# Generating realistic data for developing artificial neural network based SOC estimators for electric vehicles

1st Alexis Kalk

Institute of Electrical Engineering (ETI) Karlsruhe Institute of Technology (KIT) Karlsruhe, Germany alexis.kalk@kit.edu

2<sup>nd</sup> Oleg Birkholz APL Automobil-Prüftechnik Landau GmbH Landau, Germany oleg.birkholz@apl-landau.de

3<sup>rd</sup> Jiaming Zhang

4<sup>th</sup> Christian Kupper

5<sup>th</sup> Marc Hiller

Institute of Electrical Engineering (ETI) Institute of Electrical Engineering (ETI) Institute of Electrical Engineering (ETI) Karlsruhe Institute of Technology (KIT) Karlsruhe Institute of Technology (KIT) Karlsruhe Institute of Technology (KIT) Karlsruhe, Germany

Karlsruhe, Germany christian.kupper@kit.edu Karlsruhe, Germany marc.hiller@kit.edu

Abstract—Tracking the state of a lithium-ion battery in an electric vehicle (EV) is a challenging task. In order to tackle one aspect of this task, we choose a data-driven approach for estimating the State of Charge (SOC), which is one of the most import parameters. In this context, the quality of the provided data is of utmost importance. Usually, standardized driving profiles are used to generate current profiles which are then applied to battery cells during testing. However, these standardized driving profiles exhibit significant deviation from real-world conditions, which can considerably affect the learning and validation performance of data-driven approaches. In this paper, we first propose a test profile generator which generates realistic current profiles for EV battery testing. Second, to demonstrate the effect of the proposed test profiles a multilayer perceptron (MLP) based SOC estimator is presented. Finally, we compare the results to the standardized driving profiles.

Index Terms-State of Charge, Artificial Neural Networks, Realistic Driving Cycle, SOC Estimation, Lithium-Ion Battery

### I. Introduction

Lithium-ion batteries (LIBs) are extensively employed as electrochemical energy storage in electrical vehicles (EVs) thanks to their high energy and power density, energy efficiency and long cycle life [1]. Battery management systems (BMS) are required to ensure the safe and efficient operation of LIBs [2]. One of the most important tasks of BMSs is the estimation of State of Charge (SOC). SOC is a crucial indicator which defines the currently available capacity  $(Q_{act})$ relative to the maximum capacity  $(Q_{nom})$  of LIBs [3], which is expressed in percentage as follows:

$$SOC = \frac{Q_{act}}{Q_{nom}} \times 100\% \tag{1}$$

SOC cannot be measured directly and has to be estimated with the help of special algorithms and battery state

variables such as battery voltage, current and temperature [13]. Accurate and reliable SOC estimation improves user experience, prevents overdesign of LIBs and enables more efficient and safe usage of the installed capacity [14]. Since the energy storage system is one of the major cost factors of EVs [15], each improvement of SOC estimation has a direct influence on the economic efficiency and customer acceptance of EVs. Therefore, developing SOC estimation algorithms is one of the most popular research areas in battery domain and numerous methods were proposed in the literature [16]. However, accurate and reliable SOC estimation is still a challenging task. A large number of influencing factors like temperature, history of cell operation (hysteresis effects), aging and nonlinear voltage characteristics of lithium-ion cells make the accurate estimation challenging. All these parameters interact with each other and form a complex, time-varying nonlinear system [1]. In addition, the proposed methods have to deal with extra limitations of implementation like the measurement errors of the BMS. The state-of-the-art estimation algorithms are model based. Plett [17] introduced in 2004 for the first time extended Kalman filter (EKF) for state estimation of LIBs and it has become one of the most commonly used methods. Nevertheless, several data-based or artificial intelligence (AI) methods were proposed in the last decade to overcome the problems of model based algorithms. They offer new possibilities to eliminate difficulties and weaknesses of mathematical modeling of nonlinear systems. Among others, variations of artificial neural networks (ANN) are widely implemented. Like the other data-based methods, ANNs learn from the data, hence the quality and the quantity of the data is crucial. According to the computational learning theory, the data set has to be representative of the problem [18]. To generate experimental battery data, standardized driving

profiles (SDPs) like UDDS, US06, WLTP, DST etc. are used intensively. Table I summarizes some examples of applied ANN architectures and used test profiles for SOC estimation. Numerous different SDPs are proposed in the literature [19]. They are specifically designed and stylized speed profiles for the reproducible measurement of vehicle emissions and fuel consumption of conventional vehicles. Additionally, the configured tests assume a road slope of zero. In this context, they can mimic the complex real life load of EVs and batteries to a limited extent. Therefore, the generalization performance of the developed algorithms can be validated only partially. This validation problem can be overcome by using real-drive measurements like [20] used. But unfortunately, generating real-drive data is excessively time and resource consuming.

To train and test the proposed algorithms, experimental data has to be divided randomly into two disjoint data sets [21]. The number of the required data points to train an algorithm is proportional to its complexity. Because of SDPs limited distance, same cycles are repeated in the generated data.

Repeating the same test profiles builds unrealistic patterns in the collected data. These patterns can cause an overlapping between some data points in training and test data sets. Hence, the performance of the algorithm is likely be overestimated. In [9–12] different profiles are used to generate distinct training and test data sets.

Vehicle simulators are standard tools for vehicle design. In order to generate realistic driving profiles, as an alternative to real-drive and SDP tests, simulators can be cost efficiently used with online data sources. Morlock et al. [22] proposed a similar approach for the prediction of EV's energy consumption.

In this paper, we propose a test profile generator which generates realistic current profiles for EV battery testing. Our algorithm combines an EV simulator with online real-time traffic and route information. This approach provides a reliable and cost efficient way to generate training and test data for ANN based SOC estimators for EVs.

In the following, we first introduce the profile generator algorithm. Second, experimental setup and cell tests for data collection is described. Then, to demonstrate the effect of the proposed test profiles a multilayer perceptron (MLP) based SOC estimator is presented. In the final section, the results and the conclusions are given.

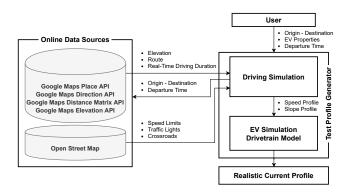


Fig. 1. Structure of Algorithm

#### II. DESIGN OF REALISTIC PROFILE GENERATOR

The proposed algorithm consists of two stages: driving simulation stage and EV simulation stage. The algorithm implemented in MATLAB/Simulink uses online data sources. Fig. 1 shows the structure of the implemented algorithm. At the driving simulation stage, speed and slope profiles are generated for a given route considering speed limits, traffic lamps, road intersections and real-time traffic information. At the EV simulation stage, realistic current profile is derived from the speed and slope profiles with an in-house developed EV drive train model according to the selected vehicle parameters. The consideration of slope allows to simulate the EV's current demand more realistically.

## A. Driving simulation

First, the algorithm collects road and real-time traffic information from online services according to the user inputs i.e. origin, destination and departure time. To this end, we combine Google Maps Platform APIs (GMP) [23] and Open Street Map (OSM) API [24] data. Both platforms support fetching data with HTTP requests via their APIs.

The fastest route is calculated by GMP according to the real-time traffic information. The route information includes encoded polyline representation of the route, which is an approximate path consisting finite number of points described with their coordinates. GMP also provides route length, real-time duration in traffic, base duration (derived from historical data), elevation and road type. Additionally, we obtain data

TABLE I		
EXAMPLES OF TEST PROFILES AND ANN ARCHITECTURES FOR SOC ESTIMATION	V	

Reference	ANN Architecture	Training data	Test data
[4]	MLP	Discharge Test, UDDS	Discharge Test, UDDS
[5]	NARX	US06	US06
[6]	Recurent NARXNN-LSA	FUDS, US06	FUDS, US06
[7]	LSTM-RNN	Combination of 10 different SDPs	Combination of 10 different SDPs
[8]	BPNN	DST, FUDS	DST, FUDS
[9]	FFNN	US06	Pulse test
[10]	Improved BP neural network	1C Discharge test	DST
[11]	Time-delay neural network (TDNN)	Discharge Test, HPPC	DST, FUDS, US06
[12]	MLP + UKF	DST	FUDS, US06

about speed limits, location of traffic lights and intersections from OSM.

- 1) Calculation of Slope Profile: The distance between two polyline points is calculated with Haversine formula using their latitude and longitude information. The slope between these points is calculated by division of elevation difference by distance. Thereafter, the slope profile along the route is generated iteratively.
- 2) Calculation of Speed Profile: The calculation of the speed profile begins with the calculation of an average speed profile. Therefore, the route is divided into  $1000\,\mathrm{m}$  segments. The real time driving duration in traffic of each segment depending on the user-defined departure time is used to calculate the average speed profile  $v_{av}$ . The speed limits are also considered in this calculation.

Saumeister et al. [25] compared mean traffic speeds from the map provider HERE with real drive measurements and assessed the uncertainty of the speeds. To model realistic driver behavior, we added speed deviations to the speed profile for speeds higher than  $50 \,\mathrm{km}\,\mathrm{h}^{-1}$ . In their study, Schwarz and Ghorbani [26] treated speed fluctuations as velocity noise and modeled them as oscillations. They classified velocity oscillations into three categories: Low-Frequency (LF) Noise, which is caused by factors such as terrain topology, traffic congestion, speed limits, and construction areas; Medium-Frequency (MF) Noise, which is caused by road topology, traffic flow, and driving behavior; and High-Frequency (HF) Noise, which is caused by road conditions, lane changes, rapid driving maneuvers, and spontaneous reactions. As our basic speed profile  $v_{av}$  takes into account the effects of the LF noise, we modified the proposed equation to incorporate both MF and HF noise to reflect the influence of factors such as traffic flow, driving behavior, and driving maneuvers. The implemented equation is displayed in (2), where the amplitudes  $A_{FS}^i$ , frequencies  $\omega_{FS}^i$ , and phase  $\phi_{FS}^i$  for each frequency spectrum (FS) i are modeled as normally distributed random variables. The mean values  $\mu$  and variances  $\sigma^2$  of the normal distributions  $\mathcal{N}$  are inherited from [27]. The speed deviation  $v_{dev}$  is calculated for each segment and added to the speed profile  $v_{av}$  to get the first version of the realistic speed profile  $v_{Rpre,1}$ .

$$v_{dev}(t) = \sum_{i=1}^{2} A_{FS}^{i} sin(\omega_{FS}^{i} t + \phi_{FS}^{i})$$
 (2)

$$v_{Rpre,1}(t) = v_{av} + v_{dev} \tag{3}$$

The generated profile  $v_{Rpre,1}$  is then expanded to include information about traffic lights and intersections. Rittger et al. [28] investigated driving behavior at intersections with traffic lights and found that drivers start slowing down approximately 80 m before a red light. Based on this finding, we implemented a random braking process starting 80 m before intersections and traffic lights. The probability of stopping is defined based on the ratio between the duration of the route with and without traffic. The updated speed profile called  $v_{Rpre,2}$ .

To avoid sudden and unrealistic changes in the speed profile  $v_{Rpre,2}$ , we incorporated processes for acceleration, deceleration, and braking into the profile. Acceleration (m s<sup>-2</sup>) is modeled using an exponential function of speed (m s<sup>-1</sup>) as described in (4) [29]. Bokare and Maurya [30] investigated deceleration and proposed a two-step function (5), where the deceleration rate is considered a constant value for speeds higher than  $100 \, \mathrm{km} \, \mathrm{h}^{-1} \, (27.7 \, \mathrm{m} \, \mathrm{s}^{-1})$  and a second-order polynomial is used for lower speeds. For braking, we implemented the step function (6) depending to the speed during braking proposed by Roenitz et al. [31].

$$a_{acc}(v) = 1.90exp(-0.04v)$$
 (4)

$$a_{dcc}(v) = \begin{cases} 0.005v^2 - 0.154v - 0.493 & \text{if } v < 27.7\\ -1.686 & \text{if } 27.7 \le v \end{cases}$$
 (5)

$$a_{br}(v) = \begin{cases} -0.628 & \text{if } v < 1\\ -(0.001296v + 0.175)9.806 & \text{if } 1 \le v \le 6.25\\ -(-0.1188v + 0.330)9.806 & \text{if } 6.25 < v < 15.2\\ -1.686 & \text{if } 15.2 \le v \end{cases}$$
(6)

The final realistic speed profile  $v_R(t)$  is generated by applying the demonstrated acceleration, deceleration, and braking processes into the profile  $v_{Rpre.2}$ .

## B. EV Simulation

The in-house developed ETI drive train model for EVs is employed to simulate the battery current. It consists of the models of a battery, DC-AC inverter, permanent-magnet synchronous motor and mechanical elements such as car body, tires, differential, spur gear etc. In addition to the predefined SDPs, the simulator also allows to use custom current and slope profiles as inputs.

Motor			
Type Power Nominal Torque	Permanent-magnet synchronous motor $100\mathrm{kW}$ $220\mathrm{N}\mathrm{m}$		
	Battery		
Nominal Voltage Capacity Max. Charge Current Max. Discharge Current	400 V 40 kW h 1 C 3 C		
	Chassis		
Weight Drag coefficient Tire radius Frontal area	$1400 \mathrm{kg}$ 0.29 0.3687 m 2.195 m <sup>2</sup>		

TABLE III
GENERATED REALISTIC DRIVING PROFILES

No.	Origin - Destination	Departure Time	Duration	Length	$\overline{v}$	$v_{max}$
1.a	Berlin, Mitte - Berlin, Neuköln	06.02.2023 (Mon) 00:53	27.8 min	16.15 km	$34.3  \mathrm{km}  \mathrm{h}^{-1}$	$80.9 \mathrm{km}\mathrm{h}^{-1}$
1.b	Berlin, Mitte - Berlin, Neuköln	08.02.2023 (Wed) 15:58	$56.8\mathrm{min}$	$16.15\mathrm{km}$	$17.2{\rm km}{\rm h}^{-1}$	$80.4  \mathrm{km}  \mathrm{h}^{-1}$
2.a	Frankfurt, Innenstadt - Frankfurt, Nieder-Erlenbach	05.02.2023 (Sun) 22:12	$22.8\mathrm{min}$	$14.53\mathrm{km}$	$38.5  \mathrm{km}  \mathrm{h}^{-1}$	$113  {\rm km}  {\rm h}^{-1}$
2.b	Frankfurt, Innenstadt - Frankfurt, Nieder-Erlenbach	06.02.2023 (Mon) 17:13	$32.2\mathrm{min}$	$14.53\mathrm{km}$	$27.4{\rm km}{\rm h}^{-1}$	$109.4{\rm km}{\rm h}^{-1}$
3.a	Frankfurt, West - Frankfurt, Mitte-Nord	05.02.2023 (Sun) 22:14	$13.3\mathrm{min}$	$4.9\mathrm{km}$	$22  {\rm km}  {\rm h}^{-1}$	$50  {\rm km}  {\rm h}^{-1}$
3.b	Frankfurt, West - Frankfurt, Mitte-Nord	06.02.2023 (Mon) 17:19	$16\mathrm{min}$	$6.1\mathrm{km}$	$22  {\rm km}  {\rm h}^{-1}$	$60  {\rm km}  {\rm h}^{-1}$
4	Karlsruhe - Berlin	05.02.2023 (Sun) 21:47	$386\mathrm{min}$	$677.5\mathrm{km}$	$104  {\rm km}  {\rm h}^{-1}$	$139  \mathrm{km}  \mathrm{h}^{-1}$
5.a	Frankfurt - Würzburg	08.02.2023 (Wed) 01:23	$78.8\mathrm{min}$	$119.86\mathrm{km}$	$89.2{\rm km}{\rm h}^{-1}$	$122.7{\rm km}{\rm h}^{-1}$
5.b	Frankfurt - Würzburg	09.02.2023 (Thu) 01:58	$77.8\mathrm{min}$	$119.86\mathrm{km}$	$91.3{\rm km}{\rm h}^{-1}$	$128.4{\rm km}{\rm h}^{-1}$
6	Karlsruhe - Zugspitze	05.02.2023 (Sun) 08:30	$207.7\mathrm{min}$	$332.2\mathrm{km}$	$95.2{\rm km}{\rm h}^{-1}$	$136.2{\rm km}{\rm h}^{-1}$
	WLTP NEDC		30 min 19.7 min	23.1 km 11 km	$46.3\mathrm{km}\mathrm{h}^{-1}$ $33.6\mathrm{km}\mathrm{h}^{-1}$	$131.3\mathrm{km}\mathrm{h}^{-1}\\120\mathrm{km}\mathrm{h}^{-1}$

## III. DESIGN OF EXPERIMENT

#### A. Lithium-Ion Battery Cell

As for the chosen cell to perform testing, we used a commercially available cylindrical cell Samsung INR21700-40T. It has a nominal capacity of  $Q_{nom}=4\,\mathrm{Ah}$  and a nominal voltage of  $U_{nom}=3.6\,\mathrm{V}$  and delivers a maximum constant current of  $I_{max}^{const}=35\,\mathrm{A}$ . The cells are charged using the recommended CCCV (Constant Current Constant Voltage) charging method with a  $2\,\mathrm{A}$  constant current rate, a maximum voltage of  $4.2\,\mathrm{V}$  and a cut-off current of  $200\,\mathrm{mA}$ .

#### B. Generating Realistic Test Profiles

In order to generate realistic test profiles, different trips are simulated at different times. The simulated EV has a  $100\,\mathrm{kW}$  permanent-magnet synchronous motor and  $40\,\mathrm{kW}\,\mathrm{h}$  battery with  $400\,\mathrm{V}$  nominal voltage. Table II summarizes the applied drive train model parameters. Table III presents the generated realistic driving profiles and standardized driving profiles, highlighting key characteristics such as average speed  $\overline{v}$ , maximum speed  $v_{max}$ , duration and length. Fig. 2 shows the results of an exemplary trip between Karlsruhe and Landau, Germany. Additionally, current profiles of WLTP and NEDC are also generated with the ETI drive train model. In all cases, the battery current was scaled-down to cell level for cell testing.

## C. Battery Experimental Setup and Tests

In order to generate proper data sets, cell tests are executed with different scenarios. The cells are cycled first with two different SDPs, namely WLTP and NEDC. Additionally, generated current profiles of ten realistic trips are used. All the tests are executed at room temperature  $(25\,^{\circ}\mathrm{C})$ .

The testing hardware used at APL Automobil-Prüftechnik Landau GmbH is a combination of a temperature chamber equipped with an electrical DC controller. The temperature chamber, on the one hand, is commercially available by CTS [32] where the chamber volume is  $1000\,\mathrm{L}$  and the temperature can range from  $-40\,^{\circ}\mathrm{C}$  to  $180\,^{\circ}\mathrm{C}$  by a temperature gradient up to  $3.5\,\mathrm{K\,min}^{-1}$ . The electrical DC controller, on

the other hand, is a commercially available battery tester by Keysight Technologies [33] with 12 DC channels per chamber providing  $6\,\mathrm{V}$  voltage, from  $100\,\mathrm{A}$  (in this case) up to  $300\,\mathrm{A}$  current and  $0.6\,\mathrm{kW}$  to  $1.8\,\mathrm{kW}$  electrical power. In terms of accuracy,  $\pm 1\,\mathrm{mV}$ , concerning voltage, and  $\pm 0.05\,\%$  of measured value with an offset between  $20\,\mathrm{mA}$  and  $60\,\mathrm{mA}$ , concerning electric current, can be achieved. Additionally, the temperature can be measured within a tolerance of  $\pm 1\,\mathrm{K}$  per test channel. During all tests, the voltage, the current and the temperature at the cell surface is tracked and measured.

Reference values for SOC are calculated by (1) with a high accuracy Coulomb-counting approach in order to determine the currently available capacity  $(Q_{act})$ . The initial maximum available capacity of the cell  $(Q_{nom})$  must be periodically adjusted due to degradation effects of the cell. The current maximum available capacity  $(Q_{max,act})$  is determined through specially designed capacity tests, which are conducted intermittently between the driving profile tests and are utilized for SOC calculation. Additionally, the ratio of the current capacity to its initial capacity represents the State of Health (SOH) of the cells, and it is calculated by:

$$SOH = \frac{Q_{max,act}}{Q_{nom}} \times 100\% \tag{7}$$

To minimize the impact of aging effects on results, we utilized cells with similar SOH by generating test data. Fig. 3 shows the cells at the test bench.

### D. Generated Data sets for Training and Testing

The first data set is generated from the cell tests with WLTP. It's split into training and testing data sets with an 70-30 ratio. The cell test measurements of NEDC and realistic current profiles are used to generate test data sets for the validation of the SOC estimator trained with WLTP. All data sets are normalized to improve the learning performance.

#### IV. DESIGN OF MLP BASED SOC ESTIMATOR

A multi-layer perceptron (MLP) neural network algorithm is implemented using Python and Keras for SOC estimation. The network consists of two hidden layers, one input layer, and one

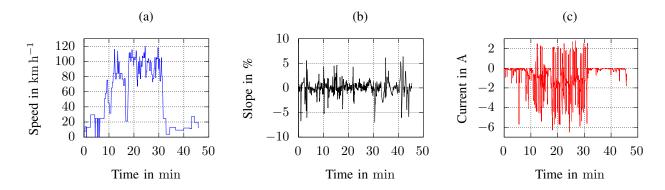


Fig. 2. Exemplary speed (a), slope (b) and cell current (c) profile

output layer. The input variables include cell voltage, current, and temperature, while the output is the SOC. The hidden layers contain 10 and 5 nodes, respectively. Equation (8) describes the general formula for calculation of the output y of a neuron i in a layer k. The output is calculated by applying the activation function f to the weighted sum of the outputs  $y_j^{(k-1)}$  from the neurons in the previous layer (k-1). The weights connecting neurons j from the previous layer to neuron i in layer k are represented by  $w_{ij}^{(k)}$ . Additionally, the bias term for neuron i in layer k is denoted by  $b_i^{(k)}$ . Rectified linear unit (ReLU) is used as activation function for the nodes within the hidden layers. Equation (9) presents the ReLU function, in which the output ranges from 0 to positive infinity. For the output node, a linear activation function is employed, which is typical for addressing regression problems.

In this study, the mean squared error (MSE) is employed as the loss function for the model. The MSE calculates the average squared difference between the predicted values  $\hat{y}_i$  and reference  $y_i$  values. Mathematical description of the



Fig. 3. Cell testing at APL Automobil-Prüftechnik Landau GmbH

MSE is presented in (10). To optimize the model during the training process, the Adam (Adaptive Moment Estimation) optimization algorithm is employed.

$$y_i^{(k)} = f\left(\sum_{j=1}^n w_{ij}^{(k)} y_j^{(k-1)} + b_i^{(k)}\right)$$
(8)

$$\sigma(x) = \max(0, x) \tag{9}$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
 (10)

### V. EXPERIMENTAL RESULTS

To demonstrate the benefits of the proposed method, the MLP-based SOC estimation algorithm presented in the previous section is first trained and validated using the WLTP test data set with a 70% - 30% split. Subsequently, the trained SOC estimator is tested with the experimental data from the driving cycles outlined in Table III. These driving profiles simulate both highway and city driving trips at various times, including rush-hour and off-peak periods. This diverse representation enables a more comprehensive evaluation of the estimator's performance under a range of realistic driving scenarios. The root mean square error (RMSE) of SOC estimation is used as the performance criterion and calculated using (11). The relative estimation error (REE) is calculated with (12) as the relative difference between the RMSE values of the actual test profile (RMSE<sub>TP</sub>) and the WLTP test (RMSE<sub>WLTP</sub>), expressed as a percentage of the RMSE value of the test profiles.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (SOC_i - \widehat{SOC}_i)^2}$$
 (11)

$$REE = \frac{RMSE_{TP} - RMSE_{WLTP}}{RMSE_{TP}} \times 100\%$$
 (12)

Table IV summarizes the performance of the WLTP-trained algorithm under proposed driving profiles according to the RMSE values. As anticipated, the lowest RMSE is achieved for

TABLE IV
THE PERFORMANCE OF THE WLTP-TRAINED SOC ESTIMATOR UNDER
REALISTIC TEST PROFILES

Test Profile	RMSE	Relative Estimation Error to WLTP-Validation
WLTP	1.31%	-
NEDC	1.44%	9.03 %
1.a	1.91%	31.45%
1.b	1.92%	31.77%
2.a	1.97%	33.5%
2.b	1.86%	29.57%
3.a	1.95%	32.82%
3.b	2.18%	39.91%
4	1.74%	24.71%
5.a	1.95%	32.82%
5.b	1.70%	22.94%
6	1.87%	29.95%

the WLTP test. The performance under the NEDC test is comparable to that of the WLTP, which can be attributed to both being standardized profiles. However, the performance under realistic profiles is notably worse. All the realistic tests indicate an overestimation of the performance of SOC estimation with the WLTP test. A mean deviation of 30 % is observed under realistic conditions compared to the established approach.

#### VI. CONCLUSION

In this paper, a method for generating realistic data for the development of SOC estimators for EVs is presented. The experimental results demonstrate that the established approach, which relies on standardized driving profiles, may not accurately represent real driving conditions and may not suffice to be used for data-driven modeling approaches. Consequently, the performance of the developed algorithm is often overestimated. The proposed method can be utilized for the validation and testing of estimation algorithms that employ data-driven approaches. In addition to the performance validation, the proposed method can also be used to produce realistic training data sets in a very cost-efficient way and in large numbers. The proposed method is verified through an exemplary case, where a MLP based SOC estimator is developed, implemented, and tested to demonstrate the capabilities of the approach. In future work, the profile generator will be extended to automatically generate extensive data sets with combined trips, based on driving behavior statistics.

## VII. ACKNOWLEDGMENT

This work was supported by the Helmholtz Association under the program "Energy System Design". The experimental data was provided by APL Automobil-Prüftechnik Landau GmbH.

#### REFERENCES

- [1] D. Linden, Ed., *Handbook of batteries* (McGraw-Hill handbooks), 3, ed. New York: McGraw-Hill, 2002.
- [2] R. Korthauer, Ed., *Handbuch Lithium-Ionen-Batterien*. Berlin: Springer Vieweg, 2013.

- [3] J. Garche and C. K. Dyer, Eds., *Encyclopedia of electrochemical power sources*, Amsterdam, 2009.
- [4] U. Primadusi, A. I. Cahyadi, D. Prasetyo, and O. Wahyunggoro, "Backpropagation neural network models for lifepo4 battery," *AIP Conference Proceedings*, vol. 1755, no. 1, p. 090 009, 2016.
- [5] M. S. H. Lipu, A. Hussain, M. H. M. Saad, A. Ayob, and M. A. Hannan, "Improved recurrent narx neural network model for state of charge estimation of lithium-ion battery using pso algorithm," in 2018 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE), 2018, pp. 354–359.
- [6] M. S. H. Lipu, M. A. Hannan, A. Hussain, M. H. M. Saad, A. Ayob, and F. Blaabjerg, "State of charge estimation for lithium-ion battery using recurrent narx neural network model based lighting search algorithm," *IEEE Access*, vol. 6, pp. 28 150–28 161, 2018.
- [7] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6730–6739, 2018.
- [8] M. A. Hannan, M. S. H. Lipu, A. Hussain, M. H. Saad, and A. Ayob, "Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm," *IEEE Access*, vol. 6, pp. 10069–10079, 2018.
- [9] S. Tong, J. H. Lacap, and J. W. Park, "Battery state of charge estimation using a load-classifying neural network," *Journal of Energy Storage*, vol. 7, pp. 236–243, 2016.
- [10] C.-W. Zhang, S.-R. Chen, H.-B. Gao, K.-J. Xu, and M.-Y. Yang, "State of charge estimation of power battery using improved back propagation neural network," *Batteries*, vol. 4, no. 4, 2018.
- [11] M. S. Hossain Lipu, M. A. Hannan, A. Hussain, A. Ayob, M. H. M. Saad, and K. M. Muttaqi, "State of charge estimation in lithium-ion batteries: A neural network optimization approach," *Electronics*, vol. 9, no. 9, 2020.
- [12] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for li-ion batteries using neural network modeling and unscented kalman filter-based error cancellation," *International Journal of Electrical Power and Energy Systems*, vol. 62, pp. 783–791, 2014.
- [13] G. L. Plett, *Battery management systems*. Boston: Artech House, 2016, vol. 2: Equivalent-circuit methods.
- [14] A. Thaler, *Automotive battery technology*, D. Watzenig, Ed., Cham, 2014.
- [15] A. Kampker, *Elektromobilitaet : Grundlagen einer zukunftstechnologie*, D. Vallee and A. Schnettler, Eds., Berlin, Heidelberg, 2013.
- [16] M. Hannan, M. Lipu, A. Hussain, and A. Mohamed, "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications:

- Challenges and recommendations," *Renewable and Sustainable Energy Reviews*, vol. 78, pp. 834–854, 2017.
- [17] G. L. Plett, "Extended kalman filtering for battery management systems of lipb-based hev battery packs: Part 3. state and parameter estimation," *Journal of Power Sources*, vol. 134, no. 2, pp. 277–292, 2004.
- [18] K.-L. Du and M. N. S. Swamy, *Neural networks and statistical learning*, London, 2019.
- [19] T. Barlow, S. Latham, I. S. McCrae, and P. Boulter, "A reference book of driving cycles for use in the measurement of road vehicle emissions," *TRL Published Project Report*, 2009.
- [20] D. Jimenez-Bermejo, J. Fraile-Ardanuy, S. Castano-Solis, J. Merino, and R. Álvaro-Hermana, "Using dynamic neural networks for battery state of charge estimation in electric vehicles," *Procedia Computer Science*, vol. 130, pp. 533–540, 2018.
- [21] S. J. Russell and P. Norvig, *Artificial intelligence : a modern approach* (Prentice Hall series in artificial intelligence), 2. ed., internat. ed. Upper Saddle River, NJ: Prentice Hall, 2003.
- [22] F. Morlock, B. Rolle, M. Bauer, and O. Sawodny, "Forecasts of electric vehicle energy consumption based on characteristic speed profiles and real time traffic data," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1404–1418, 2020.
- [23] Google Maps Platform Google Developers, https://developers.google.com/maps, [Accessed 14-Dec-2022].
- [24] OpenStreetMap, https://www.openstreetmap.org, [Accessed 14-Dec-2022].
- [25] S. Sautermeister, M. Falk, B. Baeker, F. Gauterin, and M. Vaillant, "Influence of measurement and prediction

- uncertainties on range estimation for electric vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2615–2626, 2018.
- [26] V. Schwarzer and R. Ghorbani, "Drive cycle generation for design optimization of electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 89–97, 2013.
- [27] C. Gutenkunst, "Praediktive routenenergieberechnung eines elektrofahrzeugs," German, Ph.D. dissertation, Karlsruher Institut fuer Technologie (KIT), 2020, 196 pp.
- [28] G. Rittger Lena and Schmidt, C. Maag, and A. Kiesel, "Driving behaviour at traffic light intersections," *Cognition, Technology and Work*, vol. 17, pp. 593–605, 2015.
- [29] A. Mehar, S. Chandra, and S. Velmurugan, "Speed and acceleration characteristics of different types of vehicles on multi-lane highways," *International Journal* of *Transport Economics*, Engineering and Law, 2013.
- [30] P. Bokare and A. Maurya, "Acceleration-deceleration behaviour of various vehicle types," *Transportation Research Procedia*, vol. 25, pp. 4733–4749, 2017, World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016.
- [31] E. Roenitz, A. Happer, R. Johal, and R. Overgaard, "Characteristic vehicular deceleration for known hazards," in *International Congress and Exposition*, SAE International, Mar. 1999.
- [32] *Clima Temperatur Systeme (CTS GmbH)*, https://www.cts-umweltsimulation.de.
- [33] *Keysight Technologies*, https://www.keysight.com/de/de/home.html.