IEEE Open Journal of
Vehicular Technology

Received 21 December 2020; accepted 2 March 2021. Date of publication 11 March 2021;
date of current version 9 April 2021. The review of this paper was arranged by Associate Editor Junmin Wang.

Digital Object Identifier 10.1109/OJVT.2021.3065529

Probabilistic Prediction of Energy Demand and Driving Range for Electric Vehicles With Federated Learning

ADAM THOR THORGEIRSSON [1], STEFAN SCHEUBNER [2], SEBASTIAN FÜNFGELD [3], AND FRANK GAUTERIN [1]

[1]Karlsruhe Institute of Technology, Institute of Vehicle System Technology, 76131 Karlsruhe, Germany
[2]Department of E-Mobility Charging Infrastructure, EnBW AG, 76131 Karlsruhe, Germany
[3]Dr. Ing. h. c. F. Porsche AG, Department of Energy Management, R&D Center Weissach, 71287 Weissach, Germany

CORRESPONDING AUTHORS: ADAM THOR THORGEIRSSON; FRANK GAUTERIN (e-mail: adam.thorgeirsson@partner.kit.edu; frank.gauterin@kit.edu)

This work was supported by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

ABSTRACT Today's drivers of battery electric vehicles must deal with limited driving range in a sparse charging infrastructure. An accurate prediction of energy demand and driving range is therefore important and enables reliable routing and charge planning applications. Predictions of energy demand entail uncertainty, which can be considered directly with the use of probabilistic prediction algorithms. Machine learning algorithms are frequently applied in this context, but data used to train these algorithms are often distributed over a fleet of connected vehicles. Federated learning can be applied in this setting, but predictive uncertainty is typically not considered. We apply an extension of the federated averaging algorithm to learn probabilistic neural networks and linear regression models in a communication-efficient and privacy-preserving manner. We demonstrate the performance advantage of probabilistic prediction models over deterministic prediction models using proper scoring rules. Furthermore, we show that federated learning can improve the standard, driver-individual learning. Using probabilistic predictions, variable safety margins based on destination attainability can be applied, leading to increased effective driving range and reduced travel time.

INDEX TERMS Electric vehicles, energy demand prediction, probabilistic predictions, range estimation.

I. INTRODUCTION

The call for low or zero emissions vehicles, along with improved battery technology, makes the battery electric vehicle (BEV) a serious candidate for the replacement of internal combustion engine powered vehicles (ICEVs). Despite the advantages of such vehicles, they have not gained significant popularity among the general public. Due to limited charging-infrastructure and the inevitably shorter driving range, BEV drivers may experience range anxiety, which is the fear that the energy storage will run out before reaching the destination [1]. In order to eliminate range anxiety and increase the usability of BEVs, there is a need for applications that help drivers in arriving safely at their destinations without excessive time or cost. The primary goals of such applications are to maximize the effective driving range and to accurately predict this range. Drivers tend to reserve up to 20% of the battery capacity as a safety margin [2], i.e., the utilization of the available battery energy is poor. The utilization of the battery strongly depends on the calibration of the driving range prediction [3]. A central challenge in this context is the prediction of future energy demand. The energy demand prediction (EDP) is not only used to display remaining driving range [4], but also for other purposes such as the estimation of a destination's attainability [5], time or energy optimal routing with charge planning [6], [7], energy optimal control [8], [9], BEV fleet management systems [10] and charging infrastructure planning [11].

The EDP and driving range estimation rests upon information about the driver, vehicle, route, traffic and other environmental factors. Frequently, machine learning (ML) algorithms are used to compute the predictions [12]. Because of the high number of influence factors, large amounts of

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

predictive data are required for an accurate prediction. Few researchers have addressed the issue of uncertainty of these predictions [5]. Probabilistic predictions compute probability densities for the target variable, so that uncertainty is directly taken into account. The required predictive data come from different sources in a distributed system, which comprises a network of connected vehicles and backend infrastructure in the cloud. A prediction algorithm utilizing data from this network must guarantee the privacy of the users and be able to function without excessive computation and communication overhead. To this end, we can apply federated learning (FL), which is a ML scheme where each end device learns from local data. A centralized server creates a global model by aggregating the model weights received from the devices at regular intervals [13]. The global model is then sent back to the devices where the learning continues. federated learning (FL) algorithms, such as federated averaging (FedAvg), are typically applied when a large dataset is desired, but sharing data between devices is not possible or too expensive. Recently, an extension of FedAvg with predictive uncertainty was presented, called FedAvg-Gaussian (FedAG). There, uncertainty is introduced in the aggregation step of the algorithm by treating the set of local weights as a posterior distribution for the weights of the global model [14].

This paper presents the application of FedAvg and FedAG to the prediction of the energy demand of a BEV on a planned route. We show an efficient way to learn probabilistic ML models, evaluate and accentuate the advantages of probabilistic EDPs and demonstrate their effect on battery utilization and travel time. The paper is organized as follows: An overview of related work is given in Section II. In Section III, the system architecture and available predictive data are presented. The EDP algorithms and federated learning schemes are described in Section IV and the validation of the prediction is shown in Section V. The benefit in safety margin and travel time is discussed in Section VI before the paper is concluded in Section VII.

II. BACKGROUND AND RELATED WORK

Current practice in energy demand prediction (EDP) is to use information from the vehicle, such as driving speed, acceleration, and historic energy consumption together with predictive information about the planned route from a traffic and routing database (TRDB). TRDB information comprises static map data, e.g., road slope, legal speed limit, and dynamic data such as live traffic. The prediction itself is typically performed using mechanistic models based on physical principles [6], [15]–[17]. In recent years, ML algorithms have been trained to find the relation between the available predictive information and the resulting energy consumption [18]–[20]. The main advantage of ML algorithms is that an exact modeling of the mathematical relation between a feature and the target variable is not necessary, or rather, the ML algorithm automatically creates this model. Additionally, hybrid models, combining a mechanistic model and ML, can be applied [5], [7].

In the context of driving range, energy demand and BEV routing, few articles have addressed predictive uncertainty. Oliva *et al.* describe remaining driving range as a random variable, where the remaining battery energy is estimated with an unscented Kalman filter and the driving profile is predicted with a Markov chain. With that, a probability density function for the remaining driving range is computed [21]. Ondruska and Posner trained linear models to describe the mean and the variance of the energy consumption based on road segment features. Thereby, two deterministic models are used to calculate the parameters of a normal distribution for the prediction of energy consumption [22]. Scheubner *et al.* used a multi-linear regression (MLR) model to compute a stochastic velocity prediction, which is then used to predict a probability distribution for the energy consumption using a physical model and a sequential Monte Carlo simulation [5]. Furthermore, the uncertainty of EDPs has been considered in BEV routing applications [23]–[26].

Data-driven predictions such as with ML algorithms benefit from a rich training dataset [27]. A few articles have proposed sharing data between vehicles and the cloud, so that a user can benefit from the experience of other users, ultimately leading to more accurate predictions. Grubwinkler *et al.* proposed an energetic road map created through crowd-sourcing by collecting information on energy consumption of BEVs while driving a road segment [28]. Tseng and Chau applied the concept of *participatory sensing* to gather crowd-sourced data for the prediction of vehicle energy demand [29]. Straub *et al.* presented another approach for creating an energetic road map, by collecting crowd-sourced driving profiles where the gaps in data coverage were eliminated using ML methods [30].

By applying FedAG to the EDP problem, the advantages of crowd-sourcing can be extended to probabilistic models in an efficient and privacy preserving manner. Recent publications showed the application of FL in vehicle-to-vehicle (V2V) communications [31], in autonomous driving [32], and in traffic flow prediction [33]. To the best of our knowledge, FL has not yet been applied in EDP for BEVs.

III. SYSTEM DESIGN AND DATA

The digital ecosystem in which the EDP operates is a distributed system of connected vehicles and backend infrastructures in the cloud. In this distributed system, large amounts of data can be used to learn ML models, which typically have high computational requirements. The central challenge is to make use of information in the distributed system to enable accurate and robust probabilistic predictions, while considering aspects such as privacy protection and lean communications.

In our previous work [34], we demonstrated the importance of system architecture and module placement for the performance and user experience of driving range prediction and charge planning software. By placing the prediction algorithm parts intelligently across the vehicle and cloud, the performance can be increased. Following that, the prediction algorithm presented in this work can be implemented in an

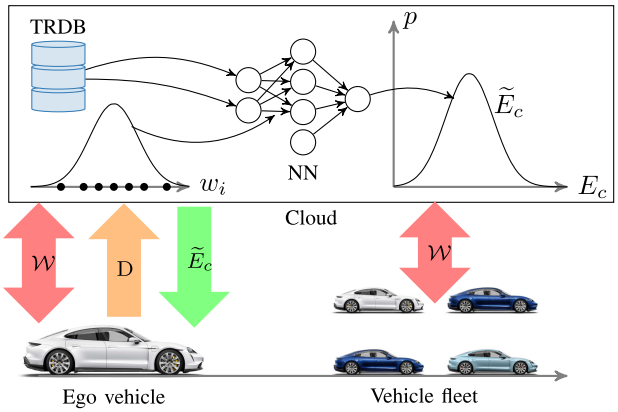

FIG. 1. Schematic overview of the distributed system.

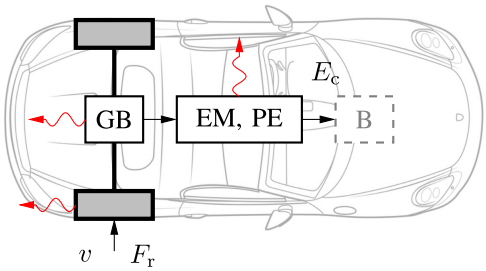

FIG. 2. Powertrain model with input variables v, F_r and output variable E_c. The red arrows indicate simulated component losses.

efficient system architecture. The learning of the models is performed in the vehicle, so that training data remains in the vehicle. Thereby, the communication between the vehicle and the cloud covers only the transfer of the model weights. Furthermore, the predictions are computed in the cloud, so that the transmission of predictive data from the cloud to the vehicle is reduced to the final predictions. In that way, the amount of data transferred between the vehicles and the cloud is minimized. Fig. 1 shows an overview of the distributed system. The ego vehicle and the vehicle fleet share their model weights $\mathcal{W}$ in a central backend in the cloud, where a probabilistic neural network (NN) is built. When a destination D is entered in the ego vehicle's navigation system, the route and predictive information is queried in the TRDB and a probabilistic EDP $\widetilde{E}_c$ is computed with a NN.

A. MEASUREMENT DRIVES AND POWERTRAIN MODEL

In this work, we use a dataset first presented in [5]. The dataset includes 20 real world measurement drives performed by 10 different drivers. All relevant data is logged in the vehicle with a sampling rate of 10 Hz. To generate unified driving data from the pool of measurements with different vehicles, a simulation model for the powertrain of an electric vehicle is used. The simulation model calculates the power P and energy E_c drawn from the battery based on velocity v and driving resistance F_r. Fig. 2 shows a schematic overview of the powertrain model. Based on efficiency maps for components such as the gearbox (GB), electric motor (EM), and power electronics

TABLE 1. Test Run Data

Number of drivers	10
Number of drives	20
Number of segments	1×10^4
Total distance	1896 km
Total duration	22.2 h
Mean velocity	23.7 m s^{-1}

(PE), component losses are computed. These losses are denoted by red arrows in Fig. 2. For a complete description of the model, we refer the reader to [5]. An overview of the test run data is shown in Table 1.

B. MAP AND TRAFFIC DATA

To complement the driving data measured in the vehicles, map and traffic data are acquired to match the driven routes. Using the GPS traces from the measurement drives, the measured data can be matched to a map. Using the IDs of the road segments that form the driven route, the TRDB can be queried to obtain static map data as well as real-time traffic information. The TRDB includes a list of properties such as road slope α, street class Λ, mean traffic speed $\overline{u}$, road curvature κ, legal speed limit v_{lim}, segment length l etc. The TRDB does not only report the mean traffic speed but also information on its distribution, such as standard deviation σ_u and percentile values $P_i(u)$ in steps of 5% [35]. A further aspect of traffic is the traffic phase. The three-phase traffic theory divides traffic into *free flow*, *synchronized flow*, and *wide moving jam* [36]. A method to classify the traffic phase directly in the vehicle was presented in [5]. Using this method, the estimated traffic phase is included in the dataset. Contrary to the measured driving data, map and traffic have a much lower spatial resolution, where a typical segment length is 200 m.

C. VELOCITY PERCENTILE ESTIMATION

An important factor in the energy consumption pattern is the driving speed. In this work, we rely on the velocity reported by the TRDB. As different drivers may exhibit different driving styles and cruise at different speeds in free flowing traffic, we individualize the velocity predictions. To this end, we observe to which percentile of the velocity distribution the driver belongs on a complete trip. By minimizing the squared error between ego vehicle speed and percentile values of the traffic speed distribution, the best matching percentile can be found:

$$\rho_d = \underset{i}{\operatorname{argmin}} \left(v - P_i(u)\right)^2 , \tag{1}$$

where ρ_d is the percentile that best matches driver d, v is the speed of the ego vehicle, and $P_i(u)$ is the i-th percentile of the traffic speed distribution u. As the traffic speed distribution is very narrow in the case of a traffic jam, we only look at synchronized flow and free flow to determine the best fitting percentile.

For each of the drives, (1) is used to find the best fitting percentile. Fig. 3 presents the results of the velocity prediction.

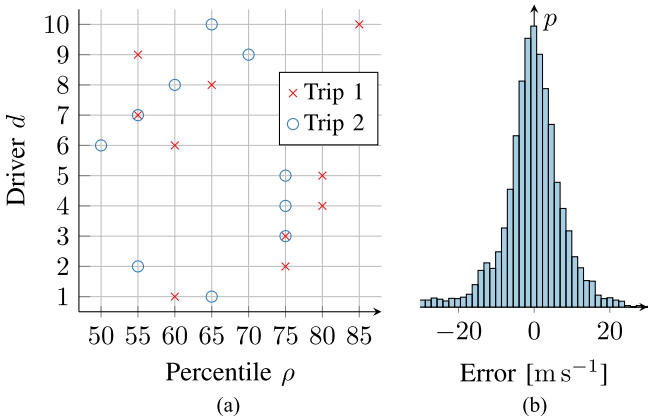

FIG. 3. Velocity percentile observation and velocity prediction error. (a) Observed velocity percentiles of 10 drivers during trip 1 and trip 2. (b) Velocity prediction error for all drivers.

Fig. 3(a) shows the observed velocity percentiles ρ_d for all 10 drivers. The drivers tend to drive faster than the median traffic speed. Most drivers tend to drive consistently, i.e., the velocity percentiles of trips 1 (×) and 2 (○) are close to each other. However, drivers 2 and 10 have significant inconsistencies between trips 1 and 2. Fig. 3(b) shows a histogram of the velocity prediction error $e = v - P_\rho(u)$. The mean value of the error distribution is 0 $\mathrm{ms^{-1}}$ and the prediction is therefore unbiased.

IV. ENERGY DEMAND PREDICTION ALGORITHM

The task of the EDP algorithm is to predict the energy demand for a planned route from start to destination. The route consists of multiple road segments and for each of the segments, the energy demand is predicted based on the features corresponding to the segment. In the probabilistic approach, the EDP algorithm computes a probability density for each of the segments. The total EDP is the sum of the EDPs for the individual segments. The sum of random variables γ and δ is defined as the convolution of their probability density functions:

$$p_{\gamma+\delta}(x) = \int_{-\infty}^{\infty} f_\gamma(y) f_\delta(x-y)\mathrm{d}y = (f_\gamma * f_\delta)(x)\,. \tag{2}$$

For a route with segments $S_1, S_2, \ldots, S_N$ and predictions $\widetilde{E}_{c,1}, \widetilde{E}_{c,2}, \ldots, \widetilde{E}_{c,N}$ the probability density for the total EDP is

$$p_{\widetilde{E}_c}(x) = (p_{\widetilde{E}_{c,1}} * p_{\widetilde{E}_{c,2}} * \ldots * p_{\widetilde{E}_{c,N}})(x)\,. \tag{3}$$

According to the central limit theorem, the sum of independent random variables tends toward a normal distribution and the total EDP is

$$p_{\widetilde{E}_c}(x) = \mathcal{N}(\mu_{\widetilde{E}_c}, \sigma^2_{\widetilde{E}_c})\,, \tag{4}$$

where $\mu_{\widetilde{E}_c}$ is the mean value and $\sigma^2_{\widetilde{E}_c}$ is the variance of the normal distribution [5]. To describe the energy demand on a road segment as a function of the available data, we apply two types of regression models, a linear regression (LR) and a neural network (NN). Both models can be used as probabilistic models with random weights $p(\mathbf{w})$.

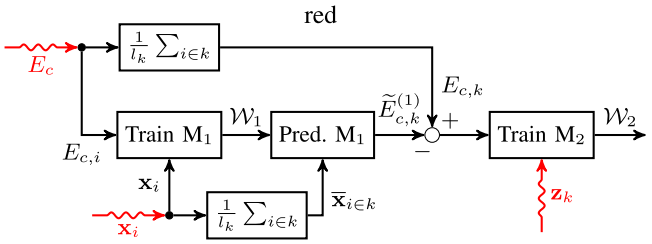

FIG. 4. Block diagram showing a schematic overview of the training process of the two-scale regression model.

The length of the road segments is not uniform. Furthermore, the training data measured in the vehicle is measured with a high sampling frequency (10–100 Hz). Therefore, the data exhibit certain irregularities. To make the most out of the available data, we propose a learning scheme operating on two scales. One part of the model is updated with the sampling frequency of the vehicle measurement data while a second part is updated in accordance to the lower, event based frequency of road segment changes. In the following, the two-scale method and the application of FedAG are presented.

A. TWO-SCALE REGRESSION

To optimally learn the regression model using unstructured data, two regression models are applied. The first model (M_1) is learned continuously with a data stream (10 Hz) to describe the current energy consumption. The second model (M_2) is learned based on the road segments and tries to correct the prediction of the first model. Fig. 4 displays a block diagram of the ML process. $E_{c,i}$ is the vehicle's measured energy consumption at time i and is the target variable for M_1. Feature vector x_i includes the variables measured by the vehicle at time i. Thereby, model weights $\mathcal{W}_1$ are learned. Simultaneously, the mean values of features $\mathbf{x}_i$ on segment k are calculated

$$\overline{\mathbf{x}}_{i\in k} = \frac{1}{l_k}\sum_{i\in k}\mathbf{x}_i l_i, \tag{5}$$

where l_k is the length of segment k and l_i is the distance driven from time $i-1$ to time i. Using the updated weights $\mathcal{W}_1$ and features $\overline{\mathbf{x}}_{i\in k}$, M_1's estimation of the energy consumption on segment k, $\widetilde{E}_{c,k}$ is computed. The difference of the true energy consumption $E_{c,k}$ and the estimation $\widetilde{E}_{c,k}$ delivers the target variable for M_2. Based on the feature vector $\mathbf{z}_k$, weights $\mathcal{W}_2$ are learned. The first model's features $\mathbf{x}_i$ are:

- v vehicle speed,
- α road slope,
- κ road curvature,
- ϕ traffic phase,
- $\overline{u}_h$ historic mean traffic speed,
- $\overline{u}_c$ current mean traffic speed.

The second model's features $\mathbf{z}_k$ are $\overline{\mathbf{x}}_{i\in k}$ and additionally:

Algorithm 1: FedAvg-Gaussian (FedAG). C is the fraction of devices used in each round, K is the total number of devices, $\mathcal{D}_k$ is the data observed by device k, B is the batch size, E is the number of local epochs, η is the learning rate, and l is the squared loss function.

Server executes:
initialize $\mathbf{w}_0$
for each round $t = 1, 2, \ldots$ **do**
  $m \leftarrow \max(C \cdot K, 1)$
  $S_t \leftarrow$ (random set of m clients)
  for each client $k \in S_t$ **do**
    $\mathbf{w}_{t+1}^{(k)} \leftarrow \text{ClientUpdate}(k, p(\mathbf{w}_t|\mathcal{D}))$
  end
  $p(\mathbf{w}_{t+1}|\mathcal{D}) \leftarrow \mathcal{N}(\mu_{\mathcal{D}}(\mathbf{w}_{t+1}^{(k)}), \sigma_{\mathcal{D}}^2(\mathbf{w}_{t+1}^{(k)}))$
  return $p(\mathbf{w}_{t+1}|\mathcal{D})$ to clients
end

ClientUpdate$(k, p(\mathbf{w}|\mathcal{D}))$: *// Run on client k*
$\mathcal{B} \leftarrow$ (split $\mathcal{D}_k$ into batches of size B)
$\mathbf{w} \leftarrow \mathbb{E}(p(\mathbf{w}_t|\mathcal{D}))$
for each local epoch $i = 1$ **to** E **do**
  for batch $b \in \mathcal{B}$ **do**
    $\mathbf{w} \leftarrow \mathbf{w} - \eta \nabla \ell(\mathbf{w}; b)$
  end
end
return $\mathbf{w}$ to server

- $\overline{v}_k - \overline{v}_{k-1}$ segment speed difference,
- $\sigma_v^{(k)}$ segment speed standard deviation,
- l_k segment length.

The predictions step is limited to the road segments, as the predictive data is only reported on that scale. The final EDP is the sum of the predictions computed with M_1 and M_2:

$$\widetilde{E}_{c,k} = \widetilde{E}_{c,k}^{(1)} + \widetilde{E}_{c,k}^{(2)}. \tag{6}$$

B. FEDERATED LEARNING WITH PREDICTIVE UNCERTAINTY

To learn the proposed regression models including predictive uncertainty, we apply FedAG, which is shown in Algorithm 1 [14].

The central part of the algorithm is the aggregation step, where a Gaussian is fitted to the set of client weights $\mathbf{w}$. In this work, the posterior distributions are found by calculating the mean value $\mu_{\mathbf{w}}$ and variance $\sigma_{\mathbf{w}}^2$ of weights $\mathbf{w}^{(k)}$. Subsequently, the posterior distributions for the weights $p(\mathbf{w}|\mathcal{D})$ are returned to the clients. The clients use the expected value $\mu_{\mathbf{w}}$ of the weight posterior distributions for further training, but the predictive distributions are computed with

$$p(y|\mathbf{x}, \mathcal{D}) = \int p(y|\mathbf{x}, \mathcal{D}, \mathbf{w}) p(\mathbf{w}|\mathcal{D}) \mathrm{d}\mathbf{w}\,. \tag{7}$$

Since the integral is typically intractable in non-linear models, Markov chain Monte Carlo (MCMC) is used to compute an approximation. In summary, with FedAG, a probabilistic prediction model can be created in an efficient manner, which benefits from a rich data basis of a vehicle fleet, while minimizing communication overhead and preserving the privacy of the users. In case of an unstable internet connection, a client cannot send and receive updates from the server until a stable connection is restored, i.e., the federated learning becomes *asynchronous* [37]. In this work, we assume a stable connection between the vehicles and the server at all times.

C. FEDERATED LEARNING WITH CLUSTERING

Not all drivers and vehicles exhibit the same driving behavior and energy consumption patterns. Therefore, a single, global model might not be the best choice for the EDP. An alternative is to generate several federated models, each of which acts as a global model for a subset of drivers. A cluster analysis can be executed to divide the set of drivers into subsets. Drivers can then be assigned to these subsets by observing their driving behavior and properties of their vehicles. In this work, we use aggregated data from the drivers to create two driver clusters with *k-means clustering* [38]. The features used in the clustering are:

- observed velocity percentile ρ,
- relative positive acceleration [39],
- relative velocity in free flowing traffic v/v_{lim},
- distribution of observed traffic phases.

The following driver subsets are generated by the cluster analysis:

$$\mathcal{S}_1 = \{2, 4, 7, 8, 9\}\,,$$
$$\mathcal{S}_2 = \{1, 3, 5, 6, 10\}\,.$$

The drivers in $\mathcal{S}_1$ can cooperate in learning one model and the drivers in $\mathcal{S}_2$ learn a separate model. FedAG with clustering is denoted by FedAG-Clustering (FedAGC) with FedAvg-Clustering (FedAvgC) as the deterministic counterpart. With the availability of a larger dataset with more variety, additional features, e.g., the type of vehicle, geographical region, or the distribution of observed temperature, could be included.

V. PREDICTION VALIDATION

To validate the algorithms presented in (IV), the data presented in (III) is used. We apply a leave-one-out cross validation where the scheme depends on the learning algorithm. FL algorithms effectively have access to training data from the entire vehicle fleet, whereas conventional ML algorithms, e.g., stochastic gradient descent (SGD), can typically only access data observed by the respective vehicle. In the following, we validate and compare the learning algorithms FedAG, FedAGC, FedAvg, and conventional driver-individual SGD. The algorithms are applied to a linear regression (LR) and a NN. FedAGC is not applied to the NN, as a NN is able to learn more sophistic dependencies than a linear model and benefits from a larger data basis. The NN has two hidden layers, each containing 50 hidden units. $E = 40$ passes over the available training data are done. In FedAG, $K = 10$ devices denote the 10 drivers, each of which with $C = 1$ and batch size $B = 1$. The training of a the NN is a non-convex optimization and

TABLE 2. Performance Evaluation All Algorithms on All Drives With Mean CRPS and RMSE

	MCRPS [kW h]		RMSE [kW h]	
	LR	NN	LR	NN
SGD	1.2019	1.2078	1.6791	1.5048
FedAvg	0.9594	0.8088	1.1978	**0.9811**
FedAvgC	0.9333	-	1.1578	-
FedAG	0.6750	**0.5628**	1.1978	**0.9811**
FedAGC	0.6594	-	1.1578	-

$t > 1$ rounds are usually required to ensure convergence. We report the results after $t = 5$ rounds, but further rounds do not improve the results significantly. The training of the LR is a convex optimization and no more than $t = 1$ rounds are needed for the training to converge. For FedAG, appropriate precision parameters for the variance of the target variable are estimated using the variance of the training data.

A. PROPER SCORING RULES

To evaluate the performance of the prediction algorithms, proper scoring rules are required. Scoring rules assess the quality of probabilistic predictions by comparing the predictive distribution and the true observation. A scoring rule S is proper if the expected score is optimized by issuing the true distribution of observations as the prediction. In this work, we regard scores as negatively oriented, i.e., a better prediction leads to a lower score. The requirement for a scoring rule S to be proper is thus

$$\mathbb{E}_{y\sim P}\left[S(P, y)\right] \leq \mathbb{E}_{y\sim P}\left[S(Q, y)\right], \tag{8}$$

where y is the true observation of the target variable, P is the true distribution of y and Q is a predictive distribution. The equality in (8) only applies when $Q = P$ [40]. The continuous ranked probability score (CRPS) is a proper scoring rule for density predictions of continuous variables

$$\mathrm{CRPS}(Q, y) = \int_{-\infty}^{\infty} (Q(x) - H(x - y))^2 \mathrm{d}x, \tag{9}$$

where H is the Heaviside step function. CRPS can be directly compared with the mean absolute error (MAE) of deterministic predictions. Futhermore, CRPS is expressed in the unit of the target variable, e.g., [kW h]. In the following, CRPS is used as the main performance indicator in the evaluation of the prediction algorithms.

B. PREDICTION PERFORMANCE EVALUATION

Table 2 shows the mean CRPS (MCRPS) and root mean square error (RMSE) for all algorithms on all drives. Boxplots for distribution of the CRPS on all drives for the algorithms are shown in Fig. 5. The performance of the algorithms in terms of CRPS and RMSE increases with increasing algorithm complexity and the NN trained using FedAG achieves the best performance. For the LR, FedAGC slightly improves the results of FedAG. Generally, the application of FL increases the performance significantly. Finally yet importantly,

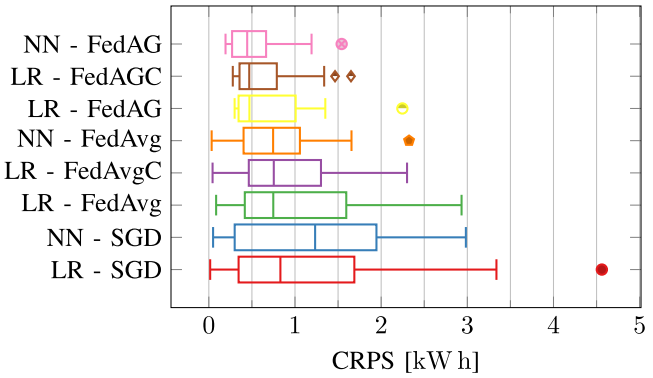

FIG. 5. Boxplots showing the distribution of the CRPS on all test drives for the prediction algorithms.

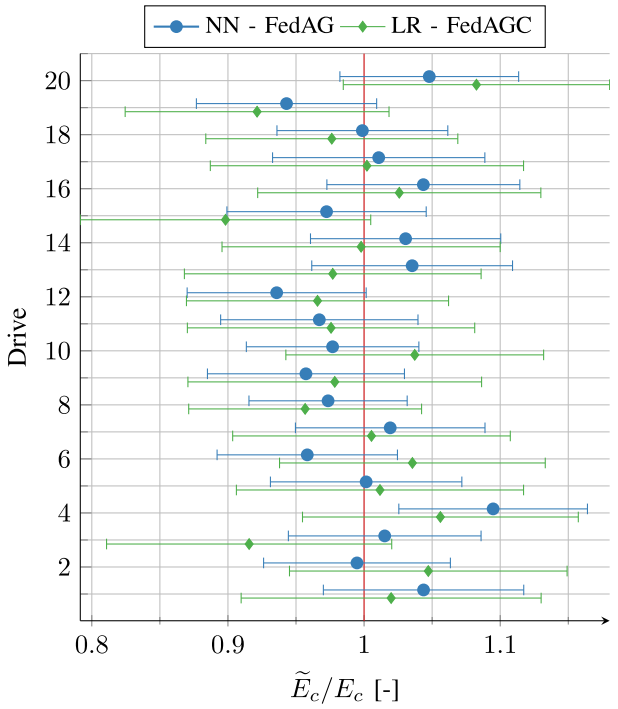

FIG. 6. Mean values and confidence intervals of the predictions computed with a NN and a LR trained using FedAG and FedAGC, respectively, normalized by the true energy consumption.

the probabilistic prediction algorithms achieve a much smaller CRPS than their deterministic counterparts.

A further visualization of the results of the two best performing algorithms is shown in Fig. 6. The figure shows the mean values and 95% confidence intervals of the predictions computed with a NN and a LR trained using FedAG and FedAGC, respectively. The predictions are normalized with the true energy consumption of the respective drive. The observed energy consumption rarely matches the mean value exactly, but falls within the confidence intervals in all drives.

The prediction for an exemplary drive (Nr. 18) using the best algorithm, NN-FedAG is shown in Fig. 7. The green band represents a 95% confidence interval for the accumulated

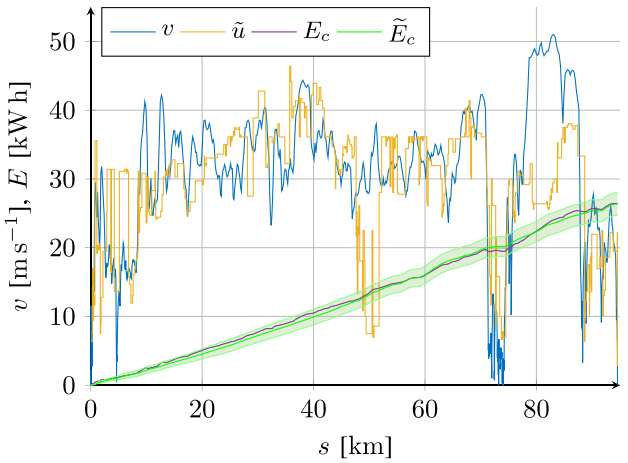

FIG. 7. Predicted and observed velocity profile and predicted and observed accumulated energy consumption of drive 18, computed with NN - FedAG.

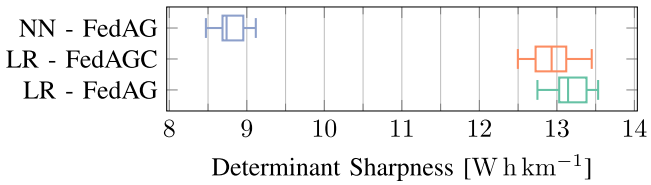

FIG. 8. Boxplots showing the distribution of the DS of the probabilistic EDP algorithms.

energy consumption at each point in the drive. The measured energy consumption is shown in purple. Additionally, the predicted traffic speed percentile value is shown in yellow and the observed driving speed is shown in blue. In Fig. 3(a), driver 9 displayed a moderate inconsistency in driving speed (55th and 70th percentiles). In Fig. 7, the velocity prediction fails to predict high driving speed of up to more than 50 ms^{-1} at around 80 km. Nevertheless, the measured driving speed deviates a little from the predicted velocity and the observed energy consumption always lies within the confidence interval of the prediction.

C. SHARPNESS

The sharpness of a prediction is a measure for the concentration of the predictive distribution. One way to measure sharpness of normally distributed predictions is the determinant sharpness (DS) defined as

$$\mathrm{DS} = \det(\Sigma)^{1/2\,d}\,, \tag{10}$$

where Σ is the covariance matrix of the predictive distribution of dimension $d \times d$. The EDP is univariate ($d = 1$) and the DS therefore reduces to the standard deviation of the predictive distribution. Fig. 8 shows boxplots displaying the distribution of the determinant sharpness of the predictions on all drives for the three probabilistic algorithms. The NN computes significantly sharper predictive distributions than the LRs in all drives. The clustering in FedAGC brings a marginally significant benefit in sharpness compared to FedAG when tested with a two-sample Kolmogorov-Smirnov test.

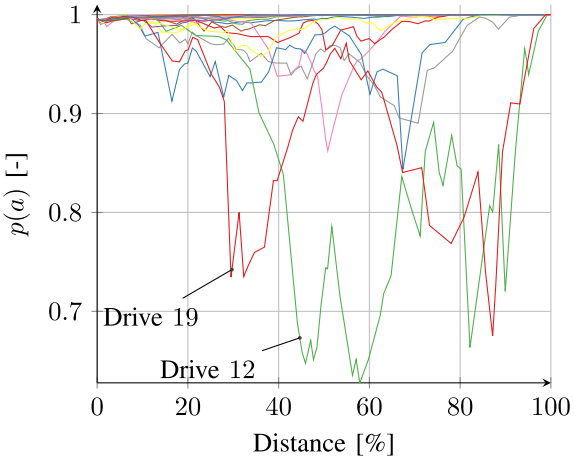

FIG. 9. Destination attainability $p(a)$ over the course of all drives based on the EDP computed with a NN trained using FedAG.

D. DESTINATION ATTAINABILITY

With a probabilistic EDP and a known available battery energy, the probability of reaching a destination, i.e., destination attainability $p(a)$, can be calculated [5]. However, this is not possible with a deterministic EDP. The available battery energy is a variable that cannot be measured directly, but is estimated with some uncertainty [41]. The attainability can thus be calculated with

$$p(a) = p(\widehat{E}_b \geq \widetilde{E}_c) = p(\widehat{E}_b - \widetilde{E}_c \geq 0)\,, \tag{11}$$

where $\widehat{E}_b$ is the estimated available battery energy. Additionally, the amount of energy needed to achieve $p(a) = 0.99$ can be calculated using the inverse of the normal cumulative distribution function Φ:

$$\widetilde{E}_{c,p} = \mu_{\widetilde{E}_c} + \sigma_{\widetilde{E}_c}\Phi^{-1}(p)\,. \tag{12}$$

With (12), the amount of energy to be charged in order to reach a destination can be computed. An important feature of the prediction and attainability estimation is that the destination is ultimately reached. To analyze this, we compute the energy needed for $p(a) = 0.99$ with (12) for each drive, set the initial battery energy $\widehat{E}_b$ to this value and observe the attainability $p(a)$ during the trip. Fig. 9 shows the progression of the destination attainability over the course of all drives. In some drives, the attainability exhibits fluctuation, e.g., in drives 12 and 19, $p(a)$ is significantly lower than 0.99 at times. The gradient of a sharp prediction's cumulative distribution is proportionally large, so that a single maneuver, e.g., strong acceleration during overtaking, can have a significant impact on the attainability. However, the attainability converges to 1 when the destination is approached and the destination is reached in all drives. The linear models trained using FedAG and FedAGC are also able to accurately estimate the attainability.

TABLE 3. Calibration Error Measures for the Destination Attainability With Probabilistic EDP Algorithms

	ECE	MCE	RMSCE
LR-FedAG	0.0426	0.1323	0.0520
LR-FedAGC	0.0322	0.0965	0.0396
NN-FedAG	0.0446	0.1576	0.0576

E. CALIBRATION

The value $p(a)$ can also be called the confidence of the attainability estimation and the observed ratio of drives in which the destination is reached can be denoted as accuracy. If the confidence always matches the accuracy, the prediction is well calibrated [42]. A measure for the calibration of the attainability decision is the difference in expectation between confidence and accuracy

$$\mathbb{E}\left[\left|\mathbb{P}\left(\widetilde{Y}=Y|\hat{P}=p\right)-p\right|\right], \tag{13}$$

where the accuracy term $\mathbb{P}(\widetilde{Y}=Y|\hat{P}=p)$ is the probability of the prediction $\widetilde{Y}$ being equal to observation Y given the estimated confidence $\hat{P}=p$ of the predictor. A perfect calibration, although impossible, is when the expected difference is zero [43]. Using (12) and the observed energy consumption, the accuracy for different p-values can be computed. In our application, accuracy β is the empirical frequency of successful trips given EDP $\widetilde{E}_{c,p}$ and confidence p

$$\beta(p)=\frac{1}{N_D}\sum_j \mathbb{1}\left(\widetilde{E}_{c,p}^{(j)}\geq E_c^{(j)}\right), \tag{14}$$

where N_D is the total number of drives. The expected calibration error (ECE) is defined as mean difference between accuracy and confidence

$$\mathrm{ECE}=\frac{1}{N_p}\sum_i |\beta(p_i)-p_i|, \tag{15}$$

where N_p is the number of confidence levels p tested. The maximum calibration error (MCE) is the maximum difference

$$\mathrm{ECE}=\max_i |\beta(p_i)-p_i|. \tag{16}$$

Finally, the idealized root mean square calibration error (RMSCE) is defined as

$$\mathrm{RMSCE}=\sqrt{\frac{1}{N_p}\sum_i \left(|\beta(p_i)-p_i|^2\right)}. \tag{17}$$

Table 3 shows the ECE, MCE and RMSCE values for the probabilistic prediction algorithms. The LR trained with FedAGC has the lowest calibration errors, followed by the LR and NN trained with FedAG. The ranking of the algorithms is thus not the same as according to the prediction performance in terms of CRPS and RMSE. Fig. 10 shows a reliability diagram visualizing the expected sample accuracy of the attainability estimation as a function of the confidence of the prediction. The black, straight line with slope 1 is the ideal calibration. NN-FedAG tends to be slightly under-confident for $p<0.5$ but slightly over-confident for $p>0.5$. Guo *et al.* discovered that modern NNs are often poorly calibrated [42]. A poorly calibrated prediction can not only lead to a driver being stranded with an empty battery, but also to a significantly higher travel time if the prediction tends to be under-confident. Nonetheless, all three probabilistic EDP algorithms exhibit a sufficient calibration.

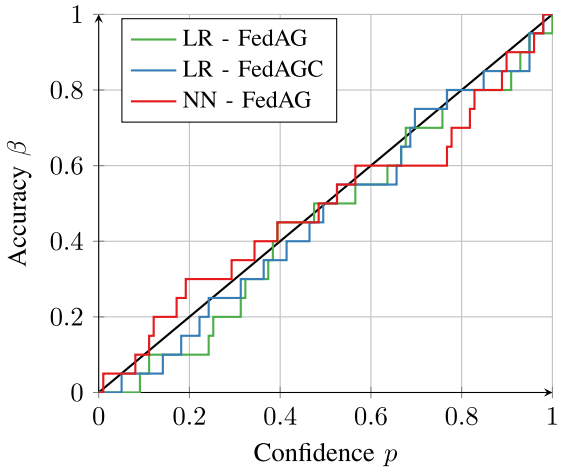

FIG. 10. Reliability diagram for the destination attainability estimation using the probabilistic EDP algorithms.

VI. SAFETY MARGIN AND TRAVEL TIME

A central task of the EDP is to enable certain decision making for attainability and charge planning. The requirement is to predict the energy demand so that a destination can be reached safely without an unnecessary large safety margin Δb_E. A safety margin is the proportion of battery energy reserved in case of an inaccurate prediction. A robust EDP should thus maximize the probability of attaining the destination while minimizing the safety margin, which in turn maximizes the effective driving range of the vehicle. The user primarily experiences how far he can drive without charging and how fast he can travel from A to B. Hence, the user experience is positively influenced by an appropriate safety margin. The safety margin is closely related to the sharpness of the prediction and a sharp prediction leads to a smaller safety margin than a less sharp prediction. In the following, we analyze the safety margins resulting from the EDPs and their impact on travel time.

A. SAFETY MARGIN

With probabilistic predictions, the safety margin can be directly derived from the predictive distribution. The difference between the mean value and the $p=0.99$ value of the predictive distribution can be seen as a safety margin. Using (12), these values can be calculated and the safety margin $\Delta b_E^{(p)}$ is

$$\Delta b_E^{(p)}=1-\frac{\widetilde{E}_{c,0.5}}{\widetilde{E}_{c,0.99}}, \tag{18}$$

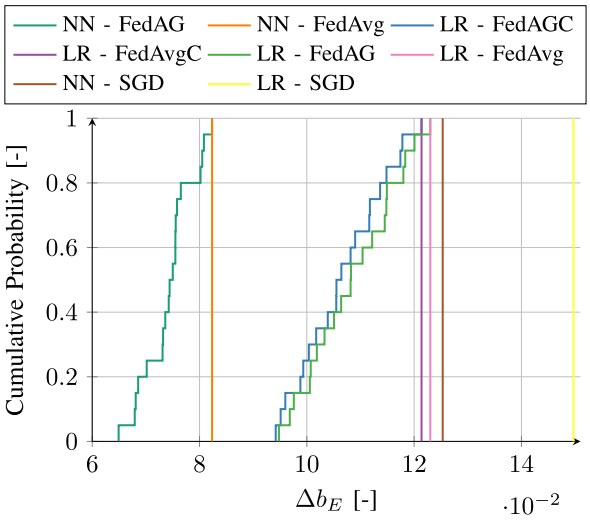

FIG. 11. Empirical cumulative probability distributions of safety margins Δb_E of the EDP algorithms.

TABLE 4. Simulation Data Overview

Number of trips	452
Mean trip distance	438.3 km
Total distance	220 891 km
Total number of charging stops	900
Mean velocity	27.5 m s^{-1}
Battery capacity	90 kW h
Mean consumption	214 W h km^{-1}
Start SoC	50 %

TABLE 5. Safety Margin and Charging Time Results

	Δb_E	Charging time
LR - SGD	0.1497	100 %
NN - SGD	0.1253	95.1 %
LR - FedAvg	0.1230	94.6 %
LR - FedAvgC	0.1214	94.2 %
LR - FedAG	$\mathcal{N}(0.1078, 0.0084^2)$	90.7 %
LR - FedAGC	$\mathcal{N}(0.1066, 0.0080^2)$	90.5 %
NN - FedAvg	0.0823	86.4 %
NN - FedAG	$\mathcal{N}(0.0743, 0.0047^2)$	84.7 %

where the superscripted (p) denotes that the safety margin is based on a probabilistic prediction. A deterministic prediction includes no information about the uncertainty of the prediction and a safety margin can not be derived directly. In a previous publication, we suggested calculating the safety margin $\Delta b_E^{(d)}$ based on the maximum probable error

$$\Delta b_E^{(d)} \geq \left(\frac{\mathrm{d}E_c/\mathrm{d}s}{e_{max}} + 1 \right)^{-1}, \tag{19}$$

where $\mathrm{d}E_c/\mathrm{d}s$ is the mean consumption, e_{max} is the maximum probable error in terms of energy per distance and the superscripted (d) denotes that the safety margin is based on a deterministic prediction. We observe that when (19) is applied to the mean values of the probabilistic EDPs, a safety margin similar to the maximum value of the probabilistic safety margins is found:

$$\Delta b_E^{(d)} \approx \max \left[\Delta b_E^{(p)} \right]. \tag{20}$$

Fig. 11 shows empirical cumulative probability distributions for the resulting safety margins of the EDP algorithms. The ranking of the algorithms is the same as according to CRPS. Predictions with a NN lead to lower safety margins than with a LR and the probabilistic FedAG leads to lower safety margins than FedAvg and SGD. A disadvantage of a constant, deterministic safety margin is that is frequently too large. An unnecessarily large safety margin reduces the effective driving range and reduces the possibilities for feasible routing and charge planning strategies [26].

B. INFLUENCE ON TRAVEL TIME

The safety margin determines the amount of reserved battery energy. The smaller the safety margin, the further a BEV can drive before a charging stop needs to be planned. Thereby, a faster charging point might be attainable. Additionally, a planned charging stop may be shorter, since with a smaller safety margin, the energy needed for the continuation of the trip may be smaller. Additionally, the driving time may be reduced as well, if a more convenient charging point (CP) is attainable with greater effective driving range. The safety margin has therefore a direct influence on charging and travel time. To quantify this, a stochastic framework was developed to analyse the influence of different vehicle parameters [3]. The framework includes the real road and charging infrastructure, in which virtual routes can be defined based on mobility patterns and population data. Using traffic data, speed profiles for the routes are generated. BEV powertrain and battery models are included to allow a calculation of the energy demand for the routes. With route planning, EDP and charge planning, the fastest route is computed. In turn, driving time and charging time can be measured. For a more detailed description of the paper, we refer the reader to [3]. Using this same stochastic framework, we simulate 452 random long-distance trips in Europe and North America and use the total time spent charging as a performance indicator. Table 4 shows an overview of the simulation data.

Table 5 shows the simulated safety margins Δb_E based on Fig. 11 and the resulting total charging time as a percentage of a benchmark algorithm LR-SGD. With probabilistic predictions, the safety margins follow a normal distribution. The results for charging time in Table 5 show that a decreased safety margin Δb_E leads to a decrease in charging time. The advantage of a probabilistic EDP, such as with FedAG, over a deterministic EDP can be seen as well. A further analysis of the charging time benefit of probabilistic EDPs can be seen in Fig. 12, where the distribution of difference in charging time over the complete route collective is shown. In the case of LR, the mean reduction in charging time is approximately 4.7% when predictive uncertainty is considered. For a NN, including predictive uncertainty leads to a mean charging time reduction of 2.3%. The predictions with NNs are generally

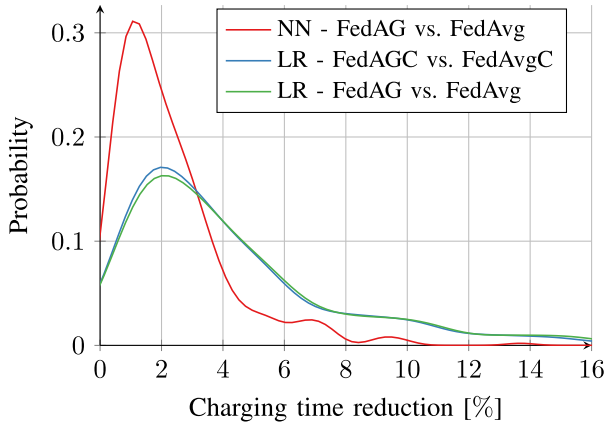

FIG. 12. Proportional charging time reduction between probabilistic and deterministic EDP algorithms.

significantly sharper than those performed with LRs. Furthermore, the variance of sharpness and Δb_E is lower. This leads to the somewhat smaller charging time reduction. Nevertheless, considering predictive uncertainty explicitly improves charging and travel time, especially in regions with sparse charging infrastructure.

VII. CONCLUSIONS AND FURTHER WORK

A network of connected BEVs and backend infrastructure in the cloud constitute a distributed system with various sources of information relevant for the energy demand prediction. By applying federated learning and computing a probabilistic prediction, the uncertainty of the distributed data is considered in a communication efficient and privacy preserving manner. With a multi-scale regression, the prediction models can be trained using data measured in the vehicles while the predictions are computed with data from TRDB directly in the cloud. The energy demand predictions are validated with real driving data and the performance is measured with proper scoring rules. The performance of the probabilistic predictions is superior to conventional deterministic predictions. Furthermore, a non-linear model (NN) achieves higher performance in terms of CRPS and RMSE than a linear model (LR). A probabilistic prediction allows the estimation of destination attainability, i.e., the probability of reaching a destination using the available battery energy. The calibration of this estimation is sufficient and the error between accuracy and confidence is low for all algorithms. A further advantage of an accurate, probabilistic energy demand prediction is the variable safety margin. This leads to a better utilization of the battery energy and increases the effective driving range. Additionally, this translates into a shorter travel and charging time on long distance trips. Our further work includes more research on how global, federated models can be personalized for the participating drivers. The analysis of system design, network usage, and transmission time might prove an important area for future research [44]. Moreover, an ever-expanding database of real driving data can be used to confirm the presented advantages of federated learning and probabilistic models for BEV driving.

REFERENCES

[1] M. Eisel, I. Nastjuk, and L. Kolbe, "Understanding the influence of in-vehicle information systems on range stress—Insights from an electric vehicle field experiment," *Transp. Res. Part F: Traffic Psychol. Behav.*, vol. 43, pp. 199–211, 2016.

[2] T. Franke, I. Neumann, F. Bühler, P. Cocron, and J. Krems, "Experiencing range in an electric vehicle—Understanding psychological barriers," *Appl. Psychol.: Int. Rev.*, vol. 61, no. 3, pp. 368–391, 2012.

[3] A. T. Thorgeirsson, S. Scheubner, S. Fünfgeld, and F. Gauterin, "An investigation into key influence factors for the everyday usability of electric vehicles," *IEEE Open J. Veh. Technol.*, vol. 1, pp. 348–361, Oct. 2020.

[4] Y. Zhang, W. Wang, Y. Kobayashi, and K. Shirai, "Remaining driving range estimation of electric vehicle," in *Proc. IEEE Int. Electric Veh. Conf.*, Greenville, SC, USA, 2012, pp. 1–7.

[5] S. Scheubner, A. Thorgeirsson, M. Vaillant, and F. Gauterin, "A stochastic range estimation algorithm for electric vehicles using traffic phase classification," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6414–6428, Jul. 2019.

[6] 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 Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1404–1418, Feb. 2020.

[7] L. Thibault, G. D. Nunzio, and A. Sciarretta, "A unified approach for electric vehicles range maximization via eco-routing, eco-driving, and energy consumption prediction," *IEEE Trans. Intell. Veh.*, vol. 3, no. 4, pp. 463–475, Dec. 2018.

[8] Z. Yi and P. H. Bauer, "Optimal speed profiles for sustainable driving of electric vehicles," in *Proc. IEEE Veh. Power Propulsion Conf.*, 2015, pp. 1–6.

[9] Z. Yi and P. H. Bauer, "Energy aware driving: Optimal electric vehicle speed profiles for sustainability in transportation," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 1137–1148, Mar. 2019.

[10] A. Fotouhi, N. Shateri, D. S. Laila, and D. J. Auger, "Electric vehicle energy consumption estimation for a fleet management system," *Int. J. Sustain. Transp.*, vol. 15, no. 1, pp. 40–45, 2020.

[11] Z. Yi and P. H. Bauer, "Optimization models for placement of an energy-aware electric vehicle charging infrastructure," *Transp. Res. Part E: Logistics Transp. Rev.*, vol. 91, pp. 227–244, Jul. 2016.

[12] S. Deepak, A. Amarnath, G. KrishnanU, and S. Kochuvila, "Survey on range prediction of electric vehicles," in *Proc. Innov. Power Adv. Comput. Technol. (i-PACT)*, 2019, pp. 1–7.

[13] B. McMahan, E. Moore, D. Ramage, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist.*, ser. Proceedings of Machine Learning Research, A. Singh and J. Zhu, Eds., vol. 54. Fort Lauderdale, FL, USA: PMLR, Apr. 20–22, 2017, pp. 1273–1282.

[14] A. T. Thorgeirsson and F. Gauterin, "Probabilistic predictions with federated learning," *Entropy*, vol. 23, no. 1, Dec. 2020, Art. no. 41.

[15] Z. Yi and P. H. Bauer, "Adaptive multiresolution energy consumption prediction for electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10515–10525, Nov. 2017.

[16] S. Grubwinkler, M. Kugler, and M. Lienkamp, "A system for cloud-based deviation prediction of propulsion energy consumption for EVs," in *Proc. IEEE Int. Conf. Veh. Electron. Saf.*, 2013, pp. 99–104.

[17] A. Jayakumar, F. Ingrosso, G. Rizzoni, J. Meyer, and J. Doering, "Crowd sourced energy estimation in connected vehicles," in *Proc. IEEE Int. Electric Veh. Conf.*, Dec. 2014, pp. 1–8.

[18] S. Sun, J. Zhang, J. Bi, and Y. Wang, "A machine learning method for predicting driving range of battery electric vehicles," *J. Adv. Transp.*, vol. 2019, pp. 1–14, Jan. 2019.

[19] A. Fukushima, T. Yano, S. Imahara, H. Aisu, Y. Shimokawa, and Y. Shibata, "Prediction of energy consumption for new electric vehicle models by machine learning," *IET Intell. Transport Syst.*, vol. 12, no. 9, pp. 1174–1180, 2018.

[20] C. D. Cauwer, W. Verbeke, T. Coosemans, S. Faid, and J. V. Mierlo, "A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions," *Energies*, vol. 10, no. 5, May 2017, Art. no. 608.

[21] J. A. Oliva, C. Weihrauch, and T. Bertram, "Model-based remaining driving range prediction in electric vehicles by using particle filtering and markov chains," in *Proc. IEEE World Electric Veh. Symp. Exhib. (EVS27)*, Barcelona, Spain, 2013, pp. 1–10.

[22] P. Ondruska and I. Posner, "Probabilistic attainability maps: Efficiently predicting driver-specific electric vehicle range," in *Proc. IEEE Intell. Veh. Symp. Proc.*, Dearborn, MI, USA, 2014, pp. 1169–1174.

[23] M. W. Fontana, "Optimal routes for electric vehicles facing uncertainty, congestion, and energy constraints," Ph.D. dissertation, Massachusetts Inst. of Technol., Cambridge, MA, USA, 2013. [Online]. Available: http://hdl.handle.net/1721.1/84715

[24] Z. Yi and P. H. Bauer, "Optimal stochastic eco-routing solutions for electric vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 12, pp. 3807–3817, Dec. 2018.

[25] S. Pelletier, O. Jabali, and G. Laporte, "The electric vehicle routing problem with energy consumption uncertainty," *Transp. Res. Part B: Methodol.*, vol. 126, pp. 225–255, Aug. 2019.

[26] G. Huber, K. Bogenberger, and H. van Lint, "Optimization of charging strategies for battery electric vehicles under uncertainty," *IEEE Trans. Intell. Transp. Syst.*, to be published, doi: 10.1109/TITS.2020.3027625.

[27] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Hoboken, NJ, USA: Wiley, 2012.

[28] S. Grubwinkler, T. Brunner, and M. Lienkamp, "Range prediction for evs via crowd-sourcing," in *Proc. IEEE Veh. Power Propulsion Conf.*, 2014, pp. 1–6.

[29] C.-M. Tseng and C.-K. Chau, "Personalized prediction of vehicle energy consumption based on participatory sensing," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3103–3113, Nov. 2017.

[30] T. Straub, M. Nagy, M. Sidorov, L. Tonetto, M. Frey, and F. Gauterin, "energetic map data imputation: A machine learning approach," *Energies*, vol. 13, no. 4, p. 982, Feb. 2020, Art. no. 982.

[31] S. Samarakoon, M. Bennis, W. Saad, and M. Debbah, "Federated learning for ultra-reliable low-latency V2V communications," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2018, pp. 1–7.

[32] S. R. Pokhrel and J. Choi, "A decentralized federated learning approach for connected autonomous vehicles," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, 2020, pp. 1–6.

[33] Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato, and S. Zhang, "Privacy-preserving traffic flow prediction: A federated learning approach," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7751–7763, Aug. 2020.

[34] A. T. Thorgeirsson, M. Vaillant, S. Scheubner, and F. Gauterin, "Evaluating system architectures for driving range estimation and charge planning for electric vehicles," *Softw.: Pract. Experience*, vol. 51, no. 1, pp. 72–90, 2021.

[35] HERE GlobalB.V, "Traffic API developer's guide," Accessed: Jul. 2020. [Online]. Available: https://developer.here.com/documentation/traffic/dev_guide/topics/api-reference.html

[36] B. S. Kerner, "Three-phase traffic theory and highway capacity," *Phy. A: Stat. Mechanics Appl.*, vol. 333, pp. 379–440, 2004.

[37] M. R. Sprague *et al.*, "Asynchronous federated learning for geospatial applications," in *ECML PKDD 2018 Workshops, ECML PKDD 2018. Communications in Computer and Information Science*, A. Monreale et al., Eds., Dublin, Ireland: Springer, 2019, pp. 21–28.

[38] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[39] E. Ericsson, "Independent driving pattern factors and their influence on fuel-use and exhaust emission factors," *Transp. Res. Part D: Transport Environ.*, vol. 6, no. 5, pp. 325–345, 2001.

[40] T. Gneiting, F. Balabdaoui, and A. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. Roy. Stat. Soc.: Ser. B. (Stat. Methodol.)*, vol. 69, no. 2, pp. 243–268, 2007.

[41] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *J. Power Sources*, vol. 226, no. 1, pp. 272–288, 2013.

[42] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 1321–1330.

[43] M. P. Naeini, G. F. Cooper, and M. Hauskrecht, "Obtaining well calibrated probabilities using Bayesian binning," in *Proc. 29th AAAI Conf. Artif. Intell.*, Austin, TX, USA, 2015, pp. 2901–2907.

[44] K. Bonawitz *et al.*, "Towards federated learning at scale: System design," in *Proc. Mach. Learn. Syst.*, A. Talwalkar, V. Smith, and M. Zaharia, Eds., 2019, pp. 374–388.

ADAM THOR THORGEIRSSON received the B.Sc. degree in mechanical engineering from the University of Iceland, Reykjavik, Iceland, in 2014, and the M.Sc. degree in mechanical engineering from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2017. He is currently a Doctoral Candidate with Dr. Ing. h.c. F. Porsche AG, Ludwigsburg, Germany, and supervised by Frank Gauterin (KIT).

STEFAN SCHEUBNER received the B.Sc. degree in mechanical engineering from the Karlsruhe Institute of Technology, Karlsruhe, Germany, and the M.Sc. degree in mechanical engineering from the Korean Advanced Institute of Science and Technology, Daejeon, South Korea, in 2011 and 2015, respectively. From 2015 to 2020, he worked on energy management functions for BEVs with Porsche AG, Stuttgart, Germany. Since then, he has been concerned with DC fast-charging infrastructure for BEVs with EnBW AG, Karlsruhe, Germany.

SEBASTIAN FÜNFGELD received the M.Sc. degree in mechanical engineering from the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2014. Since 2014, he has been with Dr. Ing. h.c. F. Porsche AG, Ludwigsburg, Germany. His research interests include intelligent control and predictive systems, including modeling of driver and traffic behavior, stochastic forecasting, and stochastic optimization.

FRANK GAUTERIN received the Diploma degree in physics from the University of Münster, Münster, Germany, in 1989, and the Dr. rer. nat. degree (Ph.D.) in physics from the University of Oldenburg, Oldenburg, Germany, in 1994. From 1989 to 2006, he was in different R & D positions with Continental AG, Hanover, Germany, leaving as the Director of NVH Engineering (noise, vibration, harshness). Since 2006, he has been a Full Professor with the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. He is currently the Head with the Institute of Vehicle System Technology, KIT, and a Scientific Spokesperson with KIT Center Mobility Systems. His research interests include vehicle control, vehicle dynamics, vehicle NVH, vehicle suspension, tire dynamics and tire-road-interaction, vehicle concepts, vehicle modeling, and identification methods.