [5] in a real charging infrastructure. The dataset contains over 30.000 charging processes and time series with charge schedules and the resulting charging current. However, the dataset does not include SoC over time and the EV's model.

In this work, we make use of a charging process dataset containing EV models and the accompanying time series data.

Literature gap. Related work on smart charging frequently positions integrating nonlinear charge profiles as future work [2, 5, 49]. For example, an analysis of a practical application of smart charging [5] describes the difference between applied schedules and measured EV power and proposes reducing the difference by integrating nonlinear charge profiles. There is thus a lack of work on integrating nonlinear charge profiles with real-time smart charging.

To address the literature gap, we propose taking into account nonlinear charge profiles by integrating charge profile predictions with smart charging. We use a data-driven approach with machine learning models trained on historical time series data of EV charging processes. We propose an approach based on data gathered by the charge point operator thus ensuring the applicability of the approach in practice.

3. Method

This section introduces the methods and techniques applied for data driven smart charging. Data-driven smart charging is based on two pillars as visualized in figure 2: model preparation and model application.

Model preparation is a one-off activity for training a regression model which is capable of predicting charging power based on SoC. The process of training the regression model begins with the collection of historical charging process data. In this work, we use a historical dataset [50] with charge process data as described in table 1. The data preparation methods for the raw dataset include data cleansing and one-hot encoding and are described in detail in subsection 3.1. Subsection 3.2 explains how the prepared dataset is used to train the regression model.

Once trained, the regression model is ready to be applied for real-time smart charging. During daily operations the charge point operator applies smart charging and repeats the process of scheduling a charge plan for each EV when it connects to the charging infrastructure. The smart charging algorithm is described in detail in subsection 3.3. The role of the regression model in the algorithm is to predict the charging power throughout the charging process and thus contribute to the prioritization among EVs.

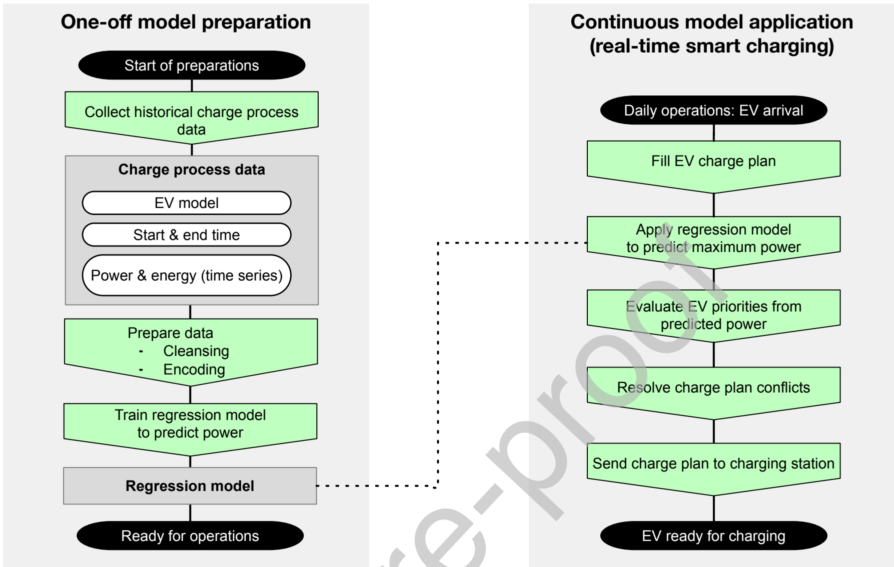


Figure 2: Overview of proposed methods

3.1. Data preparation

This section discusses the methodology we use to clean and preprocess a dataset of meter values collected from charging stations in order to train regression models. We use the term *charging process* to refer to the complete session of an EV charging. We use the term *meter value* to refer to a single data point of a time series. For example, a measured power of 10kW at a charging station at 10:18 is one meter value. Each charging process is associated with one time series of power over time. In contrast to related work such as [32], we do not simulate this time series but work with real data. We use the R programming language for dataset preparation.

Dataset structure. To begin with, we gather EV model data published by manufacturers (table 5). We then combine the manufacturer's EV model data with the charging process data and the meter value data. Meter value data is gathered with a granularity of 1 minute. Table 1 shows the raw attributes of

Table 1: Dataset raw attributes

Attribute	**Sample value**
EV data (18 unique models)	
Battery capacity	41.000Wh
Maximum power	22080W
Three phase charging	True
Charging process data (10.595 processes)	
Arrival time	08:12:37
End time	17:48:01
Car model	Renault Zoe (2018)
Total power consumption	26120Wh
Meter value data (Granularity: 1 minute, 5.2 million data points)	
Timestamp	09:01:37
Power	22.080W

the dataset.

Dataset characteristics. The dataset consists of 10.595 uncontrolled charging processes gathered from employee charging at the workplace. The processes stem from the years 2016-2018 and from 1001 EVs charging at 338 charging stations in 8 different cities. The EV fleet is heterogeneous and is composed of 18 unique EV models, each of which has different values regarding maximum charging rate, battery capacity, three phase charging (see table 5). BEVs typically charge on three phases while PHEVs usually charge on a single phase. The mean energy charged per charging processes is 7.01kWh while the mean duration is 7 hours and 17 minutes.

Dataset preprocessing. Next, we describe the steps we perform for cleaning and preprocessing the dataset. To begin with, we estimate SoC assuming a charging efficiency η of 0.85. Based on studies [51, 52, 53] on EV charging, this is an optimistic value. EV charging efficiencies computed from experimental data include values of 85% [51], between 60% and 85% [52] or between 64%

and 88% [53]. In practice, charging efficiency is not constant but a function of attributes of EV charging such as heat, SoC and power. For the sake of simplicity, we assume a constant value.

We use equation (1) to estimate the SoC $z_{n,t}$ for EV n at t. We use the time series X of a charging process to compute the SoC as the difference between the EV's capacity Q_n (Wh) and missing relative capacity based on the previously charged energy X_t (Wh) taking into account efficiency η.

$$z_{n,t} = Q_n - \eta * \frac{X_{end} - X_t}{Q_n} \tag{1}$$

Data cleansing. First, we remove processes without an associated car model. Such processes can be a result of backend system tests. Next, we remove processes with negative SoC values. Negative SoC values are generated by equation (1) in processes where our efficiency estimate (85%) is too optimistic. We also limit processes to those with a maximum length of 24 hours. We assume longer stays contain irregular charging behaviour not in the scope of charging at the workplace. Similarly, we remove processes shorter than 10 minutes.

We remove charging processes with missing data due to charging station connectivity problems. Furthermore, we remove charging processes with a flat charging profile. We assume processes with a flat charging profile were stopped before reaching full SoC and thus have no second charging profile stage. We identify such processes by analyzing whether the values of the time series are all within 95% of the maximum value of the time series.

The remaining cleaned dataset consists of 10.595 charging processes.

Regarding the time series, we remove meter values after power reaches 0W for the first time. Artifacts such as remote air conditioning sometimes results in a power draw greater zero just before the end of the process. This leaves us with 1.2 million data points (from 5.2 million originally).

Feature engineering. Table 2 shows the preprocessed attributes of the dataset. We use the standard approach of one-hot encoding categorical features. In one-hot encoding each value of a categorical feature is converted to its own Boolean feature and the original categorical feature is removed. We one-hot encode the categorical variable EV model leading to 18 individual Boolean features. In regression models requiring only numerical features we use values of 0 and 1 for the Boolean features.

Table 2: Sample data point in cleaned and preprocessed dataset (1.2 million data points) used for training regression models. The car model is included as a one-hot encoded feature leading to 18 individual Boolean features.

Feature name	**Sample value**
Input features (23 features)	
Is model BMW i3?	false
Is model Renault Zoe?	true
...	false
Is model Mercedes Benz C 350e?	false
Three phase charging?	true
Is BEV?	true
Is PHEV?	false
State of charge	0.3892
Arrival time (seconds after midnight)	29557
Target feature	
Power (W)	21358

3.2. Learning regression models

Training regression models is a proven data-driven methodology for estimating one output feature given a set of input features. In the context of battery-aware smart charging the charging power is the feature to be predicted via a regression model. This section discusses how we train different regression models and which baseline we use to predict charging power given the EV model, SoC and arrival time from the preprocessed dataset described in table 2. We train a linear regression model, an XGBoost regression model as well as a neural network.

Constant. We use a constant predictor as a baseline with the simplest possible predictions and to represent prior work [14]. For the constant predictor, we use the maximum power draw per EV model according to manufacturers (see table 5).

Linear regression. Linear models represent a common benchmark for regression models. We fit the model in equation (2) with predicted power y, 23 features x_i and error ϵ_j per data point j.

$$y = \beta_0 + \sum_{i=1}^{m} \beta_i * x_i + \epsilon_j \tag{2}$$

We also fit a second linear regression model on data where the target feature y has been logarithmically transformed to maximize prediction accuracy. The motivation for logarithmically transforming power y is that the power in the second stage of charging profiles is often modelled as an exponential decrease [49, 54]. In section 5 both linear regression models are evaluated independently.

Neural networks. Recent related work on predicting or modelling battery behaviour with neural networks includes [55, 56]. Both approaches use neural networks to predict SoC taking into account battery voltage, current and temperature.

In this work we use Keras [57] for training neural networks. To begin with, we use the commonly applied standardization approach where each input feature has a mean value of 0 and a standard deviation of 1. The neural network consists of two hidden layers, each of which has 128 neurons with rectified linear activation functions. Weights are regularized using L1 regularization. Because we use the neural network for regression on a single target feature (power) the final layer is a single neuron. An 80-20 train and test split was used for evaluation. Additionally introducing dropout layers, a higher number of hidden layers or a different number of neurons did not result in an improved MAE.

XGBoost. Gradient boosting machines are a popular regression model for structured data. XGBoost (eXtreme Gradient Boosting) in particular has become widespread [58]. In the domain of battery technology, work such as [21] uses gradient boosted trees to predict the remaining useful life (RUL) of batteries. Similarly, [22] uses XGBoost to predict SoC. Both approaches require internal variables such as battery voltage and current. However, charge point operators do not have access to internal battery variables.

We perform hyperparameter tuning via grid search using the `mlr` package in R. To avoid overfitting, hyperparameters are tuned using 10-fold cross validation. For reference, the best parameters based on the grid search are `booster=gbtree`, `nrounds=100`, `eta=0.1`, `max_depth=9`, `min_child_weight=1`, `gamma=0`, `colsample_bytree=1` and `objective=reg:linear`.

3.3. Integration with smart charging algorithm

In this work we improve on the real-time smart charging algorithm introduced in prior work [14]. The algorithm is a heuristic method which creates individual charge plans for a set of vehicles. It represents a practically oriented method with goals similar to the method proposed in [5].

The algorithm solves the decision problem of assigning charging power to EVs over time. Charging capacities are limited by fuses in the installed charging infrastructure and ultimately by the connection to the power grid. The goal of the algorithm is to maximize the average SoC over all vehicles while respecting charging capacities of the infrastructure. The output of the algorithm is one charge plan per vehicle. The charge plan is a time-series discretized into 15 minute timeslots and specifies the charging power for each time slot. From the perspective of the CPO, a charge plan specifies desired values for charging power. On a technical level, the charging station implements the charge plan as an upper bound to EV power draw. The actual power draw is controlled by the EV's BMS.

The flow of the algorithm is depicted in figure 3 and involves the following scheduling and prioritization steps. Each time a vehicle begins a charging process the smart charging algorithm is triggered and computes a charge plan for the newly connected vehicle. The scheduling step iterates over all timeslots and includes a timeslot in the vehicle's charge plan as long as the vehicle has a charging need. The initial charge plan is optimistic in the sense that it uses the earliest available timeslots for each vehicle. The prioritization step deals with a conflict resolution when capacity allocations reach the limits of the infrastructure. Allocations are reassigned between vehicles and between timeslots. When reassigning charging capacities, vehicles with higher charging urgency are preferred. The charging urgency is associated with the charging priority and is related to the remaining parking time Δt and the state of charge $z(t)$ as detailed in equation 3 and initially introduced in prior work [14]. Each vehicle's departure time $t_{\text{departure}}$ is assumed to be known in advance. Vehicles are first prioritized by whether and how much they are charged above their minimum required SoC z_{min} at time t by using a large artificial constant M (e.g., $M = 10^5$). Vehicles' charged energy is normalized by the maximum current $I(t)$ they can draw.

$$priority(t) = \begin{cases} \frac{Q*(z_{min}-z(t))}{|\Delta t|*I(t)+\epsilon}, & \text{if } z(t) < z_{min} \\ \frac{Q*(1-z(t))}{|\Delta t|*I(t)+\epsilon} - M, & \text{otherwise} \end{cases} \tag{3}$$

$$\text{where } \Delta t = t_{\text{departure}} - t \tag{4}$$

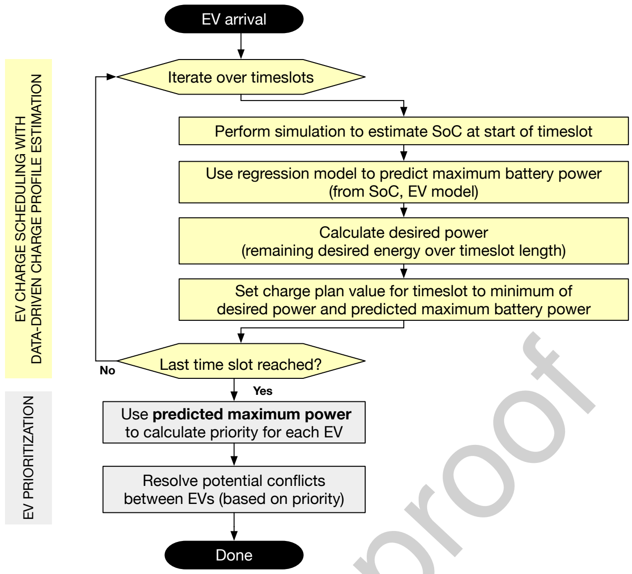


Figure 3: Smart charging algorithm integrating the data-driven charge profile simulation

This work enhances the algorithm so that it takes into account charge profiles such as CCCV when computing charge plans. Prior work [14] creates charge plans assuming each EV is able to draw its maximum power irrespective of SoC and charge profile. Figure 4 depicts a sample CCCV process without a charge plan, with a charge plan as per previous work [14] and with a charge plan from this work. In the constant current phase of the charging profile the vehicle follows the charging plan directly. In the constant voltage phase, the current drawn follows a steadily declining curve which is only affected by the charge plan if the charge plan current is below the CV current. When the vehicle draws less power than is reserved by the charge plan then charging capacity is wasted from allocation to other vehicles.

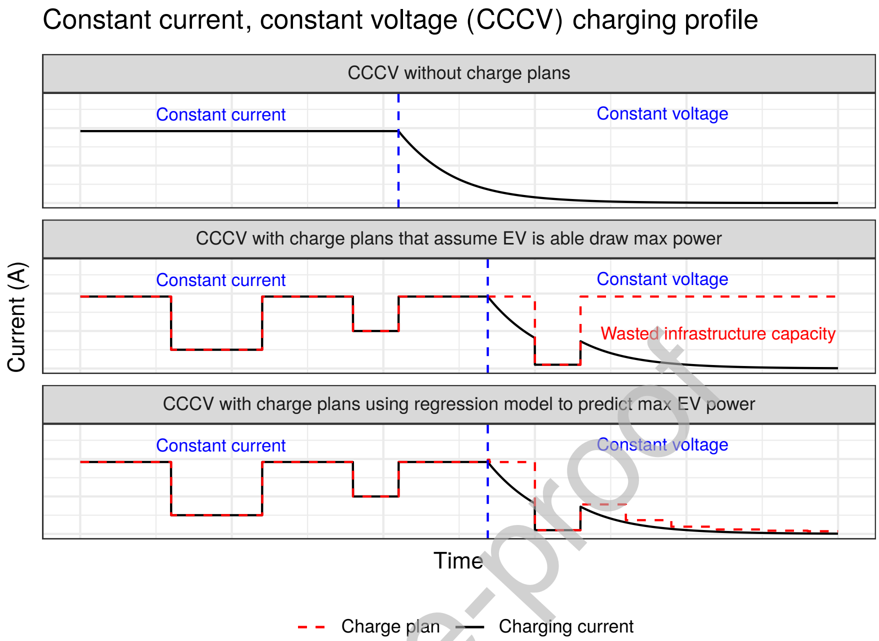


Figure 4: Constant current, constant voltage (CCCV) charging with and without charge plans. The charge plan specifies an upper bound for the EV's charging current.

The improvement proposed in this work consists of reducing wasted capacity by considering the anticipated power drawn during the creation of charge plans. The approach is to introduce a data-driven charge profile simulation which integrates a regression model as introduced in section 3.2. The regression model is used to predict the maximum power drawn by the vehicle at any point in time based on features including the state of charge. The smart charging algorithm uses the estimated maximum power to reserve a more realistic charging capacity. Consequently, the resulting charge plan reflects the actual charge profile more closely as visualized in the last diagram within figure 4. The strength of integrating a regression model is that it allows to reflect arbitrary charge profiles and is not restricted to CCCV charging only.

Regarding performance requirements of the algorithm: In a real life setting, vehicles connect and disconnect from the charging infrastructure frequently. In

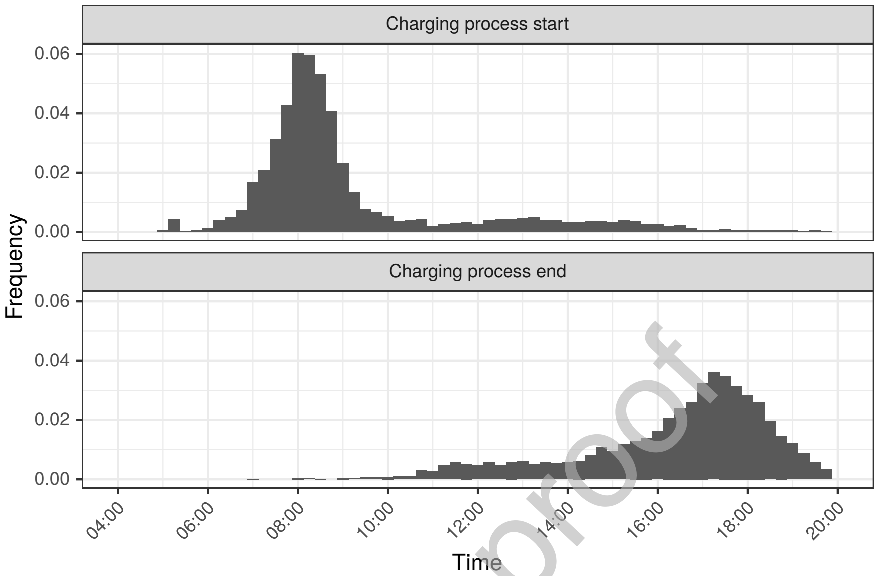


Figure 5: EV arrival and departure time distribution

the scenario of workplace charging the frequency distribution of arrivals and departures is characterized by a peak of arrivals between 8 and 9 am and a peak of departures after 5 pm. In the dataset [50] used in this work the median arrival is at 8:22 am while the median departure is at 5:01 pm. Figure 5 shows the distributions for EV arrival and departure times.

During peak arrival several new charging processes are started each minute. The fact that every connection triggers a computation of charge plans and potentially a reallocation of charging capacities motivates the need for a real-time capable algorithm. The smart charging algorithm is designed to perform with sub-second response times for charging infrastructures with 300 charging stations while achieving comparable smart charging quality compared to an approach involving a mixed integer linear programming model [14].

4. Experimental setup

We use a discrete event simulation implemented in Java. The simulation models EVs and a three phase charging infrastructure. We trigger reoptimization on new events. Events include new EV arrivals, EVs reaching 100% SoC, EV departures and changes to the charging infrastructure. Reoptimizing includes recomputing charge plans via the algorithm discussed in 3.3

For predicting individual EV charge profiles we use a pure Java implementation[1] of XGBoost. XGBoost models are trained with the programming language R as discussed in section 3.2. For each simulation run we select a random historical charging process per EV and simulate the charging profile via a lookup table consisting of SoC and power draw. We use a seed when selecting the random historical charging processes to ensure reproducible results. A seed is used in pseudo-random number generators as a base value. The generator will always generate the same sequence of numbers with a given seed. In other words, the results of a simulation are reproducible with a given seed because the same historical charging processes will be selected.

Table 3 shows the simulation parameters we use for simulations. The number of cars is variable and is increased for each simulation to show the impact of the infrastructure bottleneck (30kW). With a smart charging algorithm this infrastructure should be sufficient to fully charge roughly 40 EVs considering the main business hours of 08:00-17:00. The average energy consumption per session is 7 kWh and the infrastructure could be used for 270kWh spread out over 9 hours.

Each charging station is rated for a three phase BEV (22kW). We set the number of charging stations to allow for every EV to plug in. We use a constant charging efficiency of 0.85 as discussed in section 3.1. Finally, table 5 shows the car model data published by manufacturers that we use.

[1] https://github.com/komiya-atsushi/xgboost-predictor-java version 0.3.1

Table 3: Simulation parameters: 600 simulations total

Parameter	**Values**
Number of cars	{5, 10, ..., 100}
Number of charging stations	100
Charging station power rating	22kW
Charging efficiency	0.85
Power prediction method	{Constant, XGBoost, ...}
RNG seed	{0, 1, ..., 9}

5. Experimental results

This section discusses results from the simulations in section 4. First, in section 5.1 we analyze the performance of the different regression models by comparing error metrics. Based on the error metrics we discuss the suitability of the regression models for integration in the smart charging algorithm. We then show six examples of charging profiles and the accompanying predictions. Finally, we analyze feature importance in the best performing model (XGBoost).

In section 5.2 we directly measure the effect of the integrated regression models on smart charging by comparing charging output measured as mean final SoC. We evaluate the relevance of individual features within the model and present an ablation study to further quantify possible adverse effect when omitting features from the model. Lastly, we analyze two simulations in more detail and the influence of integrating regression models on the computation time of smart charging.

5.1. Regression models

Table 4: Predicting charging profile power: Error metrics per prediction method

Method	MAE	RelMAE
Baseline (predict maximum power)	2077.63	1.00
Linear regression	656.13	0.32
Linear regression (on log-power)	765.97	0.37
Neural Network	151.28	0.07
XGBoost (ablation study)	145.96	0.07
XGBoost	126.21	0.06

Comparison of regression models. In this work, we use the metrics mean absolute error (MAE) and the relative MAE (RelMAE) to quantify regression model results. The relative MAE is computed by dividing the MAE of a method with the MAE of a baseline [59]. We use the EV's maximum power as the baseline.

Table 4 shows the MAE and RelMAE per regression model. To begin with, the baseline which assumes the EV will always draw its maximum power has the highest MAE (2077.63) thus motivating the use of regression models. Next, the standard linear regression model as well as the model fitted to logarithmically transformed power both perform poorly with an MAE of 656.13 and 765.97 respectively. The high MAE of both linear regression models can be explained by the intrinsic nonlinear relation between SoC and power [2].

More sophisticated regression models which are able to deal with nonlinear relations between attributes show better results. Neural networks and XGBoost show a comparable MAE of 151.28 and 126.21. However, XGBoost shows slightly better performance. As described in section 3.2 we report the average MAE on each test set during 10-fold cross validation. We use the best performing regression model (XGBoost) for the simulations of EV fleets in section 5.2.

As discussed in section 3.1 in this work we assume an efficiency value of $\eta = 0.85$. A sensitivity study with varying efficiency values showed no significant difference in prediction accuracy.

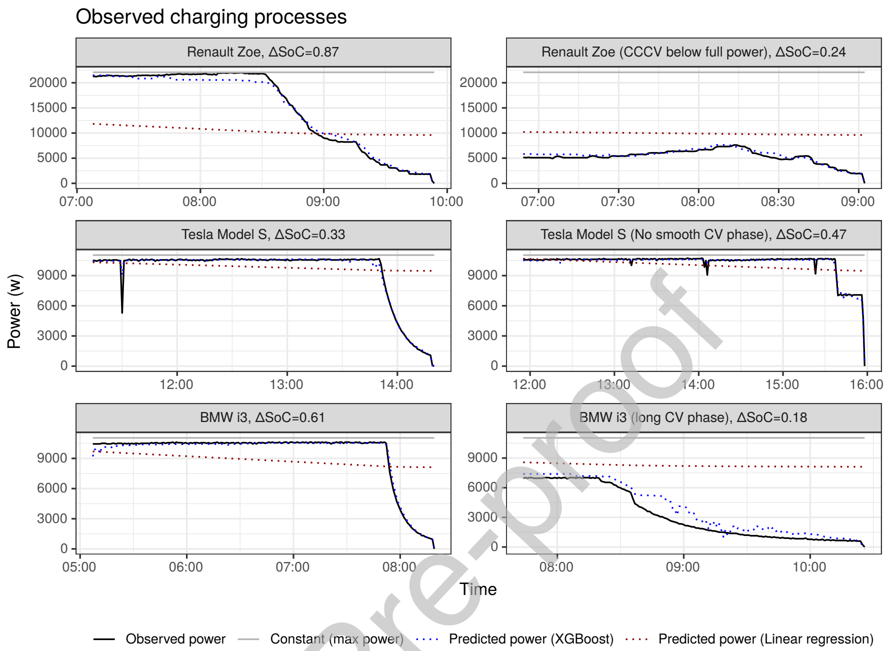


Figure 6: Observed CCCV processes and predicted power using different regression models

Charge profile prediction examples. In the following we discuss six examples of charging processes to visualize the inherent nonlinearity in charging profiles as well as regression model prediction accuracy. Figure 6 shows the six processes involving three popular EV models from the fleet. To show charging profile diversity in the dataset we present three processes with a high prediction accuracy (left column) and three processes with a low prediction accuracy (right column).

Interestingly, when observing the first stage in the charge profile of each process, the Renault Zoe appears to implement CCCV while the Tesla Model S and the BMW i3 appear to implement constant-power, constant-voltage (CPCV). We base this observation on the steady increase of power in the first stage of CCCV. In comparison, the CPCV processes show a first stage with recognizably constant power.

For each of the six charging processes we show the charge profile predictions for two regression models. XGBoost shows significantly better prediction accuracy compared to a linear regression model which corresponds to the lower MAE in table 4 (126.21 vs 656.13). The diversity in such sample charging processes underscores the need for data-driven approaches to take into account processes that do not follow the expected theoretical charge profiles.

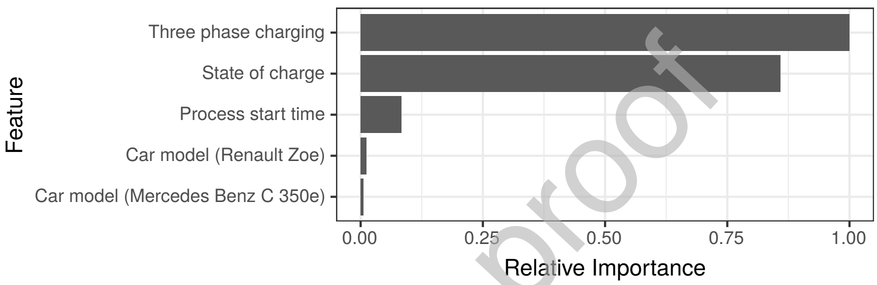


Figure 7: XGBoost feature importance

Feature importance. Analyzing the feature importance in trained regression models presents another approach to evaluation of regression results. For tree-based models such as XGBoost, feature importance expresses the impact on the regression when splitting on the feature. In a single decision tree, the *gain* for a feature expresses how well splitting on that feature improves results. An XGBoost model consists of multiple trees and the gain is averaged over all trees. We refer to [58] for a more in-depth explanation of feature importance in XGBoost models.

Figure 7 shows the feature importance for the trained model. The most important feature is whether the EV is able to use three phases followed by the state of charge and the charging process' start time. The relative importance of *three phase charging* is explained by the fact that vehicles charging on three phases draw significantly more power than vehicles charging on only one phase, namely roughly by factor three. The fact that *state of charge* ranks high in feature importance emphasizes that the power drawn changes significantly with the SoC. In particular, the power draw decreases significantly towards high SoC.

5.2. Impact of integrated regression models on smart charging

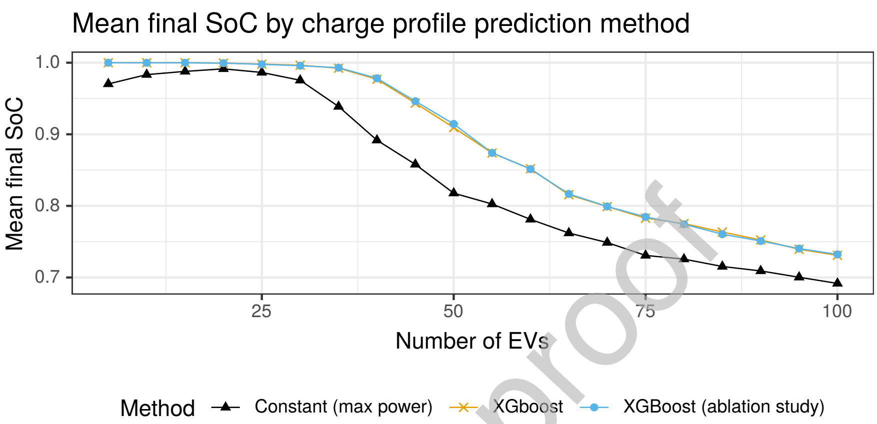


Figure 8: Experimental results: Mean final SoC for different EV fleet sizes

Scheduling quality. Figure 8 shows the mean final SoC for different EV fleet sizes. The mean final SoC is the average SoC after charging computed over all EVs in the fleet. Each point in the plot is the average result of 10 simulations (600 simulations total). All methods show a decline after 35 EVs in the mean final SoC because of the infrastructure bottleneck (30kW). However, each method shows a different mean final SoC because the infrastructure is used more or less effectively. For example, the difference in SoC with 40 EVs is large. Assuming a constant battery model leads to a mean final SoC of 87%. In comparison, a mean final SoC of 96% is reached when using XGBoost as a charge profile predictor.

Ablation study. An ablation study can be used to determine the influence of certain features on ML models by retraining the models without said features. In the following we discuss whether the knowledge of the EV's actual *car model* is necessary or whether it is sufficient to know the EV's type (BEV/PHEV) and whether the EV is able to charge on three phases. The EV's model may not be available to the charge point operator, for example, in scenarios where users do not authenticate.

In this work, we perform an ablation study assuming charging features (car type, three phase charging) are available via organisational measures. For example, parking spaces may be reserved specifically for BEVs or PHEVs and charging stations may be connected on one or three phases. The regression results in table 4 show a slightly higher MAE (145.96 vs 126.21) when training the XGBoost model without the one-hot encoded car model features.

We thus conclude the car model itself is not needed. However, the EV's charging features (BEV/PHEV, three phase charging) are needed as per the following reasoning.

Figure 9 shows power in relation to SoC and contains all data points from the dataset. There are many different charging profiles and the typical charging levels are easily recognizable as horizontal lines (single phase 3.7kW, three phase 11kW or 22kW). Intuitively, without the car type and the three phase charging characteristic there are multiple values on the y-axis per single value on the x-axis. It is thus infeasible to find a function to accurately predict power based solely on SoC which confirms the need to include further features such as the charging features discussed above.

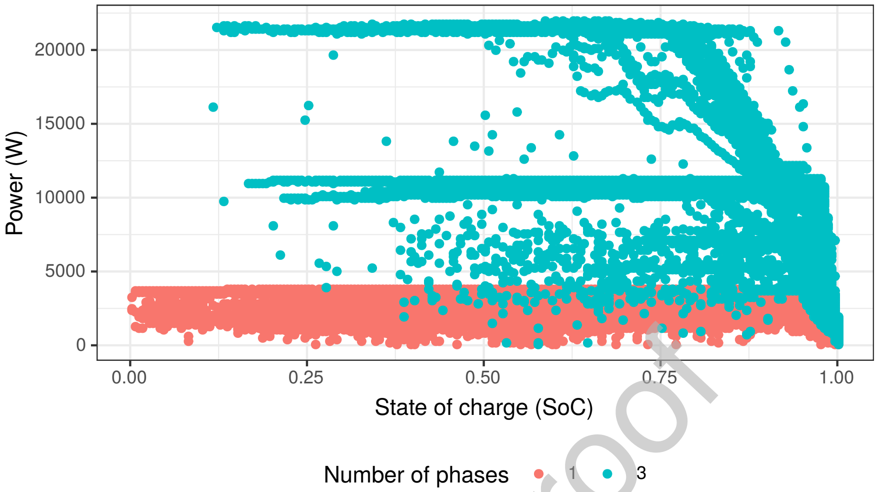


Figure 9: Power vs SoC for all data points in the dataset

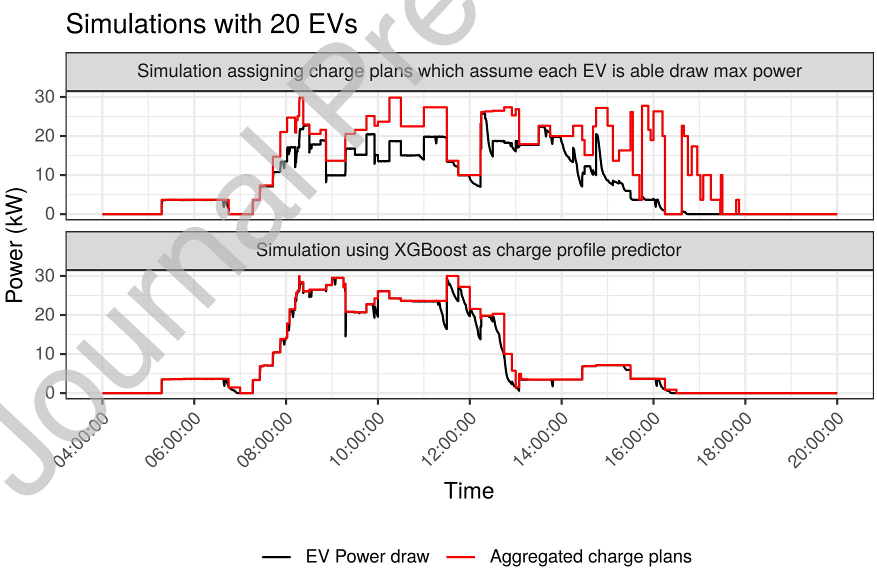


Figure 10: Single simulation runs using different predictors (20 EVs)

535 **Example simulation runs.** Figure 10 shows the aggregated consumed power for two different simulations using different predictors. Notably, when

using *Constant (max power)*, there are large differences between the actually drawn power and the aggregated charge plans. In other words, there is a gap between each EV's power and the power assigned via a charge plan. Semantically, the gap represents the difference between assuming a constant battery model and taking into account EV charge profiles.

In the upper plot of figure 10, 196kWh was planned but only 129kWh was drawn which represents a charge plan utilization of 65.8%. In the lower plot, XGBoost more accurately predicts charge profiles leading to 144kWh planned and 136kWh drawn and a charge plan utilization of 94.4%.

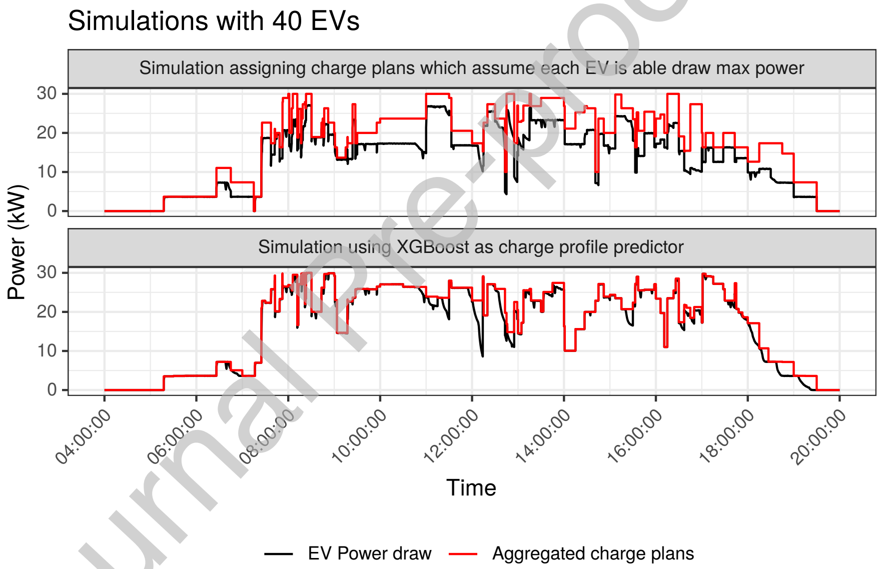


Figure 11: Single simulation runs using different predictors (40 EVs)

Figure 11 shows the aggregated consumed power for two simulations with 40 EVs. With 40 EVs, the impact of a lower charge plan utilization is more pronounced and leads to 21% more energy being drawn in total (259kWh vs 213kWh) when using the more accurate prediction method (XGBoost). That is, the accuracy of the charge profile prediction impacts how effectively the infrastructure is used. Consequently, using XGBoost leads to leads to a significantly higher mean final SoC of 98% compared to the baseline of *constant (max power)* of 90%.

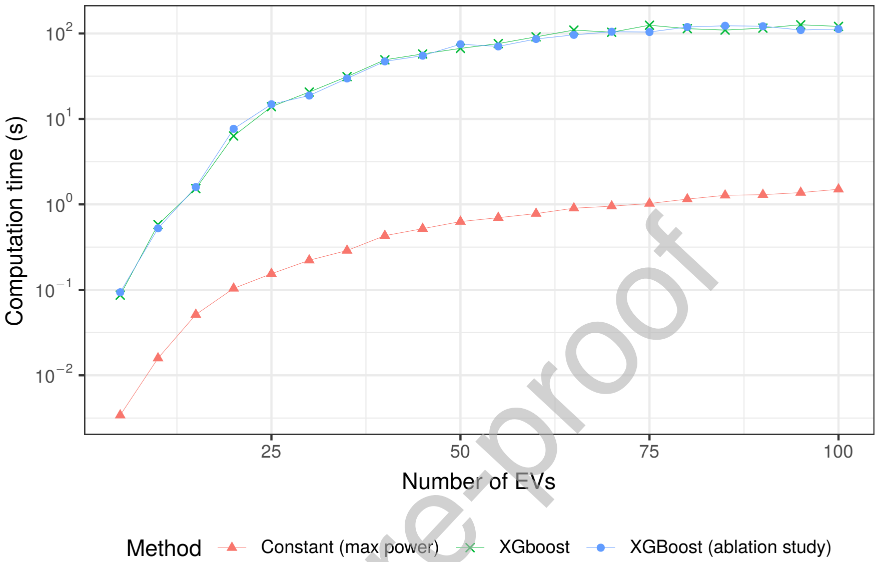


Figure 12: Computation time of the smart charging algorithm per integrated prediction method

Smart charging computation time. Figure 12 shows the average computation time per rescheduling operation of the smart charging algorithm. In other words, how long does an EV wait, on average, before receiving a charge plan? The computation time includes the time applying the regression model as well as computing charge plans.

In a real-time context, charge plans should be computed as quickly as possible in order to make the most of the EV's stay. Both methods show increasing computation time depending on the number of EVs. Using *Constant (max power)* is clearly faster. The smart charging algorithm performs many charge profile predictions with the integrated XGBoost model. The computation time of each XGBoost prediction is roughly 30ns and is independent of the number of EVs. The number of EVs influences how often the prediction method is used.

For the smart charging algorithm, the overall computation time remains in the order of 100s for fleets of up to 100 EVs.

6. Discussion

Using a data-driven approach to reflect charging behavior in smart charging leads to improved charge plans. Here, improvement is understood as reaching a higher mean final SoC for the vehicle fleet by using the limited charging infrastructure more efficiently. The EV type (BEV versus PHEV, three phase charging) and SoC data is sufficient to train a regression model reflecting charge profiles. The smart charging approach proposed in this work reaches an improvement of up to 21% in charging power consumed in total over an approach which does not take into account charging behavior. An ablation study showed how contrary to initial intuition, the EV's model is neither necessary nor does it produce a significant improvement in the performance of the regression model.

From a computation time perspective, the integration of the regression model is negligible with an average runtime of 30ns per prediction.

With regard to the impact of prediction accuracy: If the power prediction is too low the EV will draw less power than it is capable of. However, the smart charging algorithm can assign the unused infrastructure capacity to other EVs if available. If the power prediction is too high the EV will follow its normal charging profile and draw less power than specified by the charge plan. Consequently, infrastructure capacity is reserved which cannot be assigned to other EVs.

The studied scenario of workplace charging generalizes to other smart charging applications that involve EVs topping up batteries in limited charging infrastructures. Such applications include charging delivery fleets over night and topping up EV batteries at gyms, airports or retail locations. It can be argued that the main gain of integrating the regression model is leveraged when vehicles approach 100% SoC. However, it is to be expected that with an increasing number of charging opportunities frequent topping up becomes common charging behavior.

The fact that the regression model for the charge profile prediction is agnostic of EV models makes it versatile and applicable to use with any EV. Regarding continuous SoC data, the vehicles' SoC is not commonly accessible in practice due to pending implementations of the ISO15118 [60] standard in charging stations and EVs. The missing access to SoC is expected to resolve with time. We describe how we estimate SoC for historical charging processes in this work in section 3.1.

Additional features such as EV internal temperature, battery SoH and charging efficiency may improve the accuracy of the regression models. However, such features are not reported to the charging station and thus not available to the CPO. Furthermore, we omit external temperature from the set of regression model features. Retraining the regression models with external temperature as an additional feature did not improve accuracy.

The limitations of the proposed approach relate to the availability of historical data. A large enough dataset with car models and meter data from charging processes is necessary to train the regression model for predicting charge profiles. The trained model only reflects the charging behaviour of the EV models included in the historical dataset. With new EV models and EVs aging over time the dataset will need to be continuously updated and the regression model will need to be retrained. Furthermore, in this work we assume we know each EV's departure time. In practice, it is difficult to reliably estimate departure times. We propose an approach to predict EV departure time with machine learning in previous work [61].

As far as alternative approaches to charge profile prediction are concerned, time series forecasting may at first seem to suggest itself. However, the application of charge plans influences the course of each time series making the integration of time series forecasting with real-time smart charging conceptually problematic.

7. Conclusion

In this work we propose the integration of a regression model for charge profile prediction in a smart charging algorithm. The regression model is trained on a large historical dataset of charging processes and predicts the power drawn by the EV over the course of the charging process. A data-driven regression model is more practical to infer charge profiles than traditional battery models such as ECMs and physical models. Battery internal parameters such as current, voltage or state of health are required for such analytical models but are not publicly available. Additionally, the regression model is trained on arbitrary charge profiles and thus not restricted to a single charge profile such as CCCV.

We address the research question of how integrated EV charge profile predictions impact smart charging. We show how with our approach EVs charge up to 21% more energy and reach a 9 percentage point higher mean final SoC in a limited infrastructure. Consequently, more energy can be delivered without the need for costly and time-consuming upgrades of the charging infrastructure.

Furthermore, an ablation study shows that the EV model is not a necessary attribute for considering charge profiles in heterogeneous EV fleets. However, EV characteristics are required including the type (BEV or PHEV) and the number of the phases used for charging.

Future work includes studying the impact on smart charging of how well regression models generalize to new EV models as they enter the fleet. Progress in battery technology leads to EV models with new charge profiles not contained in the historical dataset.

Appendix

Car model	**Type**	**Three phase charging**	**Capacity (kWh)**	**Max power (kW)**
Audi A3 e-tron	PHEV	×	8.8	3680
BMW 225xe	PHEV	×	7.6	3680
BMW 330e	PHEV	×	12.0	3680
BMW 530e	PHEV	×	12.0	3680
BMW i3	BEV	✓	42.2	11040
Hyundai Kona 150kW	BEV	✓	64.0	11040
MINI Cooper S E Countryman	PHEV	×	7.6	3680
Mercedes Benz B250e	BEV	✓	28.0	11040
Mercedes Benz C 350e	PHEV	×	7.0	3680
Mercedes Benz E 300de	PHEV	×	13.5	7360
Mercedes Benz E 350e	PHEV	×	6.2	3680
Mercedes Benz GLC 350e	PHEV	×	13.5	7360
Mercedes Benz GLC 350e COUPE	PHEV	×	7.0	3680
Mercedes Benz GLE 500e	PHEV	×	8.8	3680
Nissan Leaf	BEV	×	62.0	6600
Renault ZOE	BEV	✓	41.0	22080
Smart ED	BEV	✓	17.6	22080
Smart fortwo ED	BEV	✓	17.6	22080
Smart fortwo EQ Cabrio	BEV	✓	17.6	22080
Tesla Model S	BEV	✓	75.0	11040
VW Golf GTE	PHEV	×	8.7	3680
VW Passat GTE	PHEV	×	9.9	3680
VW e-Golf	BEV	×	35.8	7200
Volvo V60 2,4 PHEV	PHEV	×	11.6	3680

Table 5: Car model data from manufacturers (2018 models)

References

[1] Z. J. Lee, T. Li, S. 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, E-Energy '19, Association for Computing Machinery, Phoenix, AZ, USA, 2019, pp. 139–149. doi:10.1145/3307772.3328313.

[2] Q. Wang, X. Liu, J. Du, F. Kong, Smart Charging for Electric Vehicles: A Survey From the Algorithmic Perspective, IEEE Communications Surveys Tutorials 18 (2) (Secondquarter 2016) 1500–1517. doi:10.1109/COMST.2016.2518628.

[3] P. Kong, G. K. Karagiannidis, Charging Schemes for Plug-In Hybrid Electric Vehicles in Smart Grid: A Survey, IEEE Access 4 (2016) 6846–6875. doi:10.1109/ACCESS.2016.2614689.

[4] J. C. Mukherjee, A. Gupta, A Review of Charge Scheduling of Electric Vehicles in Smart Grid, IEEE Systems Journal 9 (4) (2015) 1541–1553. doi:10.1109/JSYST.2014.2356559.

[5] Z. J. Lee, D. Chang, C. Jin, G. S. Lee, R. Lee, T. Lee, S. H. Low, Large-Scale Adaptive Electric Vehicle Charging, in: 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2018, pp. 863–864. doi:10.1109/GlobalSIP.2018.8646472.

[6] G. L. Plett, Battery Management Systems, Volume II: Equivalent-Circuit Methods, Artech House, 2015.

[7] Q. Lin, J. Wang, R. Xiong, W. Shen, H. He, Towards a smarter battery management system: A critical review on optimal charging methods of lithium ion batteries, Energy 183 (2019) 220–234. doi:10.1016/j.energy.2019.06.128.

[8] M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit, Z. W. Seh, Predicting the state of charge and health of batteries using data-driven machine learning, Nature Machine Intelligence (2020) 1–10doi:10.1038/s42256-020-0156-7.

[9] G. L. Plett, Battery Management Systems, Volume I: Battery Modeling, Artech House, 2015.

[10] K. N. Kumar, B. Sivaneasan, P. L. So, Impact of Priority Criteria on Electric Vehicle Charge Scheduling, IEEE Transactions on Transportation Electrification 1 (3) (2015) 200–210. doi:10.1109/TTE.2015.2465293.

[11] Z. Darabi, P. Fajri, M. Ferdowsi, Intelligent Charge Rate Optimization of PHEVs Incorporating Driver Satisfaction and Grid Constraints, IEEE Transactions on Intelligent Transportation Systems 18 (5) (2017) 1325–1332. doi:10.1109/TITS.2016.2621049.

[12] Y. Yang, Q. Jia, X. Guan, X. Zhang, Z. Qiu, G. Deconinck, Decentralized EV-Based Charging Optimization With Building Integrated Wind Energy, IEEE Transactions on Automation Science and Engineering 16 (3) (2019) 1002–1017. doi:10.1109/TASE.2018.2856908.

[13] M. Dronia, M. Gallet, CoFAT 2016 - Field test of charging management system for electric vehicle - State of the art charging management using ISO 61851 with EV from different OEMs, in: 5th Conference on Future Automotive Technology, 2016.

[14] O. Frendo, N. Gaertner, H. Stuckenschmidt, Real-Time Smart Charging Based on Precomputed Schedules, IEEE Transactions on Smart Grid 10 (6) (2019) 6921–6932. doi:10.1109/TSG.2019.2914274.

[15] S. Pelletier, O. Jabali, G. Laporte, Charge scheduling for electric freight vehicles, Transportation Research Part B: Methodological 115 (2018) 246–269. doi:10.1016/j.trb.2018.07.010.

[16] Y. A. Sha'aban, A. Ikpehai, B. Adebisi, K. M. Rabie, Bi-Directional Coordination of Plug-In Electric Vehicles with Economic Model Predictive Control, Energies 10 (10) (2017) 1507. doi:10.3390/en10101507.

[17] A. E. Trippe, R. Arunachala, T. Massier, A. Jossen, T. Hamacher, Charging optimization of battery electric vehicles including cycle battery aging, in: IEEE PES Innovative Smart Grid Technologies, Europe, 2014, pp. 1–6. doi:10.1109/ISGTEurope.2014.7028735.

[18] Y. Huang, A Day-Ahead Optimal Control of PEV Battery Storage Devices Taking into Account the Voltage Regulation of the Residential Power Grid, IEEE Transactions on Power Systems (2019) 1–1doi:10.1109/TPWRS.2019.2917009.

[19] Y. Cao, R. C. Kroeze, P. T. Krein, Multi-timescale Parametric Electrical Battery Model for Use in Dynamic Electric Vehicle Simulations, IEEE Transactions on Transportation Electrification 2 (4) (2016) 432–442. doi:10.1109/TTE.2016.2569069.

[20] A. Farmann, D. U. Sauer, Comparative study of reduced order equivalent circuit models for on-board state-of-available-power prediction of lithium-ion batteries in electric vehicles, Applied Energy 225 (2018) 1102–1122. doi:10.1016/j.apenergy.2018.05.066.

[21] S. S. Mansouri, P. Karvelis, G. Georgoulas, G. Nikolakopoulos, Remaining Useful Battery Life Prediction for UAVs based on Machine Learning, IFAC-PapersOnLine 50 (1) (2017) 4727–4732. doi:10.1016/j.ifacol.2017.08.863.

[22] R. Donato, G. Quiles, Machine learning systems based on xgBoost and MLP neural network applied in satellite lithium-ion battery sets impedance estimation, Adv. Comput. Intell. Int. J 5 (2018) 1–20.

[23] L. K. K. Maia, L. Drünert, F. La Mantia, E. Zondervan, Expanding the lifetime of Li-ion batteries through optimization of charging profiles, Journal of Cleaner Production 225 (2019) 928–938. doi:10.1016/j.jclepro.2019.04.031.

[24] Z. Song, X. Wu, X. Li, J. Sun, H. F. Hofmann, J. Hou, Current Profile Optimization for Combined State of Charge and State of Health Estimation of Lithium Ion Battery Based on Cramer–Rao Bound Analysis, IEEE Transactions on Power Electronics 34 (7) (2019) 7067–7078. doi:10.1109/TPEL.2018.2877294.

[25] A. Di Giorgio, F. Liberati, S. Canale, Electric vehicles charging control in a smart grid: A model predictive control approach, Control Engineering Practice 22 (2014) 147–162. doi:10.1016/j.conengprac.2013.10.005.

[26] R. Halvgaard, N. K. Poulsen, H. Madsen, J. B. Jørgensen, F. Marra, D. E. M. Bondy, Electric vehicle charge planning using Economic Model Predictive Control, in: 2012 IEEE International Electric Vehicle Conference, 2012, pp. 1–6. doi:10.1109/IEVC.2012.6183173.

[27] K. Qian, C. Zhou, M. Allan, Y. Yuan, Load model for prediction of electric vehicle charging demand, in: 2010 International Conference on Power System Technology, 2010, pp. 1–6. doi:10.1109/POWERCON.2010.5666587.

[28] D. Panahi, S. Deilami, M. A. S. Masoum, S. M. Islam, Forecasting plug-in electric vehicles load profile using artificial neural networks, in: 2015 Australasian Universities Power Engineering Conference (AUPEC), 2015, pp. 1–6. doi:10.1109/AUPEC.2015.7324879.

[29] M. B. Arias, S. Bae, Electric vehicle charging demand forecasting model based on big data technologies, Applied Energy 183 (2016) 327–339. doi:10.1016/j.apenergy.2016.08.080.

[30] M. Duan, A. Darvishan, R. Mohammaditab, K. Wakil, O. Abedinia, A novel hybrid prediction model for aggregated loads of buildings by considering the electric vehicles, Sustainable Cities and Society 41 (2018) 205–219. doi:10.1016/j.scs.2018.05.009.

[31] A. Ashtari, E. Bibeau, S. Shahidinejad, T. Molinski, PEV Charging Profile

Prediction and Analysis Based on Vehicle Usage Data, IEEE Transactions on Smart Grid 3 (1) (2012) 341–350. doi:10.1109/TSG.2011.2162009.

[32] K. Nandha, P. H. Cheah, B. Sivaneasan, P. L. So, D. Z. W. Wang, Electric vehicle charging profile prediction for efficient energy management in buildings, in: 2012 10th International Power Energy Conference (IPEC), 2012, pp. 480–485. doi:10.1109/ASSCC.2012.6523315.

[33] D. Zhou, F. Gao, A. Ravey, A. Al-Durra, M. G. Simões, Online energy management strategy of fuel cell hybrid electric vehicles based on time series prediction, in: 2017 IEEE Transportation Electrification Conference and Expo (ITEC), 2017, pp. 113–118. doi:10.1109/ITEC.2017.7993256.

[34] Z. Yi, M. Shirk, Data-driven optimal charging decision making for connected and automated electric vehicles: A personal usage scenario, Transportation Research Part C: Emerging Technologies 86 (2018) 37–58. doi:10.1016/j.trc.2017.10.014.

[35] G. Box, G. Jenkins, G. Reinsel, Time Series Analysis, Fourth Edition, fourth edition Edition, John Wiley & Sons, Inc, 2008.

[36] M. H. Amini, A. Kargarian, O. Karabasoglu, ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation, Electric Power Systems Research 140 (2016) 378–390. doi:10.1016/j.epsr.2016.06.003.

[37] M. Ghofrani, A. Arabali, M. Ghayekhloo, Optimal charging/discharging of grid-enabled electric vehicles for predictability enhancement of PV generation, Electric Power Systems Research 117 (2014) 134–142. doi:10.1016/j.epsr.2014.08.007.

[38] M. Majidpour, C. Qiu, P. Chu, R. Gadh, H. R. Pota, Fast Prediction for Sparse Time Series: Demand Forecast of EV Charging Stations for Cell Phone Applications, IEEE Transactions on Industrial Informatics 11 (1) (2015) 242–250. doi:10.1109/TII.2014.2374993.

[39] M. G. Flammini, G. Prettico, A. Julea, G. Fulli, A. Mazza, G. Chicco, Statistical characterisation of the real transaction data gathered from electric vehicle charging stations, Electric Power Systems Research 166 (2019) 136–150. doi:10.1016/j.epsr.2018.09.022.

[40] E. Xydas, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll, M. Barker, A data-driven approach for characterising the charging demand of electric vehicles: A UK case study, Applied Energy 162 (2016) 763–771. doi:10.1016/j.apenergy.2015.10.151.

[41] R. R. Desai, R. B. Chen, W. Armington, A Pattern Analysis of Daily Electric Vehicle Charging Profiles: Operational Efficiency and Environmental Impacts, Journal of Advanced Transportation (2018) e6930932doi:https://doi.org/10.1155/2018/6930932.

[42] M. Straka, P. De Falco, G. Ferruzzi, D. Proto, G. van der Poel, S. Khormali, Ľ. Buzna, Predicting popularity of EV charging infrastructure from GIS data, arXiv:1910.02498 [cs, econ, stat] (Oct. 2019). arXiv:1910.02498.

[43] E. C. Kara, J. S. Macdonald, D. Black, M. Bérges, G. Hug, S. Kiliccote, Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach, Applied Energy 155 (2015) 515–525. doi:10.1016/j.apenergy.2015.05.072.

[44] C. Li, C. Liu, K. Deng, X. Yu, T. Huang, Data-Driven Charging Strategy of PEVs Under Transformer Aging Risk, IEEE Transactions on Control Systems Technology 26 (4) (2018) 1386–1399. doi:10.1109/TCST.2017.2713321.

[45] John Snow Labs, Electric Vehicle Charging Network, https://datahub.io/JohnSnowLabs/electric-vehicle-charging-network, last accessed 11 Mar. 2020 (2018).

[46] L. Markram, Electric Vehicle Charging Stations: Energy Consumption &

Savings, https://bouldercolorado.gov/open-data/electric-vehicle-charging-stations, last accessed 11 Mar. 2020 (Feb. 2020).

[47] M. Muratori, Impact of uncoordinated plug-in electric vehicle charging on residential power demand, Nature Energy 3 (3) (2018) 193. doi:10.1038/s41560-017-0074-z.

[48] J. Smart, S. Schey, Battery Electric Vehicle Driving and Charging Behavior Observed Early in The EV Project, SAE International Journal of Alternative Powertrains 1 (1) (2012) 27–33.

[49] C. Z. El-Bayeh, I. Mougharbel, M. Saad, A. Chandra, D. Asber, S. Lefebvre, Impact of Considering Variable Battery Power Profile of Electric Vehicles on the Distribution Network, in: 2018 4th International Conference on Renewable Energies for Developing Countries (REDEC), 2018, pp. 1–8. doi:10.1109/REDEC.2018.8597742.

[50] SAP, Electric vehicle charging records 2016-2019, unpublished data. (2019).

[51] J. Sears, D. Roberts, K. Glitman, A comparison of electric vehicle Level 1 and Level 2 charging efficiency, in: 2014 IEEE Conference on Technologies for Sustainability (SusTech), 2014, pp. 255–258. doi:10.1109/SusTech.2014.7046253.

[52] A. Kieldsen, A. Thingvad, S. Martinenas, T. M. Sørensen, Efficiency Test Method for Electric Vehicle Chargers, in: Proceedings of EVS29 - International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, 2016.

[53] E. Apostolaki-Iosifidou, P. Codani, W. Kempton, Measurement of power loss during electric vehicle charging and discharging, Energy 127 (2017) 730–742. doi:10.1016/j.energy.2017.03.015.

[54] F. Marra, G. Y. Yang, C. Træholt, E. Larsen, C. N. Rasmussen, S. You, Demand profile study of battery electric vehicle under different charging options, in: 2012 IEEE Power and Energy Society General Meeting, 2012, pp. 1–7. doi:10.1109/PESGM.2012.6345063.

[55] E. Chemali, P. J. Kollmeyer, M. Preindl, A. Emadi, State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach, Journal of Power Sources 400 (2018) 242–255. doi:10.1016/j.jpowsour.2018.06.104.

[56] T. Zahid, K. Xu, W. Li, C. Li, H. Li, State of charge estimation for electric vehicle power battery using advanced machine learning algorithm under diversified drive cycles, Energy 162 (2018) 871–882. doi:10.1016/j.energy.2018.08.071.

[57] F. Chollet, Keras, https://keras.io (2015).

[58] T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, ACM, New York, NY, USA, 2016, pp. 785–794. doi:10.1145/2939672.2939785.

[59] N. G. Reich, J. Lessler, K. Sakrejda, S. A. Lauer, S. Iamsirithaworn, D. A. T. Cummings, Case study in evaluating time series prediction models using the relative mean absolute error, The American statistician 70 (3) (2016) 285–292. doi:10.1080/00031305.2016.1148631.

[60] J. Schmutzler, C. A. Andersen, C. Wietfeld, Evaluation of OCPP and IEC 61850 for smart charging electric vehicles, in: 2013 World Electric Vehicle Symposium and Exhibition (EVS27), 2013, pp. 1–12. doi:10.1109/EVS.2013.6914751.

[61] O. Frendo, N. Gaertner, S. Heiner, Improving Smart Charging Prioritization By Predicting Electric Vehicle Departure Time (forthcoming), IEEE Transactions on Intelligent Transportation Systems (2020). doi:10.1109/TITS.2020.2988648.

Declaration of interests

☑ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: