Driver- and Situation-Specific Impact Factors for the Energy Prediction of EVs Based on Crowd-Sourced Speed Profiles

Stefan Grubwinkler[1], Martin Hirschvogel[1] and Markus Lienkamp[1]

Abstract— This paper presents a system for the prediction of the necessary energy for selected trips of electric vehicles (EVs), which can be used for various EV assistants like range estimation. We use statistical features extracted from crowd-sourced speed profiles for the energy prediction, since they consider the varying impact factors of the individual driving style and the prevailing traffic condition. A statistical prediction model uses these features in order to predict the deviation from the mean energy consumption of the EV. Hence, the model predicts the variance of energy consumption caused for example by individual driving behavior. The results show an improvement of the energy prediction by 5.4 percentage points if the statistical features are considered. The prediction of the propulsion energy for EVs before the start of a given route has a relative mean error of 6.8%.

I. INTRODUCTION AND MOTIVATION

The limited driving range and long recharging times are two major barriers against the market penetration of electric vehicles (EVs). A reliable prediction of the necessary energy along specified routes is an important information to reduce the drivers' range anxiety [1]. Thus, various applications for EVs depend on an accurate energy prediction: eco-routing [2], [3], the calculation of the residual range [4], travel and journey planning [5], information for charging service providers [6] or advanced driver assistance systems for minimizing the energy consumption along the route [7]. The energy consumption of EVs consists of two parts:

- power train energy consumption depending on vehicle parameters and the speed profile along the route
- energy consumption of the auxiliaries depending on weather conditions (e.g. heating), status of the electric devices and the expected travel time

The vehicle parameters, which are known in advance or are estimated during the trip, mainly influence the level of energy consumption, whereas the speed profile especially influences the variance around the average energy consumption [8]. The uncertainty factor to be considered most for the estimation of auxiliary energy consumption is the travel time, since weather conditions usually are predictable for a short forecast period. Hence, the knowledge of the future speed profile along a route is essential for an accurate energy prediction. The prediction is a hurdle since various impact factors such as vehicle parameters (e.g. maximum acceleration), route (e.g. traffic signs), traffic flow and the individual driving behavior influence the future speed profile.

[1]Stefan Grubwinkler, Martin Hirschvogel and Markus Lienkamp are with the Institute of Automotive Technology, Technische Universitaet Muenchen, 85748 Garching b. Muenchen, Germany, `grubwinkler@ftm.mw.tum.de`

In the next years, the amount of vehicles being connected to a server in the back-end will increase. Thus we use a cloud-based approach to collect the speed profiles from different vehicles and drivers. Prediction models use the stored statistical features being extracted from the previous collected speed profiles to predict the variance of the energy consumption along given routes before the start of a trip.

This paper is organized as follows. Section II gives an overview about related work. Section III describes the used data, which are analyzed for their influence on traffic conditions or individual driving behavior in Section IV. Section V presents the statistical prediction model and an evaluation.

II. RELATED WORK

The impact of individual driving behavior on speed profiles and on the energy consumption of vehicles has been analyzed by several studies. In [10] and in [11] the future speed profile was predicted for only a limited horizon of a few kilometers by using real time infrastructure information or driver characteristics. [12] and [13] used a cloud-based framework to calculate the optimal speed profile for a limited prediction horizon in the back end to minimize the fuel consumption of Plug-In Hybrid Vehicles (PHEV).

In [14] and [15] methods for the characterization of the driving style are introduced. The authors used statistical features of speed profiles in regression models or clustering algorithms in order to calculate an aggregated indicator and significant clusters related to the individual driving style.

Several studies consider the individual driving behavior depending on the occurring maneuver situation. In [16] statistical features of speed profiles were used for the discrimination of driving conditions. Based on the determined conditions, the energy management strategy of Hybrid Electric Vehicles is adapted. The authors of [17] used defined maneuver classes along the navigation route for the consideration of the driver-specific influence. A self-learning algorithm adapts the parameters of the maneuver classes (e.g. average velocity), if the driver-specific values differ from the predefined ones. In [18] the complete future speed profile along a specified route is predicted by the usage of map attributes, whereas individual driving patterns are not considered.

III. OVERVIEW OF THE ENERGY PREDICTION SYSTEM

The proposed energy prediction system consists of a vehicle-specific part and a cloud-based part (see Fig. 1). The crowd-sourced information has to be used for the prediction of various types of EVs with different vehicle
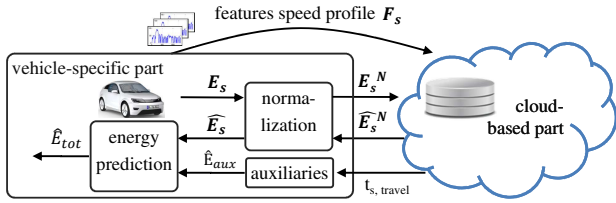
Fig. 1. Overview energy prediction system

Fig. 2. Normalization propulsion energy

parameters. Extracted statistical features from speed profiles as for example the mean acceleration or the mean speed are used for the cloud-based part, since they are comparable within certain limits. The propulsion energy $\mathbf{E_s}$ of a road segment s has to be normalized ($\mathbf{E_s^N}$) for the cloud-based prediction system. The cloud-based part predicts a normalized propulsion energy value $\widehat{\mathbf{E}}_\mathbf{s}^\mathbf{N}$, which can be transformed according to the vehicle-specific energy consumption into $\widehat{\mathbf{E}}_\mathbf{s}$ depending on the vehicle parameters.

The regenerative braking strategy varies even for the same type of vehicle and can often be changed by the driver himself (e.g. choose of eco-mode). Hence, we separate the powertrain energy in acceleration mode E_{pos} from the one by using the regenerative braking mode E_{neg}. The prediction of the energy consumption of the auxiliaries $\widehat{E}_{aux}$ is also vehicle-specific, but is not considered in this paper ($\widehat{E}_{aux} = 0$). The energy for n segments of a route can be calculated:

$$\widehat{E}_{tot} = \sum_{s=1}^{n} \widehat{E}_{s,pos} + \widehat{E}_{s,neg} + \widehat{E}_{s,aux} \tag{1}$$

A. Vehicle-specific prediction system

In the vehicle-specific part of the prediction system the vehicle parameters are considered. We use a characteristic power consumption map (PCM) of the drivetrain to describe the necessary power to overcome the driving resistances. The PCM takes into account all the necessary power including the drivetrain efficiency at the various operating points. A recursive least square algorithm adapts the PCM during the trip according to occurring parameter changes [19].

The occurrence of operating points (vehicle speed and acceleration) depends on the driving condition (e.g. urban road, highway). We defined twelve road classes according to the speed limits and road types. A probability density function (PDF_k) exists for the frequency distribution of vehicle speed and acceleration of every road class k. PCM and PDF_k are used to calculate a mean powertrain energy consumption $\varnothing \mathbf{E_k}$ for k (see Fig. 2). A detailed description is given in [19]. $\varnothing \mathbf{E_k}$ depends only on the vehicle parameters and is independent from the prevailing acceleration and speed values along a certain route, since we use a fixed PDF_k for every road class. As result, the vehicle specific impact factors as for e.g. the additional load are independent from the ones related to the speed trajectory, which are predicted separately by the cloud-based part of the system.

$\varnothing \mathbf{E_k}$ is due to the fixed probability density function PDF_k independent from the speed profile in a segment s, so that $\varnothing \mathbf{E_k}$ can be used for normalization of the measured
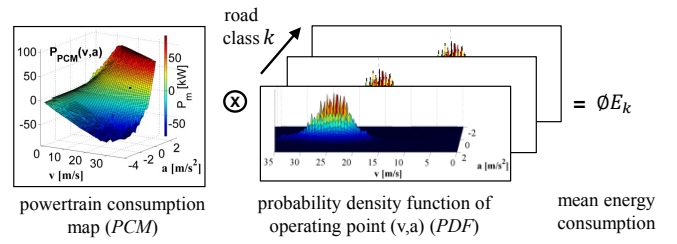
energy consumption of the powertrain $\mathbf{E_s}$ in s (a detailed description is given in [20]). τ is a fixed parameter for each type of vehicle to increase the accuracy of the normalization depending on the vehicle parameters.

$$\mathbf{E_s^N} = \frac{\mathbf{E_s}}{\tau * \varnothing \mathbf{E_k}} \tag{2}$$

B. Cloud-based prediction system

We divide the road network in fixed segments s for the calculation of the statistical features from the collected speed profiles. The segments s are longer than the ones in navigatable maps, because a minimum segment length of $150s$ of travel time is necessary for the extraction of appropriate statistical features [16]. The road network is segmented according to speed limits and road types. Hence, every segment s consists of road sections belonging to the same road class k. In [20] the criteria for the segmentation of the road network are described.

The normalized energy consumption $\mathbf{E_s^N}$ for every segment s can be divided in the mean value $\varnothing \mathbf{E_s^N}$ and the corresponding deviation $\mathbf{\Delta E_s}$ from the mean value:

$$\mathbf{E_s^N} = \varnothing \mathbf{E_s^N} * (1 + \mathbf{\Delta E_s}) \tag{3}$$

Segment-specific impact factors like traffic signs, speed limits or road curvature are considered in the mean value of $\varnothing \mathbf{E_s^N}$, since this value is updated each time the segment is traversed. According to (2), the mean energy consumption on a segment s is higher than on an average segment of the corresponding road class k, if the value for $\varnothing \mathbf{E_s^N}$ is higher than 1 (see Fig. 3). The deviation $\mathbf{\Delta E_s}$ from the normalized mean value $\varnothing \mathbf{E_s^N}$ of s takes into account the variation of the speed profile from the average speed profile in s. $\mathbf{\Delta E_s}$ is comparable for various segments, which is necessary for cloud-based systems.
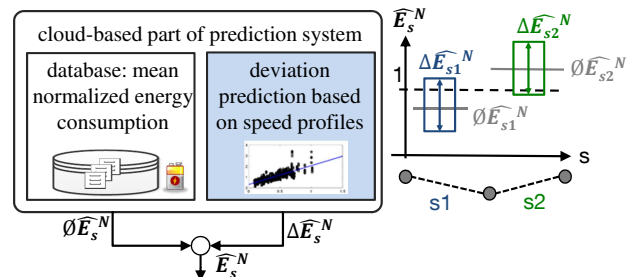

Fig. 3. Overview cloud-based part of the energy prediction system

C. Energy prediction

The system provides three prediction values, which can be converted into one another:

1) $\widehat{E}_{tot,\varnothing k}$: based on the mean energy per road class $\varnothing \widehat{\mathbf{E}}_{\mathbf{k}}$
2) $\widehat{E}_{tot,\varnothing s}$: based on the normalized mean energy per segment $\varnothing \widehat{\mathbf{E}}_{\mathbf{s}}^{\mathbf{N}}$
3) $\widehat{E}_{tot,\Delta k}$: based on the deviation from the segment average value $\mathbf{\Delta}\widehat{\mathbf{E}}_{\mathbf{s}}^{\mathbf{N}}$

The number of considered impact factors and the accuracy increases with the order in the list above. However, the complexity and the necessary information also rises. The three-step modular system can be used for a prediction with lower accuracy in case only a part of the necessary inputs for the corresponding prediction model is available.

In this paper, we introduce a statistical model to predict $\mathbf{\Delta}\widehat{\mathbf{E}}_{\mathbf{s}}^{\mathbf{N}}$, which considers the influence of traffic, driving behavior and further impact factors. The prediction methods ($\widehat{E}_{tot,\varnothing k}$ and $\widehat{E}_{tot,\varnothing s}$) are used for evaluation purposes.

IV. COLLECTED DATA

In this section, we give an overview about the collected speed profiles. Since we do not have access to real-time traffic information systems, the collected speed profiles have to be classified afterwards in order to get information about the prevailing traffic condition at the time of the recording.

A. Speed Profiles

$16,000$ tracks with a length of more than $300,000\ km$ of real world speed profiles have been collected in the urban area of Munich. GPS position and speed are sampled at one second timestamp and transmitted to the database in the back-end. We match the collected speed profiles with an OpenStreetMap (OSM) map [21], so that existing map attributes can be used and the speed profiles can be partitioned according to defined fixed segments s. The consideration of various driving styles and traffic conditions is guaranteed by the usage of a huge amount of real world speed profiles.

We use a simulation model consisting of validated component models (e.g. battery, electric drive) to calculate the energy consumption of different EVs according to the recorded speed profiles. Significant vehicle parameters can be varied.

B. Traffic Information

We apply a simple and efficient spatio-temporal method for the segment-by-segment identification of traffic conditions, which is described in [22]. This algorithm requires only GPS-data and achieves accurate results after 15-20 existing vehicle trajectories without a continuous data stream.

The method compares the temporal mean speed with the spatial mean speed in a segment s. Thresholds for the temporal and spatial speed cluster the vehicle trajectories in four quadrants (see Fig. 4). The temporal threshold considers the 5th percentile of traversal time and the average red light duration on s. The locations of the traffic lights and the corresponding red light durations are taken from the OSM-data or are estimated via crowd-sourcing from the recorded vehicle trajectories. In [22] the spatial threshold is defined
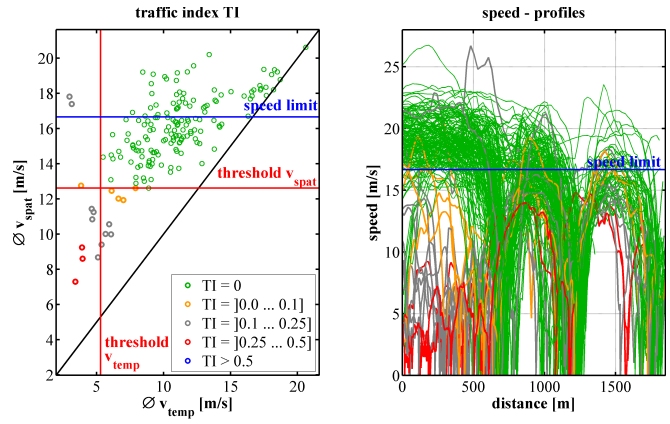

Fig. 4. Traffic index for a segment

as the 5th percentile of spatial mean speed in the right-sided subspace. Since we also use the method in non-urban areas with hardly any stops, we additionally use the speed limit and the 85th percentile of the mean spatial speed in s for the definition of the threshold.

A traffic index TI is calculated for every vehicle trajectory of s depending on the distance from the intersection point of the two thresholds, whereby all vehicle trajectories in the right upper quadrant have the value 0 (free-flowing status). TI is an indicator for the approximate traffic congestion. A comparison of the calculated TI and the corresponding speed profiles on a chosen segment is shown in Fig. 4. A detailed description and evaluation is given in [22]. We observed in our own analysis on the plausibility, that the calculated TI can be used for the development of our prediction system.

C. Data acquisition for deviation prediction

In [9], [14], [16] and [23] the influence of various statistical features of speed profiles on fuel consumption is shown. We use some of these features (see Table I), since they can be

TABLE I
STATISTICAL FEATURES

f_i	explanation
$a_\%$	part of segment s in acceleration phase
$d_\%$	part of segment s in deceleration phase
$c_\%$	part of segment s in cruising phase
$a_m * a_\%$	mean acceleration combined with acceleration part
$d_m * d_\%$	mean deceleration combined with deceleration part
a_{std}	standard deviation of acceleration
$stop_{cnt}$	number of stops in segment s
$v_{m,run}$	mean run velocity
$v_{m,spat}$	mean spatial velocity
$v_{std,run}$	standard deviation of $v_{m,run}$
PKE	positive kinetic energy related to segment s
RPA	relative positive acceleration related to segment s
RNA	relative negative acceleration related to segment s
$a_{m,low}, d_{m,low}$	mean accel. / decel. at low speed difference
$a_{m,mid}, d_{m,mid}$	mean accel. / decel. at medium speed difference
$a_{m,high}, d_{m,high}$	mean accel. / decel. at high speed difference

Fig. 5. Collected data for each road segment

Fig. 7. Distribution of principal components of a segment

used for statistical or machine-learning approaches to predict $\Delta\widehat{\mathbf{E}}_\mathbf{s}^\mathbf{N}$. We have evaluated these features concerning their correlation to $\Delta\widehat{\mathbf{E}}_\mathbf{s}^\mathbf{N}$ in [20].

The statistical features $\mathbf{F_s}$ are extracted from the recorded speed profile and the consumed or regenerated energy is normalized with $\mathbf{E_k}$ of the corresponding road class k after an EV has passed a segment s. We only store the features in the database in case they have been recorded with free-flowing traffic. The database provides the mean value and the standard deviation of the collected data of different vehicles ($\varnothing\mathbf{F_s}, \sigma\mathbf{F_s}, \varnothing\mathbf{E_s}, \sigma\mathbf{E_s}$) for each segment (see Fig. 5). It is important for the deviation prediction, that the statistics of $\mathbf{F_s}$ and $\mathbf{E_s^N}$ are updated simultaneously. Map features like traffic lights or the road classes are also available and in future real-time traffic information will be added.

V. DATA ANALYSIS

We analyze the impact of possible factors on $\Delta\widehat{\mathbf{E}}_\mathbf{s}^\mathbf{N}$ as we compare the extracted statistical features $\mathbf{F_s}$ with the corresponding mean values $\varnothing\mathbf{F_s}$ of the segment s:

$$\Delta\mathbf{F_s} = \mathbf{F_s} - \varnothing\mathbf{F_s} \tag{4}$$

The number of features has to be reduced for an efficient data analysis due to the length of the vector $\Delta\mathbf{F_s}$. Some of the features correlate with each other. For example, if the part of acceleration is above the average ($\Delta a_\% > 0$), then probably $\Delta d_\%$ or $\Delta v_{m,run}$ are below the average. Similar to the approach in [24], a principal component analysis (PCA) is applied. The PCA transforms the correlated elements of vector $\Delta\mathbf{F_s}$ into a smaller number of uncorrelated variables $\Delta\mathbf{PC_s}$, which are called principal components.

$$\Delta\mathbf{PC_s} = \text{PCA}(\Delta\mathbf{F_s}) = [\Delta pc_{1,s}, ..., \Delta pc_{n,s}]^T \tag{5}$$

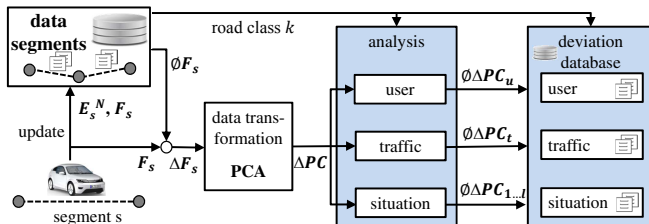The results of the PCA show that the first seven principal components represent 95% of the information of $\Delta\mathbf{F_s}$. Thus,

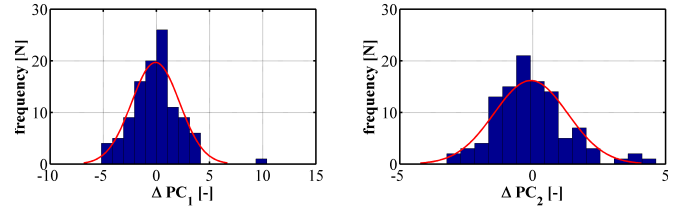

Fig. 6. Deviation database for various impact factors

the 16 elements of $\mathbf{F_s}$ can be reduced to $\Delta\mathbf{PC_s}$ with seven elements. The principal components of each segment are approximately normal distributed around the mean value, which is usually zero, since the mean value of $\Delta\mathbf{PC_s}$ of s is usually zero (see Fig. 7). We distinguish between following impact factors for the analysis of $\Delta\mathbf{PC}$:

- traffic condition
- driving-behavior
- contextual situations (daytime, weather condition, etc.)

The elements of $\Delta\mathbf{PC_s}$ are comparable, so that we can cluster several segments for the analysis. Usually we cluster the data according the road classes k. The mean deviation of the principal components $\varnothing\Delta\mathbf{PC}$ for every cluster is calculated and the results of the analysis are used for the energy prediction.

In our approach, a driver-specific model always implies the usage of the same car; in case the user drives a different vehicle another user-specific model and database is necessary. For this reason we do not have to consider the vehicle-related influence in the speed profiles (e.g. it does not matter if the mean acceleration is lower due to the driving style or due to the limited acceleration power of the vehicle).

The necessary user-, traffic- or situation-specific data ($\Delta\mathbf{PC_u}$, $\Delta\mathbf{PC_t}$ or $\Delta\mathbf{PC_{1,...l}}$) can either be stored in the back-end or on an electronic device in the vehicle (especially driver-specific data). Possible system architectures are not part of this paper and have to be discussed in future research.

A. Analysis traffic condition

The values of the principal components $\Delta\mathbf{PC_s}$ are analyzed for an exemplary chosen segment (motorway with

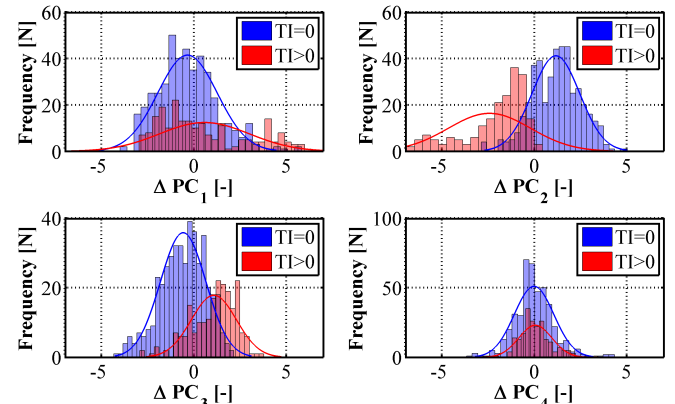
Fig. 8. Analysis traffic condition of 1 segment

speed limit $120\frac{km}{h}$) considering the calculated traffic index. We distinguish only between free-flowing ($TI = 0$) and non-free flowing status ($TI > 0$). The distribution of the first four principal components is given in Fig. 8. The reason for the lower values of Δpc_2 is the high impact of the speed on Δpc_2 which is usually lower during traffic congestion ($TI > 0$). Δpc_3 also varies, whereas the mean value of the approximated normal distribution of Δpc_4 is equal in both cases. The differences of $\mathbf{\Delta PC_s}$ between the two traffic conditions depend also on the attributes of the segment and on the kind of traffic congestion.

B. Analysis of driver-specific impact

The data sets $\mathbf{\Delta PC}$ are clustered according to the existing road classes k for the analysis of the individual driving behavior. We only use data recorded in the free-flowing traffic state ($TI = 0$) for comparability reasons. As example, we compare the mean value of two principal components of several drivers on segments of a certain road class (motorways with speed-limit between $100\frac{km}{h}$ and $120\frac{km}{h}$). We group the drivers by their average deviation of energy consumption from the mean value. Fig. 9 shows the relation between ΔE_{pos}^N and the mean value of $\varnothing\Delta pc_1$ or $\varnothing\Delta pc_2$. However, the variance of $\varnothing\Delta pc_1$ or $\varnothing\Delta pc_2$ is for most of the drivers equal to the variance of all drivers. Only driver 17, 18 or 19 have a smaller variance than the average.

We analyzed the data from the participants of the fleet-test e-Flott [25], in which about 20 participants drove both an AUDI A1-TDI (ICE) an AUDI A1-etron (PHEV) for several months. The two vehicle types have a different acceleration behavior. Fig. 10 shows the mean values of two principal components, which have been calculated of $\mathbf{\Delta F_s}$ of all segments in a certain road class. The values for Δpc_2 of the drives with the A1-TDI are for all drivers higher than the
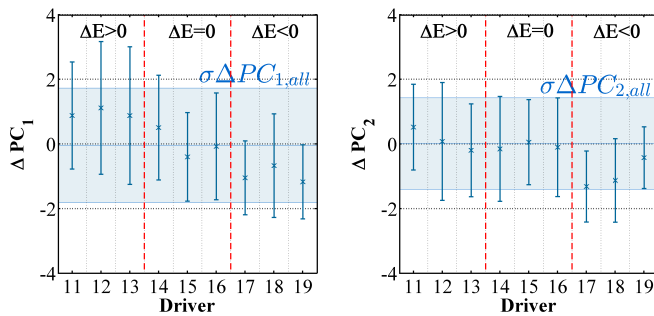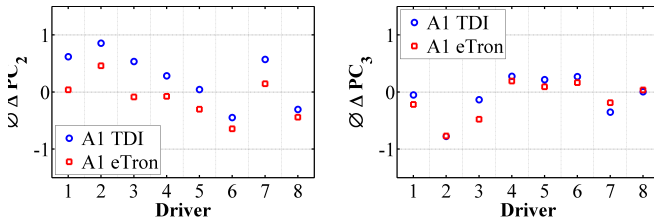
ones with the other vehicle. The driver-specific impact on Δpc_2 is still visible, since for example the values for driver 6 and 8 are lower than the average of all drivers ($\Delta pc_2 = 0$), whereas those of 2 and 7 are higher. Δpc_3 is independent of the type of vehicle. The mean values do not vary despite the usage of two different vehicles.

The results of the analysis show that the mean values of the principal components $\varnothing\mathbf{\Delta PC_u}$, which are calculated of the deviation of the extracted features $\mathbf{\Delta F_s}$ from the mean values of each road segment, can be used for the prediction of $\mathbf{\Delta\widehat{E}_s^N}$. $\varnothing\mathbf{\Delta PC_u}$ of every road class k is stored for every driver in a user-specific database.

C. Analysis of situation-specific impact factors

The impact of further situation-specific attributes on the distribution of the principal components is analyzed. Map attributes are not analyzed, since they are considered already in $\varnothing\mathbf{E_s^N}$. If possible, the data is clustered in several groups (e.g. road classes), so that the impact of the situation-specific factors on $\mathbf{\Delta PC}$ can be described in a general manner. All situation-specific factors are described with categorical variables (e.g. daylight / no daylight). We perform an one-way ANOVA-test at the 5% level if there is a significant difference of $\varnothing\mathbf{\Delta PC}$ for the two categories. We have identified the following situation-specific impact factors, which are at least significant for two principal components:

- daylight
- above-average stop likelihood at traffic lights
- rainfall
- temperature $< 5°C$
- time window (peak traffic hours, non-peak hours on weekdays, weekends, night-hours)

The light condition influence the driving behavior [26] and is exemplary analyzed in this paper. Fig. 11 shows the mean values of six principal components depending on daylight or darkness. For comparability, only measurements between $6pm$ and $10pm$ and under free-flowing traffic condition ($TI = 0$) are considered. Especially $\varnothing\Delta pc_1$ shows a significant difference between the two categories. We have analyzed this effect for various user-groups, which used the different type of vehicles. The difference of $\varnothing\Delta pc_1$ between daylight and darkness lies for all user-groups between 0.4 and 0.5 despite the variation caused by the usage of different vehicle types. We use the recognized deviation caused by the condition of daylight $\varnothing\mathbf{\Delta PC_i}$ for the prediction model.

Fig. 9.   Analysis driving behavior

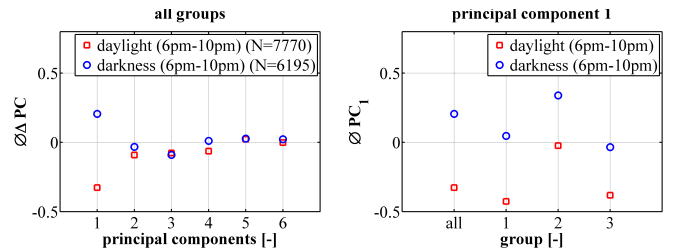Fig. 10.   Impact of different vehicles types

Fig. 11.   Impact of daylight on principal components

VI. Energy Prediction

We develop a prediction model based on the results of the principal components, which describe the deviation from the mean values of the collected features of each segment. The accuracy of the energy prediction is evaluated both segment-by-segment and for complete trips.

A. Statistical prediction model

In [20] a regression model is introduced, which uses most of the elements of the vector $\mathbf{\Delta F_s}$ to predict the energy deviation $\mathbf{\Delta \widehat{E}_s^N}$. If the statistical features $(\mathbf{\Delta \widehat{F}_s})$ can be predicted, $\mathbf{\Delta \widehat{E}_s^N}$ can be predicted precisely, since the accuracy of the regression model is high (for e.g. the absolute mean error is around 6% for $\Delta E_{s,pos}^N$). A new regression model $\mathbf{M_{reg}}$ can be defined, which uses the first seven principal components $\mathbf{\Delta PC}$:

$$\mathbf{\Delta \widehat{E}_s^N} = \mathbf{M_{reg}}(\mathbf{\Delta \widehat{PC}}) \tag{6}$$

Map features $\mathbf{M}$ such as turn-lefts or the gradient of the road are additionally used as already described in [20]. The accuracy of the new regression model decreases only by 1%. The prediction of the principal components $\mathbf{\Delta \widehat{PC}}$ is necessary as input for the regression model. We use as prediction for $\mathbf{\Delta \widehat{PC}}$ the mean values of the stored clusters of the different impact factors ($\varnothing \mathbf{\Delta PC_u}$, $\varnothing \mathbf{\Delta PC_t}$ and $\varnothing \mathbf{\Delta PC_{1,...,l}}$).

The impact of the user-specific deviation is bigger compared to the situation-specific deviation. The estimated driver-specific deviation $\mathbf{\Delta \widehat{PC}_u}$ is the basis for the realized statistical model, since the user-specific database is updated only under free-flowing traffic condition like $\varnothing \mathbf{F_s}$ and $\varnothing \mathbf{E_s^N}$. The impact of traffic congestion is considered additionally. The weighting factors α_u and α_t are used to weight the traffic and user-specific impact (see Fig. 12). α_u for the user-specific deviation $\mathbf{\Delta \widehat{PC}_u}$ decreases with increasing traffic congestion. The expected deviations $\mathbf{\Delta \widehat{PC}_i}$ of further situation-specific impact factors like the weather condition are taken into account additionally to the driver-specific impact. The corresponding weighting factors β_i depend on the frequency of the situation i in the calculation of $\varnothing \mathbf{F_s}$ and $\varnothing \mathbf{E_s^N}$.
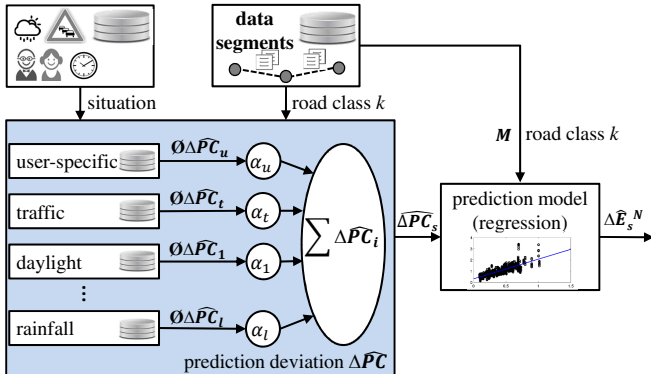

Fig. 12.   Prediction model

B. Evaluation methodology

The described energy prediction model is evaluated based on the segments of 1000 randomly chosen tracks. We have only chosen trips with a minimum length of $20km$ for the evaluation, so that short trips consisting of only a few segments are not considered. The consumption of the auxiliaries depend extremely on weather conditions and is not taken into account, since we consider only impact factors concerning the speed profile in the proposed prediction model. We evaluate the prediction accuracy at the beginning of the trip, since most of the applications for EVs require a prediction in advance.

The energy prediction $\widehat{E}_{tot,\Delta s}$ is compared with the mean energy per road class $\widehat{E}_{tot,\varnothing k}$ and the one based on the mean energy per segment $\widehat{E}_{tot,\varnothing s}$. Currently available EVs often use road classes and map attributes for the initial prediction before the start of a trip. A map-based approach, in which all relevant map features of the segment are considered will not be more accurate than $\widehat{E}_{tot,\varnothing s}$. Hence, $\widehat{E}_{tot,\varnothing s}$ and $\widehat{E}_{tot,\varnothing k}$ are possible comparative values for the evaluation of the introduced deviation prediction system.

The relative error (RE) between consumption E_{act} and the various prediction methods $\widehat{E}_{tot,m}$ is calculated with regard to the mean energy consumption of the road class $\varnothing E_k$, that the results are independent of vehicle attributes. We refer the difference between prediction and consumption E_{act} to $\varnothing E_k$ instead of E_{act}, since E_{act} is often zero in regenerative braking mode. We assume that $\varnothing E_k$ is known exactly. RE only considers the error resulting from the deviation prediction based on the collected speed profiles:

$$RE = \frac{\widehat{E}_{tot,m} - E_{act}}{\varnothing E_k} \tag{7}$$

We calculate the mean absolute percentage error (MAPE) for the evaluation of all n segments or n tracks:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\widehat{E}_{tot,m,i} - E_{act,i}}{\varnothing E_{k,i}} \right| \tag{8}$$

C. Evaluation of the prediction based on segments

Fig. 13 shows the distribution of the RE for each segment of the chosen trips. The consideration of the driver- and situation-specific impact increases the accuracy in comparison to $\widehat{E}_{tot,\varnothing s}$. The RE is higher for E_{neg} than for E_{pos}, since the RE increases with small deviations due to the low absolute values of E_{neg}. The accuracy of E_{pos} is more important for consideration of the complete energy consumption. The distribution of $\widehat{E}_{pos,\varnothing k}$ shows a peak at $RE = 1$. This result occurs, if the regenerated energy on a segment s is zero, whereas the prediction with $\widehat{E}_{pos,\varnothing k}$ is never zero due to the used mean value of the road class.

Table II shows the MAPE of all segments for the prediction variants. The consideration of the driver-specific and situation-specific impact decreases the prediction error from 21.3% for $\widehat{E}_{tot,\varnothing s}$ to 14.9% for $\widehat{E}_{tot,\Delta s}$ for each segment.

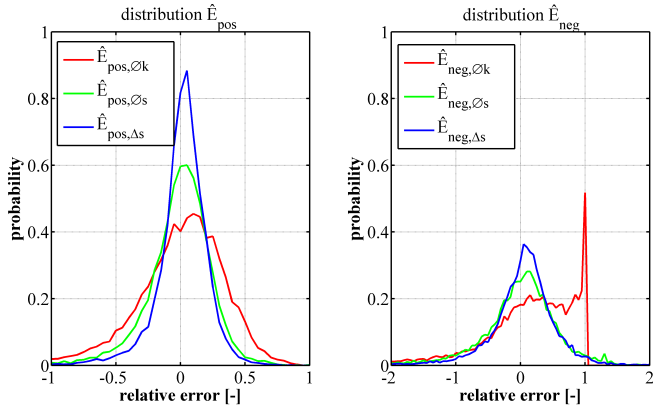The speed profile of several measurements of a segment vary due to the randomly prevailing impact factors like

Fig. 13.   Relative error RE of different methods of energy prediction

TABLE II

EVALUATION PREDICTION SEGMENT-BY-SEGMENT

MAPE	$\widehat{E}_{s,pos}$	$\widehat{E}_{s,neg}$	$\widehat{E}_{s}$
$\widehat{E}_{tot,\varnothing k}$	29.1%	62.8%	33.4%
$\widehat{E}_{tot,\varnothing s}$	19.4%	46.4%	21.3%
$\widehat{E}_{tot,\Delta s}$	14.8%	36.4%	14.9%

Fig. 14.   Energy prediction for a whole route

TABLE III

EVALUATION PREDICTION FOR WHOLE TRIPS

MAPE	$\widehat{E}_{pos}$	$\widehat{E}_{neg}$	$\widehat{E}$	$\varnothing \frac{\widehat{E}_{max}}{E_{act}}$
$\widehat{E}_{tot,\varnothing k}$	14.7%	50.3%	19.0%	1.40
$\widehat{E}_{tot,\varnothing s}$	10.9%	35.0%	12.2%	1.25
$\widehat{E}_{tot,\Delta s}$	6.0%	23.8%	6.8%	1.20

traffic lights, whose influence is considered on average with the expected mean values of the concerned impact factors $\varnothing \widehat{\Delta PC}$. In the next section, we evaluate complete trips, so that the variance of the segment-by-segment prediction is compensated.

D. Evaluation of the prediction for complete trips

Apart from the expected consumption for a complete trip, we calculate the maximum necessary energy $\widehat{E}_{max}$. We define $\widehat{E}_{max}$ as the limit to ensure, that the energy consumption E will not exceed $\widehat{E}_{max}$ in 95% of the cases. $\widehat{E}_{max}$ is important for EV users to guarantee to reach the destination with a high probability.

The stored variance of the normalized energy or the variance of the principal components can be used for the prediction of $\widehat{E}_{max}$. According to the definition of the normal distribution, we use 1.6 times the standard deviation σ above the expected value $\widehat{E}_{act}$ for the calculation of $\widehat{E}_{max}$ for one segment. The multiple of σ decreases with the number of segments of the trip. As result of an experimental simulative study, the multiple of the standard deviation can be reduced to 0.6 in case the trip has more than 25 segments.

As result, Fig. 14 shows the various prediction variants and the maximum necessary energy for the deviation prediction $\widehat{E}_{max,\Delta s}$ and the consumed energy E_{act}, which is predicted. The variance of the prediction of the single segments is balanced along the whole trip. For the exemplary chosen trip more energy is needed as on average, since $\widehat{E}_{tot,\Delta s}$ is higher than $\widehat{E}_{tot,\varnothing k}$ or $\widehat{E}_{tot,\varnothing s}$. The proposed prediction $\widehat{E}_{tot,\Delta s}$ has an relative error of less than 1% for this exemplary chosen trip.
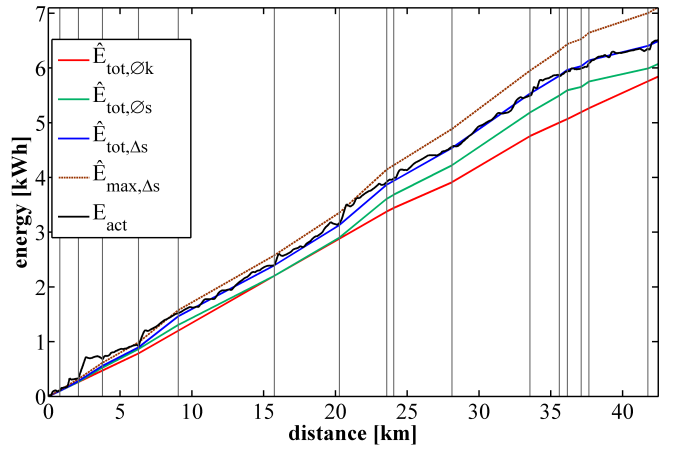
We evaluate the proposed prediction method for complete trips also with the 1000 chosen tracks. Table III shows the results for the mean values of MAPE of all trips. Compared to the results of the evaluation of single segments the MAPE is lower for all prediction variants, since the errors of the single segments are compensated. $\widehat{E}_{tot,\Delta s}$ reaches a high accuracy with an average MAPE of 6.8%. The consideration of the individual driving behavior and situation-specific impact factors in $\widehat{E}_{tot,\Delta s}$ improves the results by 5.4 percentage points compared to $\widehat{E}_{tot,\varnothing s}$, which is based on the normalized mean energy of the segments along the route.

The mean value $\varnothing \frac{\widehat{E}_{max}}{E_{act}}$ is also given in Table III, which indicates the additionally necessary energy to ensure to reach the final destination with a 95% probability. The mean value of this quotient decreases up to 1.20, which means that on average 20% of the needed energy has to be additionally reserved to prevent running out of battery. The proposed energy prediction system in this paper reduces the necessary reserved energy compared to the other prediction methods.

VII. CONCLUSION AND OUTLOOK

The impact of individual driving behavior or prevailing traffic conditions on energy consumption of EVs was considered in prediction methods, which are based on statistical features from collected speed profiles of different vehicles. Relevant features were used to predict segment-by-segment the deviation of the energy consumption from a vehicle-specific average value. We analyzed the processed statistical features regarding individual driving behavior, traffic impact

factors and prevailing weather conditions. The deviation prediction model improved the energy prediction for trips on average by 5.4 percentage points, which resulted in a mean relative error of 6.8%.

The results of the prediction model are the basis for several applications of EVs. The prediction can be shown as information for the driver before the start of a trip if a desired destination is reachable without running out of battery. Furthermore, the prediction can be used for energy management strategies (especially for PHEVs) along the desired route of the driver. The computation time for the prediction depends mainly on the data request for the needed data of the desired route, because the application of the regression model is quite fast compared to other data-mining methods. Thus the prediction of the segments can be used as link weights for navigation algorithms for the calculation of the range estimation or for the solution of eco-routing problems.

In future research, we will compare the results of this paper with the one from machine-learning algorithms, which can also used for energy prediction. Furthermore a deployment in a real fleet-test with different vehicles is necessary instead of the simulation environment used for the evaluation in this paper.

VIII. ACKNOWLEDGMENT

REFERENCES

[1] M. Nilsson, Electric Vehicles: The Phenomenon of Range Anxiety, Public Deliverable, FP7-ICT-2009-4-249105, ELVIRE, 2011.

[2] K. Boriboonsomisn, M. Barth, W. Zhu and A. Vu, Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information, IEEE Transactions on Intelligent Transportation Systems, 2012, vol. 13, no. 4, pp.1694-1704.

[3] Y. Wang, J. Jiang and T. Mu, Context-Aware and Energy-Driven Route Optimization for Fully Electric Vehicles via Crowdsourcing, Intelligent Transportation Systems, IEEE Transactions on, vol. 14, no. 3, 2013, pp.1331-1345.

[4] Y. Zhang, W. Wang, Y. Kobayashi and K. Shirai, Remaining Driving Range Estimation of Electric Vehicle, Electric Vehicle Conference (IEVC), 2012 IEEE International, 2012, pp.1-7.

[5] N. Hoch, K. Zemmer, B. Werther and R.Y. Siegwart, Electric Vehicle Travel Optimization - Customer Satisfaction Despite Resource Constraints, Intelligent Vehicles Symposium (IV), IEEE, Alcala de Henares, 2012, pp.63-68.

[6] O. Sundstroem, O. Corradi and C. Binding, Toward Electric Vehicle Trip Prediction for a Charging Service Provider, Electric Vehicle Conference (IEVC), IEEE International, Greenville, 2012, pp.1-6.

[7] T. Daun, D. Braun, C. Frank, S. Haug and M. Lienkamp, Evaluation of Driving Behavior and the Efficacy of a Predictive Eco-Driving Assistance System for Heavy Commercial Vehicles in a Driving Simulator Experiment, In Proc. of 16th Int. IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013), The Hague, 2013, pp.2379-2386.

[8] C. Dorrer, Effizienzbestimmung von Fahrweisen und Fahrerassistenz zur Reduzierung des Kraftstoffverbrauchs unter Nutzung telematischer Informationen, Doctoral dissertation, Universitaet Stuttgart, expert-Verlag, 2004.

[9] E. Ericsson, Independent driving pattern factors and their influence on fuel-use and exhaust emission factors, in Transportation Research Part D: Transport and Environment, part. 6, 2001, pp. 325-345.

[10] P. Themann and L. Eckstein, Prediction of driver behaviour by situational models in a cooperative environment to optimize energy efficiency, 9th ITS European Congress, Dublin, 2013.

[11] M. Barth, S. Mandava, K. Boriboonsomsin and H. Xia, Dynamic ECO-Driving for Arterial Corridors, Integrated and Sustainable Transportations Systems (FISTS), IEEE Forum on, Vienna, 2011, pp.182-188.

[12] J. Wollaeger, S. A. Kumer, S. Onori, D. Filev, U. Ozguner, G. Rizzoni, S. Di Carano, Cloud-computing based Velocity Profile Generation for Minimum Fuel Consumption: A Dynamic Programming based Solution, American Control Conference (AACC) 2012, pp.2108-2113.

[13] E. Ozatay, S. Onori, D. Filev, U. Ozguner, G. Rizzoni, The use of geographical, traffic management and weather information in the cloud to optimize the energy use of vehicles, In Proc. of 13th Stuttgart International Symposium, 2013, pp.42-56.

[14] C. Andrieu and G.S. Pierre, Using statistical models to characterize eco-driving style with an aggregated indicator, Intelligent Vehicles Symposium (IV), IEEE, Alcala de Henares, 2012, pp.63-68.

[15] Z. Constantinescu, C. Marinoiu and M. Vladoiu, Driving Style Analysis Using Data Mining Techniques, Int. J. of Computers, Communication & Control, vol.5, no. 5, 2010, pp.654-663.

[16] X. Huang, Y. Tan and X.He, An Intelligent Multifeature Statistical Approach for the Discrimination of Driving Conditions of a Hybrid Electric Vehicle, IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 2, 2011, pp. 453-465.

[17] M. Hamacher, J. Papajewski et al., Audi Energy Assist - Efficiency and Adaptive Controlled Consumer Experience for Plug-In-Hybrid Electric Vehicles, In Proc. of 22nd Aachen Colloquium Automobile and Engine Technology, 2013, pp.1443-1464.

[18] D. Karbowski, S. Pagerit and A. Calkings, Energy Consumption Prediction of a Vehicle along a User-Specified Real-World Trip, In Proc. from Electric Vehicle Symposium (EVS 26), Los Angeles, 2012.

[19] S. Grubwinkler and M. Lienkamp, A modular and dynamic approach for the prediction of the energy consumption of electric vehicles, In Proc. of CoFAT, Munich, 2013.

[20] S. Grubwinkler, M. Kugler and M. Lienkamp, A system for cloud-based deviation prediction of propulsion energy consumption for EVs, Vehicular Electronics and Safety (ICVES), 2013 IEEE International Conference on, Dongguan, 2013, pp.99-104.

[21] OpenStreetMap, http://www.openstreetmap.org/.

[22] J. Yoon, B. Noble and M. Liu, Surface Street Traffic Estimation, In Proc. of the 5th Conference on Mobile Systems, Applications and Services (MobiSys07), ACM, New York, 2007, pp.220-232.

[23] P. Tulpule, V. Marano and G. Rizzoni, A Statistical Approach to Asses the Impact of Road Events on PHEV Performance using Real Word Data, SAE World Congress, SAE International, no. 2011-01-0875, 2011.

[24] Q. Gong, S. Midlam-Mohler, V. Marano, G. Rizzoni and Y. Guezennec, Statistical analysis of PHEV fleet data, Vehicle Power and Propulsion Conference (VPPC), 2010 IEEE, 2010, pp.1-6.

[25] H. Hennig, Der A1 e-tron im Flottenversuch, in Proc. of COFAT, Munich, 2012.

[26] M. Kilpelaeinen and H. Summala, Effects of weather and weather forecasts on driver behaviour, Transportation Research Part F: Traffic Psychology and Behaviour, vol. 10, no.4, 2007, pp.288-299.