Data-driven charging strategies for grid-beneficial, customer-oriented and battery-preserving electric mobility

Karl Schwenk\textsuperscript{1,2}, Tim Harr\textsuperscript{2}, René Großmann\textsuperscript{2}, Riccardo Remo Appino\textsuperscript{1}, Veit Hagenmeyer\textsuperscript{1}, Ralf Mikut\textsuperscript{1}

\textsuperscript{1} Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology
Karlsruhe, Germany
E-Mail: karl.schwenk@kit.edu

\textsuperscript{2} eDrive Innovations, Daimler AG
Sindelfingen, Germany
E-Mail: karl.schwenk@daimler.com

1 Introduction

Electric Vehicle (EV) penetration, renewable energies, and customer orientation of car manufacturers [1] enables synergies between energy supply, vehicle users, and the mobility sector. However, also new challenges arise [2] which we target to examine with three types of agents and their perspectives: EV user, power supplier and car manufacturer. Although many research papers in the literature describe single perspectives, a threefold contemplation of their connections as shown in Fig. 1, is still missing. In the following, we briefly present the single perspectives and respective interactions.

To satisfy their individual needs, vehicle users expect their EVs to be just as reliable and convenient to use, as known from combustion engine cars [3, 4]. However, limited range and longer charging times of EVs complicate individual mobility, manifesting e.g. in range anxiety [2, 5]. For EV users also the power supply interaction changes, as not only home appliances but
Electric mobility also poses new issues for car manufacturers: During charging and discharging of EV batteries a degradation (battery aging) occurs [15], that correlates with a value depreciation of the entire EV. However, EV users’ satisfaction requires reliable and value-stable products, which car manufacturers aim to achieve by offering services, such as charging assistants [6]. The provided charging strategies target simplified and sustainable EV usage by considering individual customer needs and battery aging.
The remainder of this paper is structured as follows: To identify and develop missing models of the outlined problem, a general approach is presented in Section 4. Then, Section 3 describes an online learning framework for data acquisition, storage and application. In Section 4, we propose two data-driven consumption models. Finally, Section 5 gives a brief summary and outlook on future work.

2 Approach

Despite existing model predictive approaches [16, 17] that support EV users while charging, individualized charging strategies are in general still inadequately dealt with in the literature. We target to identify missing models that quantify perspectives and interactions, for which Fig. 2 shows a schematic map (the green boxes represent unexplored models requiring further evaluation).

Figure 2: Schematic map of models to describe perspectives and interactions, Automated charging algorithm (orange box) comprises missing models (green boxes)
To automatically create charging plans that satisfy the interests of power suppliers, EV users, and car manufacturers at the same time (cf. Section 1), an **Automated charging algorithm** will be required (Fig. 2, orange box). A **Charging behavior model** (cf. Section 2.1) could determine Charging Habits (CH, Fig. 2, green arrow) based on behavioral data of **Electric vehicle users**. Thereby, charging plans provided to the EV user can be adapted to individual requirements. In order to obtain user- and route-specific Estimated Consumption (EC, Fig. 2, blue arrows) for the EV, a data-driven **Consumption model** (cf. Section 4) may use driving information of the EV. The EC can also help to estimate the remaining driving range of the EV. Further, charging plans should be adapted to be battery preservative by means of a **Battery degradation model** (cf. Section 2.2) that provides a Predicted Degradation (PD, Fig. 2, red arrows) for specific charging strategies. Using the PD, car manufacturers may also monitor the condition of operating EV batteries. A **Dynamic charging energy map** (cf. Section 2.3) could calculate the Aggregated Energy Demand (AED, Fig. 2, pink arrows) based on charging information of several EVs. The **Power supplier** may then make use of the AED for improved load forecast precision. Together with information about momentary grid load conditions obtained from Power suppliers, peak shaving and load balancing applications can be included in the charging plan. An **Optimization** approach may finally conclude all the information and calculate a charging plan for the EV user.

### 2.1 Charging behavior model

The effectiveness of charging strategies depends on user acceptance, which in turn requires user-specific charging strategies [6]. Therefore, a suitable characterization of users’ driving and charging behavior is necessary [18]. Existing approaches in the literature distinguish between qualitative and quantitative ones. Psychological approaches, e.g. deducted from customer surveys [19, 2, 20] allow to characterize basic user types. Therein, *gamification* and *incentivation* [2] methods are developed to increase user acceptance. Quantitative approaches aim at predicting driving and charging behavior of EV users, e.g. to infer energy demand [21, 22, 23, 24]. However, an adequate user
integration, i.e. individualized charging strategies based on EV users’ charging habits is not represented.

To attain a more universal characterization of the EV users’ charging behavior, different user types need to be determined by analyzing behavioral data of the EV user. Thereby, different user clusters can be identified, for which individual charging habits can be determined (cf. Fig. 2, left). To clarify the EV users’ objectives, we plan a customer survey in future work. Therein, the basic requirements towards automated charging strategies are inquired from i) customers using a conventional car, ii) customers about to acquire an EV, iii) customers just recently started using an EV, and iv) customers already using an EV for a longer time. Specific questions on the users’ mobility requirements, charging habits, doubts, and expectations are supposed to reveal further requirements for individualized charging strategies. Subsequently, a data-driven analysis of charging processes and EV user information may support these findings and may allow creating a mathematical model.

2.2 Battery degradation model

The interest of car manufacturers focuses on the depreciation process of the EV and its key factors. As the battery considerably contributes to the vehicle value [25], we particularly consider battery degradation. Present-day mobility concepts (e.g. car sharing, leasing) comprise the vehicle battery to remain property of the manufacturer [26]. Value-stable batteries are thus even more relevant to operate economically. To quantify battery aging, usually the State of Health\(^1\) (SOH) is used [27, 28]. Hitherto SOH models have limited practicability due to a complex execution [29] or inaccurate State of Charge (SOC) estimations [30]. Data-driven methods, e.g. Bayesian networks and neural networks [31, 32] hold feasible alternatives for SOH estimation [33]. However, the existing approaches barely use user-related data, for which reason the influence of charging behavior on battery aging is not represented properly.

Hence, we target to evaluate the influence of charging strategies on battery degradation (cf. Fig. 2, center right). By means of a neural network re-

\(^1\)Ratio of actual usable capacity in relation to the nominal battery capacity
progression model, we target to estimate the energy consumption for driving (cf. Section 4), not taking any battery aging influences into account. The model trained on data from batteries without degradation can be used to estimate the energy consumption for EVs with aged batteries. A discrepancy between the estimation and the real consumption indicates a battery aging caused by increased internal losses. Together with a feature-augmented map, this model can additionally estimate energy consumption for charging scheduling, that also considers individual driving styles and environmental conditions. Once this battery aging estimation is mature, also a dependency on charging influences could be inferred. A concept validation has to be proceeded with both simulation and real data.

2.3 Dynamic charging energy map

To match energy usage and generation, power suppliers require information on the energy demand at a certain time and place, targeting a fault-free delivery of electric energy. This includes the demand due to charging EVs. Research papers to model the energy demand caused by charging EVs already exist. They differentiate among approaches to predict and control the grid condition [11, 34, 35, 14], and approaches for dimensioning and locating new charging stations [36, 37]. Stochastic assumptions to describe the EV user behavior are a major deficiency of these approaches.

By combining information about operating EVs, the behavior of their drivers, and the associated charging requirements, we target to aggregate the energy demand caused by charging vehicles and predict vehicles’ anticipated charging location.² EVs requiring charging can be consolidated to larger energy demands characterized by a load profile over time and location. Then, an “energy map” can be created, containing the aggregated energy demand in real time, or a prediction for future points in time (cf. Fig. 2, right). The required energy can either be acquired directly from the energy market, or the information can be passed to power suppliers to handle the predicted load accordingly. For concept validation, a simulation environment has to be developed.

²The access to EV information subjects to the present data security legislation [38].
3 Data framework

In the following, we propose a cloud-based framework to acquire, store and analyze vehicle data, as shown in Fig. 3. A telematic system records selected Controller Area Network (CAN) signals for a fleet of connected EVs during driving and charging.

Applications, e.g. distributed functions structured in microservices require a reliable internet connection to provide historical and live data. Existing devices such as data loggers store data in the car locally. However, data can only be transmitted infrequently via infrastructure-bound connections (e.g. WLAN or LAN). A real-time vehicle status can thus not be established and the large amount of data causes the data transmission to be time-intensive. For this reason, we use a single-board computer (SBC) equipped with a GPS antenna and mobile internet connection as a gateway between vehicle and back-end infrastructure. The SBC initially decodes and filters data with libraries stored...
on its internal memory. Via software packages, the SBC can be updated and controlled remotely not requiring a physical contact to the SBC. The data transmission utilizes a Message Queuing Telemetry Transport (MQTT) protocol following a publish and subscribe mechanism\textsuperscript{3}. A central MQTT broker (Fig. 3, left) organizes the data distribution. This mechanism also enables communication into the vehicle. Currently ten EVs provide 370 signals of interest, which we record with a 10-Hz sampling rate\textsuperscript{4}.

We process and store all collected data according to its structure (Fig. 3, middle). A relational database stores vehicle data and meta-information. Further, a document-oriented database allows to store semi-structured signal time series data efficiently. Additionally, a library database allows transforming vehicle specific data into a standardized format.

The stored data can subsequently be used to build models or provide stand-alone microservices with information (Fig. 3, right). Each model or microservice can make use of several data sources, or communicate among each other, respectively. A separate application database contains configurations and meta-information necessary for the microservice operation. The cloud structure allows continuity in various aspects. Models can be easily implemented, changed, monitored and deleted. Software changes can be applied and deployed continuously. Through a constant availability, the models can be fed with new data each time it is produced. This allows a continuous development of the model with a updated data stream. However, a continuous availability also constitutes a risk in terms of information security [39], that needs to be handled accordingly. The modular structure allows to create and train several models in parallel. An Application Database contains all models and their internal, learned parameters (Fig. 3, right). Via a human machine interface, the models and their learning progress can be monitored and analyzed. This allows to supervise the data acquisition, storage, and application. Furthermore, databases can be examined, microservices and models can be adapted, and errors can be handled accordingly. The interface also provides feedback to the vehicle fleet and vehicle users.

\textsuperscript{3}participants can publish information others can subscribe to
\textsuperscript{4}equals data stream of approx. 12 Megabyte per hour and vehicle.
Table 1: Input signals grouped by driving behavior, battery state and environmental condition

<table>
<thead>
<tr>
<th>Driving behavior</th>
<th>Battery state</th>
<th>Environmental cond.</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceleration torque</td>
<td>total electric current</td>
<td>geodetic altitude</td>
</tr>
<tr>
<td>brake torque</td>
<td>battery voltage</td>
<td>ambient temperature</td>
</tr>
<tr>
<td>recuperation torque</td>
<td>battery temperature</td>
<td></td>
</tr>
<tr>
<td>lateral acceleration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vehicle speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>electric current driving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>longitudinal acceleration</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 Data-driven consumption model

Here, we propose a data-driven consumption model as an exemplary application of the framework presented in Section 3. Consumption models, as used in charging scheduling, help to estimate energy demands for routes the user will drive [40]. However, existing estimators mostly utilize physical models neglecting influences of the driver and the environment. Such consumption models yield relative estimation errors between 2.52 % and 8.3 % [41, 42]. A higher model accuracy allows to compute a more reliable and thoroughly adapted charging strategy. Thus, we aim to design a model estimating the total consumed energy per driven distance for a given driving behavior, environmental condition, and battery state. Note, that we use the dimensionless State of Charge (SOC) as a metric for the consumed energy.

4.1 Data selection

To abstract driving and environmental influences we use time series data recorded during driving of the EVs. We differentiate the input signals in the three categories: driving behavior, battery state, and environmental condition (as shown in Table 1). The acceleration torque, brake torque, and recuperation torque are measured for front and rear axle, i.e. two signals for those values exist. Thus, we obtain a total signal number of $N_{\text{signal}} = 15$. 
With a variety of driving situations and environmental conditions as input, we target a universal estimator that generalizes well on arbitrary input data. The samples are sections with duration $t_{sec} = 6$ min of EV trips representing a typical ride. All trips with insufficient duration or faulty signal values are discarded beforehand.\(^5\) To avoid estimation errors due to random noise, the originally 10-Hz-sampled signals are aggregated before passing them to the model. For each signal and for all values within an aggregation time period $t_{agg} = 1$ min the mean is calculated.

With 15 signals multiplied by 6 data points per trip section we obtain the extracted features $x_1,...,x_{90}$. Note that we chose the parameters $t_{sec}$ and $t_{agg}$ according to preliminary test results.\(^6\)

For each data sample representing a trip section of $t_{sec} = 6$ min we calculate a label

$$\Gamma = \frac{\bar{e} - e}{\bar{o} - o}.$$  \hspace{1cm} (1)

Therein, $e$ is the SOC at the start of the trip section, $\bar{e}$ the SOC at the end. Similarly, $o$ is the mileage (in kilometers) at the start of the trip section, $\bar{o}$ at the end, respectively.

### 4.2 Models

To model the dependency between the input features $x_1,...,x_{90}$ and output label $\Gamma$, we design two models.

#### 4.2.1 Model A

For the first Model A we use a linear regression model

$$\hat{\Gamma} = W \cdot x + b,$$  \hspace{1cm} (2)

---

\(^5\)In following evaluations, these special cases need to be contemplated separately.

\(^6\)In future work, a more detailed evaluation on the selection of these parameters is required.
with the input features \( \mathbf{x} = (x_1, x_2, \ldots, x_{90})^\top \) of the aggregated signal time series, the weight matrix \( \mathbf{W} \in \mathbb{R}^{1 \times 90} \), and the bias \( b \in \mathbb{R} \). The weights \( \mathbf{W} \) and the bias \( b \) are chosen to yield a minimum mean squared error (cf. Section 4.3) between the training samples and the regression [43].

### 4.2.2 Model B

For the second Model B we design a neural network regression model. The input layer comprise 90 nodes representing the input features \( x_1, \ldots, x_{90} \). Further, we use five hidden layers in a triangular shape, as shown in Fig. 4. Each of the hidden layer nodes is activated through a Rectified Linear Unit (ReLU). The output layer of the neural network consists of one node representing the estimated consumption \( \hat{\Gamma} \).

For training the model we process all trips into training samples (cf. Section 4.1). Then, the data is fed to the models in batches of 32 samples over 100 epochs. The model is implemented in Python [44] using Keras [45]. We use the Adam optimizer [10] with a learning rate \( \alpha = 0.001 \).

![Diagram of Model B](image)

Figure 4: Schematic architecture Model B, green: input nodes, hidden layer nodes in grey, output node in orange, total number of nodes below each layer.
4.3 Evaluation

The models are trained with data obtained from eight out of ten EVs. For validation, the data obtained from the remaining two EVs is used, i.e. we compare the estimated consumption $\hat{\Gamma}$ with the actual consumption $\Gamma$. To quantify the model performance, we calculate the metrics Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{N} \sum_{n=1}^{N} |\Gamma_n - \hat{\Gamma}_n|,$$

Relative Mean Absolute Error (RMAE)

$$\text{RMAE} = \frac{1}{N} \sum_{n=1}^{N} \left| \frac{\Gamma_n - \hat{\Gamma}_n}{\Gamma_n} \right|,$$

Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^{N} (\Gamma_n - \hat{\Gamma}_n)^2,$$

and Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\Gamma_n - \hat{\Gamma}_n)^2}.$$

Table 2 reports the validation results according to the error metrics (3)-(16). The results show, that the neural network Model B outperforms the linear Model A in all four metrics. Using Model B, the estimation yields a mean absolute error of 0.00461 km$^{-1}$, i.e. for a trip of 1 km driven distance, the battery level change is estimated with an average discrepancy of 0.461% SOC compared to the true value. On the contrary, Model A estimates the battery level change for such a trip with an average discrepancy of 52.783% SOC. Note that an SOC of 100% represents a fully charged battery and an SOC of 0% an empty battery, respectively. Considering this fact, the estimations obtained from Model A do not provide useful information on the EV’s consumption. Showing the
Table 2: Evaluation of Model A and Model B with MAE, RMAE, MSE, and RMSE.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE [km$^{-1}$]</th>
<th>RMAE [-]</th>
<th>MSE [km$^{-2}$]</th>
<th>RMSE [km$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>0.52783</td>
<td>1.45031</td>
<td>0.44548</td>
<td>0.66744</td>
</tr>
<tr>
<td>Model B</td>
<td>0.00461</td>
<td>0.01782</td>
<td>0.00004</td>
<td>0.00651</td>
</tr>
</tbody>
</table>

Figure 5: Estimated consumption $\hat{\Gamma}$ (red dots) of Model A with validation data set compared with the true consumption $\Gamma$, ideal model output (green line), and 5% relative estimation error (green dashed lines).

results graphically further emphasizes this finding: Figure 5 illustrates the validation results of Model A. For each validation sample with consumption $\Gamma$ the estimation $\hat{\Gamma}$ that the model provided is depicted as red dot. The green line represents an ideal model behavior for comparison. The green dashed lines represent an RMAE of 5%. It can be seen, that most of the validation samples are not estimated correctly by the linear Model A, shown by the red dots that are far from the green line. Based on the low accuracy of the linear model, we
assume the real relation between the input features $x_{1,\ldots,90}$ and the consumption $\Gamma$ to be non-linear.

The results of the neural network Model B support this assumption: In the same manner as done for Model A, Fig. 6 shows the estimations for Model B. The results show a higher precision, as the estimations $\hat{\Gamma}$ are located closer to the ideal model representation, i.e. the green line. A mean relative estimation error $\varepsilon_B = 1.78\%$ (cf. Table 2) on the validation data supports this result. According to the results, the internal structure of the neural network Model B seems to represent the influences on the energy consumption with much higher precision than the linear Model A.

![Figure 6: Estimated consumption $\hat{\Gamma}$ (red dots) of Model B with validation data set compared with the true consumption $\Gamma$, ideal model output (green line), and 5% relative estimation error (green dashed lines).](image)

A further improvement of the model performance is expected if further data, e.g. the actual weight of the electric vehicle due to baggage and passengers, would be considered. Cross-validation over several vehicles should be proceeded to test the generalization of the developed model.
Note that the input data excludes signals allowing to infer the state of battery degradation. Thus, the consumption model does not consider effects of battery aging. Prospectively, battery aging can be estimated by training a similar model with data solely obtained from EVs, whose batteries have not degraded yet. Then, a consumption can be estimated for vehicles that are assumed to have an aged battery. A discrepancy between the estimation and the real consumption—beyond the known estimation error—indicates a battery aging due to increased internal losses. This method would allow an entirely data-driven battery aging estimation based on data anyway created within the vehicle operation.

5 Conclusion and perspective

In this paper, we examined perspectives and interactions of Electric Vehicle (EV) users, car manufacturers and power suppliers. We proposed a concept to quantify the objectives of all parties, aiming at automated calculation of EV charging strategies. In this context, we outlined a cloud-based framework for data acquisition, storage and application as a common basis for data-driven analyses. As an example use case, we proposed two data-driven models based on linear regression and neural-network regression to estimate specific consumption, i.e. consumed energy per driven distance. Therefore, we used time series data collected from a fleet of ten connected EVs. For the model training only eight out of ten EVs were used, while the remaining two EVs were used for validation.

The linear model seems to inadequately represent the true relation between input features and consumption. However, the neural network model with five hidden layers yields a mean relative absolute error of 1.78%. Considering the underlying data, the proposed model outperforms estimators based on physical models as described in the literature. More elaborate model validation with different data and other models, such as polynomial models, or other neural network architectures may allow further model precision improvement.

In future work, a model as proposed could estimate battery aging, if the training data is solely obtained from EVs, whose batteries have not degraded yet.
Comparing the estimated consumption with the actual consumption for EVs, which have degraded batteries, might indicate battery aging due to increased internal losses. This method would allow an entirely data-driven battery aging estimation.

Furthermore, the remaining model parts of the proposed concept require a more elaborate implementation and evaluation. The interactions among EV user, power supplier, and car manufacturer should be analyzed and described mathematically. Finally, the findings need to be combined in an optimization approach, enabling the automated calculation of individualized EV charging strategies considering momentary grid load and battery preservation. For concept evaluation, we plan to use an experimental EV fleet within measure campaigns in the Energy Lab 2.0 [47]. Concluding, the feasibility of grid-optimal, battery-preserving and individualized charging strategies needs to be investigated with adequate indexes, such as user acceptance.

References


