Abstract—In this paper, we evaluate a time series based method for predicting the first daily departure time (FDDT) of commuter vehicles. This task is relevant for the grid integration of plug-in electric vehicles (PEVs), since it allows for actively managing their electricity demand during the connection interval. Our study is based on a sample of 445 vehicle usage traces which were collected in the Puget Sound Traffic Choices Study using the Global Positioning System (GPS). We advance knowledge in the area of vehicle usage prediction in a smart grid context in four ways: First, we propose a method for selecting variable size subsets of vehicle usage traces that are more predictable irrespective of the concrete forecasting method used. Second, we show that FDDTs should not be modeled with basic theoretical probability distributions. Third, we compare a number of time series based forecasting models using the mean absolute deviation (MAD) and forecast availability to identify a model with high forecasting accuracy and high forecast availability. Fourth, we report empirical confidence intervals to demonstrate the performance of our approach and discuss implications with respect to using PEVs for demand side management. A major finding of this work is that, even for commuters, FDDTs are hard to predict using historical realizations alone. However, the forecasting accuracy varies a lot from vehicle to vehicle and weekday to weekday. Moreover, forecasting errors are skewed in a way that facilitates non-disruptive charging control but may limit load shifting potential. Finally, using the example of ex-ante knowledge about workdays, we show that further potential for improvement lies in considering more data sources.

I. INTRODUCTION

Plug-in Electric Vehicles (PEVs) have a high potential for supporting the electric grid, especially as the share of renewable energy increases [1], [2]. Most people use their cars only for relatively short trips and leave them parked in front of their homes or offices for many hours. Thus, major parts to the future PEV fleet could be leveraged as decentralized grid-connected energy storage without restricting actual usage [3].

It will be a major challenge to develop user-friendly systems that efficiently balance user and grid requirements in an uncertain environment. Accurate prediction of these requirements plays a key role in controlling the charging and discharging of PEV batteries. One highly important user requirement is maximal electric mobility. If departure times are underestimated, the charging algorithm will make sure that a sufficient state of charge is reached even before the actual departure time. However, the algorithm will not be able to use the entire demand-side management potential contained in charging jobs. This is because the algorithm expects the PEV to be unavailable for (dis)charging during a certain time interval although, in reality, it is available. If departure times are overestimated, the result is inverse: The charging algorithm may not be able to assure the required battery state of charge, but will at least take advantage of the entire period during which the PEV is connected to the grid. In any case, a more accurate prediction of departure times improves the performance of PEV load shifting algorithms.

In this paper, we propose and investigate the performance of basic methods for predicting the first daily departure times (FDDTs) of cars. Forecasts of FDDTs can be used to efficiently schedule charging or discharging in the parking period ending with the corresponding FDDT. Current PEV technology allows for complete battery recharging in less than 4 hours [4]. Nightly parking intervals are, on average, significantly longer than other parking intervals during the day, e.g. parking at work or other locations. Thus, the potential grid connection period ending with the FDDT is the most attractive for managed charging or vehicle-to-grid (V2G) operation. The more accurately FDDTs can be predicted, the more efficiently power usage or feed-in can be scheduled to maximize grid support and remaining within the constraints imposed by drivers [2].

Recent work on PEV load shifting, e.g. [2], is often based on the assumption that vehicle departure times are known. This assumption is not completely unrealistic, since drivers could be asked to specify next departure times. However, the likelihood of over- or understating this time is rather high, even if malicious intend is excluded. Moreover, having to think about and specify departure times would significantly increase user effort and thus reduce user-friendliness. User-friendliness is considered to be one of the major goals that smart grid technology has to achieve to be adopted on a large scale.

To date, the majority of research efforts have focused on modeling collective driving behavior. This type of research was conducted both by academics and industry. Academic research includes [5], [6], [7], [2]. One diligently covered industrial project is conducted at Google [8], but many other companies are developing solutions based on the insights gained from analyzing vehicle usage data.

Research about predicting driving behavior is sparse. [5]
probably uses the most advanced models of driving behavior that are based on actual data. They develop several such models, which are eventually used for simulating aggregate PEV charging profiles. The models are based on empirical probability distributions of FDDT, daily mileage, and latest daily arrival time. [5] do not predict FDDTs of individual cars or take the seasonality of weekly driving behavior into account. Instead, they combine all available data points to construct the empirical distributions required for their research.

In the area of PEV grid integration, we are first to investigate ways to predict the relevant variables using a vehicle-centric time series based approach.

II. Data

To evaluate forecasting methods, we require detailed driving data for individual vehicles. Furthermore, the required data should be available in high resolution, if possible in minutes or even seconds, and over a longer period of time, at least for several months. Previous work about PEV grid integration focused on collective driving behavior and therefore had less strict requirements with respect to time resolution and coverage of vehicle usage data. For instance, the publicly available data of the National Household Travel Survey (NHTS) contains the trip data of large numbers of vehicles and has been used for characterizing aggregate vehicle usage [9], [2]. Similar sources of driving data include the Austin and San Antonio GPS-enhanced Household Travel Survey [10], the Houston-Galveston Area GPS-enhanced Household Travel Survey [11], and the German Mobility Panel [12]. Although some of these data sets have been enhanced with limited amounts of GPS data, their goal was to collect representative travel data instead of collecting data over a long period of time.

To the best of our knowledge, there are currently two publicly available sets of vehicle usage data that fulfill the requirements of this study: the data of the Traffic Choices Study (TCS) collected by the Puget Sound Regional Council in 2005 [13], and a data set made available by researchers of the University of Winnipeg [14]. These data sets contain usage traces of 445 and 76 cars, respectively, and were collected over the time of approximately one year in second resolution. The TCS data also contains identifiers indicating the origin and destination of each trip, e.g., home to work. Due to its coverage and trip location labeling, we have chosen to use the TCS data to evaluate our FDDT forecasting approach.

III. Method

Our analysis has four steps discussed below.

Since we only consider cars used for commuting to work in this study, our first step was to eliminate 55 “non-commuter” traces from the sample. These traces contained no workday defined as a day with at least one trip to work. Furthermore, we also deleted 80 traces with a median of less than one work day per week from the data set. Thus, \( n = 310 \) vehicle usage traces remained to be analyzed.

In a second step, we applied a self-developed classification method for dividing the remaining sample into two groups labeled “well-behaved” and “variable”. The classification is based on two variables: the variability of workday FDDTs and workdays per week. Both of these variables are measured using their standard deviations. Using these two variables, each vehicle usage traces can be represented as a one point in a two-dimensional space as shown in Figure 1.

Using this representation of vehicle usage traces as starting point, our classification method works as follows. First, it draws a straight line through the zero point such that half of the points lie above and half of them below the line. The line defines the location range of the upper right corner of a rectangular box. The lower left corner of this box is the zero point of the scatter plot. The box can be scaled to encompass an arbitrary fraction of the points representing vehicle usage traces, denoted by \( \alpha \). Irrespective of \( \alpha \)‘s value, we refer to the traces that lie within the box as well-behaved. The well-behaved traces can be viewed as the training set of traces based on which we will determine the best model fit in the following.

In the third step of our analysis, we used the statistical software SPSS to test whether FDDTs can be modeled using standard theoretical probability distributions. First, we applied the Kolmogorov-Smirnov and Shapiro-Wilk test to determine whether the workday FDDTs of well-behaved commuter vehicles follow a normal distribution. The tests yielded insignificant levels of model fit. A subsequent histogram analysis revealed that this result is due to the high kurtosis and heavy right tails of the empirical FDDT distributions. We proceeded by using Q-Q plots to visually test a number of other theoretical distributions that usually better fit such data. In particular, we tested the Laplace, Exponential, and Log-normal distribution. It turned out that neither of these distributions yields an acceptable fit, either. Thus, we preliminarily concluded that modeling FDDTs using theoretical probability distributions is not advisable.

In the fourth step of our analysis, we applied a time series approach for predicting FDDTs based on historical values.

![Fig. 1. Classification of vehicle traces.](image-url)
Figure 2 shows box plots of FDDTs for a selection of 5 vehicles. The vehicle represented by the blue box plots (the first box plot in each day group) is the least well-behaved vehicle in a well-behaved set of size \([\alpha n]\), where \(\alpha = 0.01\). The red vehicle (the second box plot in each day group) is the least well-behaved at \(\alpha = 0.25\), and so forth. The box plots in Figure 2 provide a good intuition about the weekly seasonality of FDDTs: Monday to Friday FDDTs have a smaller spread and greater median compared to weekend FDDTs. Furthermore, there exist significant differences even between the Monday to Friday FDDTs. Based on these insights, we have chosen to predict the FDDT of a car \(i\) at day \(t\), \(x_i(t)\), using several forecasting models parametrized by \(s\), \(w\), \(h\), and \(m\). Our general approach consists of calculating the FDDT forecast for time \(t\), \(\tilde{x}_i(t)\), based on a set of historical values \(H_i(t, s, w, h)\).

If the parameter \(s\) is set to the “weekdays”, \(H_i(t, s, w, h)\) contains the last FDDTs of the same day type, e.g. Monday, that are found in trace \(i\) looking back until time \(t-h\). If \(s\) is set to “Mon-Fri/Sat-Sun”, \(H_i(t, s, w, h)\) contains the last FDDTs belonging to the same group, i.e. the forecast of, e.g., Mondays is based on all historic FDDTs of Mondays to Fridays, and the forecast of, e.g., Saturdays is based on all historic FDDTs of Saturdays and Sundays. Again, FDDTs until a maximum age of \(h\) days are considered. This implies that the number of values in \(H_i(t, s, w, h)\) is usually greater if \(s=\)“Mon-Fri/Sat-Sun”.

The parameter \(w\) indicates whether the fact that a FDDT marks the start of a work trip is considered in the forecast. If \(w\) is set to “true”, FDDTs of work trips are predicted solely using historic workday FDDTs and FDDTs of leisure trips are based on leisure trip FDDTs only. This implies that the number of values in \(H_i(t, s, w, h)\) is usually greater when a workday FDDT is needed for, e.g. a Monday, compared to when a workday FDDT is required for a Sunday, because only few cars are used for commuting on weekends. If \(w\) is set to “false”, ex-ante knowledge of work trip FDDTs is not available. In this case, \(\tilde{x}(t)\) is based on all available FDDTs irrespective of being work or leisure related.

Finally, parameter \(m\) defines the method applied for aggregating the historic values. If \(m\) is set to “mean”, an equally weighted average is taken, if \(m=\)“median”, the median of the values in \(H_i(t, s, w, h)\) is calculated.

Our forecasting approach thus considers the weekly seasonality of FDDTs, but ignores longer term trends, special events, or other possible influencing factors (e.g. the weather).

To evaluate the proposed forecasting approach, we extracted the FDDT time series from each vehicle usage trace and annotated them with the required parameter values.

### IV. Results

The chosen forecasting approach can be configured in a number of ways by choosing the parameters \(\alpha, s, w, h,\) and \(m\). To evaluate the approach, we calculated the mean absolute deviation of the actual FDDTs from their forecasts, \(MAD_i(s, w, h, m)\), for all well-behaved FDDT time series \(i\) and all possible parameter configurations in Table I.

\[
MAD_i(s, w, h, m) = \sum_{t=1}^{T} |x_i(t) - \tilde{x}_i(t)| / T 
\]

Where \(\epsilon_i(t, s, w, h, m)\) denotes the forecasting error at time \(t\) for time series \(i\), \(\epsilon_i(t, s, w, h, m)\). To make the results comparable, we begin the computation of the MAD at \(t=84\) irrespective of the configuration of the forecasting method: If \(h=12\), the maximum history is \(12 \times 7 = 84\) days.

The forecasting error is the difference of the FDDT realization, \(x_i(t)\), and the corresponding forecasted FDDT, \(\tilde{x}_i(t, s, w, h, m)\), i.e. \(\epsilon_i(t, s, w, h, m) = x_i(t) - \tilde{x}_i(t, s, w, h, m)\). Thus, if \(\epsilon\) is positive, the forecasting method predicted an earlier departure time, and vice versa.

For some days \(r\), the forecasting error \(\epsilon_i(r, s, w, h, m)\) cannot be computed due to missing values. Either \(x_i(t)\) does not exist because the vehicle \(i\) was not used on that day, or \(H_i(r, s, w, h)\) is an empty set. Since the size of \(H_i(t, s, w, h)\) depends on the choice of the forecasting parameters, the fraction of times when a FDDT forecast is not available, \(p_i(s, w, h, m)\), will be considered as a separate performance metric in the following. We calculate \(p_i(s, w, h, m)\) by dividing the number of times when no forecast is available despite vehicle usage on that day by the number of computable forecasting errors.

Equation 1 formally defines the \(MAD_i(s, w, h, m)\), where \(T_i\) denotes the length of time series \(i\). For time series \(i\), the function \(I_i(t, s, w, h)\) returns 1 if the forecasting error \(\epsilon_i(t, s, w, h, m)\) at day \(t\) can be computed using the method defined by the parameters \(s, w,\) and \(h,\) and 0 otherwise.

\[
MAD_i(s, w, h, m) = \frac{\sum_{t=1}^{T_i} |x_i(t) - \tilde{x}_i(t)|}{\sum_{t=1}^{T_i} I_i(t, s, w, h)} 
\]
The impact of increasing $\alpha$ has the effect implied by our analysis of the average MAD resulting from different parameter configurations yielded the following results:

The maximum memory of forecasting method, $h$, has a visible but surprisingly small impact on the average MAD. Increasing the look-back capability from 4 to 12 weeks improves the average MAD by only between 1 and 4 minutes depending on the configuration. However, increasing $h$ improves the forecast availability, which is particularly useful if $s$ is set to weekdays and $w$ is set to true. Under these circumstances, the forecast unavailability can easily surpass 10% at low levels of $h$.

According to our results, the changes of the average MAD resulting from using the “weekdays” instead of the “Mon-Fri/Sat-Sun” seasonality method is rather small, usually in the range of 1 to 2 minutes. This is remarkable considering the fact that the weekdays method’s forecasts are based on fewer values: The consideration of a more fine-grained seasonality sets make up for the smaller number of historical data points. The difference between the two types of seasonality shows more prominently if one compares forecast unavailability: Here the Mon-Fri/Sat-Sun forecasting method has a clear advantage.

The median based smoothing method always performs slightly better than the mean based one. This is due to the fact that the forecasting errors have a skewed distribution. Therefore, setting $m$=“median” can be considered optimal under all circumstances.

The table below illustrates the average MAD and percentage of missing forecasts for different parameter configurations. The table shows the average MAD for different levels of $\alpha$, weekdays, and $h$. The average MAD is calculated using the following formula:

$$MAD_i(s, w, h, m) = \frac{1}{\sum_{t=84}^{T} I_i(t, s, w, h)} \sum_{t=84}^{T} I_i(t, s, w, h) |\epsilon_i(t, s, w, h, m)|$$

To compare the effect of different parameterizations, we calculated the average MAD and percentage of missing forecasts across all time series in the well-behaved vehicle set at different levels of $\alpha$. Table II provides these average MADs for all possible parameter configurations based on the values in Table I.
forecast availability. At high levels of \( h \) and \( s \)="Mon-Fri/Sat-Sun", the impact of using ex-ante knowledge of workdays on forecast availability becomes rather small, but is still present.

Overall, our results reveal a surprisingly low influence of the forecasting parameters on the performance of the time series based forecasting method in terms of accuracy. The effect of using the weekday seasonality type and using ex-ante knowledge of workdays, however, can be critical with respect to forecast availability.

A good compromise between forecasting accuracy and availability appears to be using the Mon-Fri/Sat-Sun type of seasonality and ex-ante knowledge of workdays. The remaining parameters \( h \) and \( m \) should be set to the maximum value and median, respectively. Thus, we refer to the configuration \((s, w, h, m)\)=("Mon-Fri/Sat-Sun",true,12,median") as optimal in the following. The average MAD difference between this compromise configuration and the worst configuration in our sample can be as high as 20 minutes, which translates to a 20 to 25% improvement.

If the forecasting error followed a normal distribution with zero mean, one could calculate symmetrical confidence intervals directly from the error samples obtained using our forecasting method. In this particular case, the sample standard deviation \( \hat{\sigma} \) can be used to estimate the standard deviation of the forecasting error. A corresponding symmetrical confidence interval would then indicate that the actual value \( x_i(t) \) lies within the interval \([\hat{x}_i(t) - \delta_i; \hat{x}_i(t) + \delta_i]\) with probability \( \gamma \), where \( \delta_i = F^{-1}_i(0.5) - F^{-1}_i(1-\gamma^2) \) and \( F^{-1}_i \) is the inverse cumulative normal distribution function with zero mean and standard deviation \( \hat{\sigma} \). We tested the hypothesis that the forecasting error is normally distributed by applying the same standard statistical tests as for the FDDTs. Based on the results we obtained, we had to reject this hypothesis. Similar to the distribution of FDDTs, the empirical distributions of the
forecasts errors showed a heavy right tail and a high kurtosis. Thus, to obtain a realistic and informative representation of FDDT forecasting accuracy on the individual vehicle level, we have chosen to report confidence intervals based on the empirical distribution of FDDT forecasting errors.

An empirical forecasting confidence interval is defined as 
\[
\left[ \hat{x}_i(t) + \epsilon_i \left\lfloor \frac{1}{1 + \gamma} \right\rfloor n, \hat{x}_i(t) + \epsilon_i \left\lceil \frac{1}{1 + \gamma} \right\rceil n \right],
\]
where \( n \) is the number of observations of the forecasting error and \( \epsilon_i \) is the \( j^{th} \) element in the list of empirical forecasting errors of time series \( i \) in ascending order. In the following, to make the confidence intervals comparable, we report empirical confidence intervals of the forecasting error at levels \( \gamma = 0.95 \) and \( \gamma = 0.5 \).

Figures 3 and 4 show plots of the empirical confidence intervals across different weekdays. Since we consider FDDT prediction on the individual vehicle level, we show these confidence intervals for a selection of 4 vehicles. We selected the corresponding time series in ascending order of MAD: The first vehicle (dark blue) for which we plotted the forecasting confidence intervals is the one with the greatest MAD in a subset of \( \beta n \) vehicles with the smallest MADs are considered and \( \beta = 0.01 \). The second vehicle (red), has the greatest MAD in a set consisting of the 25% vehicles with the smallest MADs, and so on. For computing the order of vehicles that the selection is based on, we fix the parameters at the optimal level defined above.

As our previous analysis has shown, the consideration of ex-ante knowledge workdays can have a significant impact on forecasting accuracy. Therefore we report the confidence intervals for both \( w = \text{false} \) and \( w = \text{true} \) in Figures 3 and 4. Figure 3 contains plots for \( s = \text{weekdays} \), whereas Figure 4 shows plots for the “Mon-Fri/Sat-Sun” type of seasonality.

Comparing the confidence intervals of the selected vehicles, one gets an impression of the great differences of forecasting accuracy across the weekdays. As expected, the confidence intervals on weekends are usually wider compared to the other days. However, the upper boundaries of the last two confidence intervals (green and light blue) extend to weekend levels on several weekdays. Thus, the decreasing prediction accuracy at increasing levels of MAD is caused by hardly predictable late departures occurring from Monday to Friday. These could be caused by irregular leisure activities in the afternoon or evenings. Overall, the forecasts are more accurate on the negative side, i.e. the probability that our method predicts a much later FDDT is smaller than the probability that it predicts a much earlier FDDT.

Our results show that forecasting commuter FDDTs is surprisingly hard. Typical 95% confidence intervals span several hours. Furthermore, the forecasting errors are usually skewed in a way that is advantageous for achieving what [3] call non-disruptive control, because the probability that our method predicts a much earlier departure is higher. At the same time, the forecasting method is less advantageous if the goal is to maximize the grid support delivered by controlled PEV (dis)charging. This potential is also influenced by other factors, e.g. the absolute timing and distribution of grid connection time intervals and amount of energy delivered to PEVs (cf. [2]). Therefore, the impact of forecasting accuracy needs to be further evaluated based on appropriate models and figures of merit.

One possibility to increase FDDT forecasting accuracy is to consider more information than just historical FDDTs. In this way, we have shown that ex-ante knowledge of workdays has significant potential in this context. Many other types of valuable information are imaginable, e.g. weather, vacation times, or work schedules.

**REFERENCES**


