Asynchronously updated predictions of electric vehicles' connection duration to a charging station

1st Milan Straka
University of Žilina
Žilina, Slovakia
milan.straka@uniza.sk

2nd Martin Jančura
University of Žilina
Žilina, Slovakia
jancura1@stud.uniza.sk

3rd Nazir Refa
ElaadNL
Arnhem, The Netherlands

4th Ľuboš Buzna
University of Žilina
Žilina, Slovakia
lubos.buzna@uniza.sk

***Abstract*—Electric vehicles are promising to mitigate the increasing CO_2 emissions from transport, provided that renewable energy sources generate the demanded electricity. The stochasticity of renewable energy sources and charging demand require intelligent charging schemes. Smart charging achieves better performance when it is driven by reasonably accurate predictions of charging behaviour. Hence, for a smart charging scheme that dynamically updates a charging schedule, updating the predictions of charging behaviour could be beneficial. In this paper, we explore the potential to improve the accuracy of prediction models of the connection duration to a charging station by updating the predictions as the charging sessions unfold. We compare a single-model with multiple-models for regularly and irregularly spaced updates in time. The multiple-model with irregular updates achieves the best performance while improving the prediction accuracy up to $30~\%$, compared to conventional approaches. It is efficient to update the predictions with higher frequency in the very early stages of charging sessions. Later on, regular updates are sufficient.**

***Index Terms*—electric vehicles, smart charging, updated predictions, machine learning, data science**

List of symbols

- N - number of charging sessions,
- M - number of updates of a prediction,
- M_k - number of updates of a prediction of the k-th session,
- $\boldsymbol{x}$ - vector of feature values associated with one observation,
- $\boldsymbol{x}_k$ - vector of feature values associated with k-th observation (session),
- t_k^i - time offset of the i-th prediction update since the start of k-th session,
- t^1 - time offset of the first prediction,
- y_k - response variable (connection duration),
- $\hat{y}_k^i$ - estimate of the connection duration of k-th charging session resulting from the prediction update made at the time offset t_k^i since the session start.
- δ - length of the time interval between two prediction updates,
- δ_k^i - length of the time interval between i-th and $i+1$-th connection duration prediction updates of the session k,

This work was supported in part by project VEGA 1/0077/22 - Innovative prediction methods for optimisation of public service systems, in part by VEGA 1/0216/21, "Design of emergency systems with conflicting criteria with the tools of artificial intelligence",in part by APVV-19-0441 - Allocation of limited resources to public service systems with conflicting quality criteria, in part by the Spanish Ministry of Economic Affairs and Digital Transformation under the project IA4TES MIA.2021.M04.0008- Advanced intelligent technologies for a sustainable energy transition, in part by the Basque Government under the ELKARTEK program, and in part by the Operational Program Integrated Infrastructure 2014–2020 "Innovative Solutions for Propulsion, Power, and Safety Components of Transport Vehicles" through the European Regional Development Fund under grant ITMS313011V334.

I. Introduction

A. Motivation

In the recent years, the humanity is looking more intensively for options how to decrease anthropogenic CO_2 emissions, as they are correlated with global warming [1]. A promising solution to decrease CO_2 emissions are electric vehicles (EVs), when charged from renewable energy sources. The stochastic nature of the EV charging demand and renewable energy sources ask for a smart charging to coordinate the charging process. Efficient smart charging requires estimates of the future developments, e.g., predictions of the charging behaviour. Several prediction approaches already exist in the literature. However, often a single prediction is made at the arrival of an EV to a charging station and it remains valid for the whole duration of a charging session. For smart charging schemes that dynamically update a charging schedule, it would be possible to update also predictions of charging behaviour. Hence, there is potential to reach higher efficiency of dynamic smart charging, if more accurate predictions of charging behaviour can be achieved by updates.

B. Literature review

1) Prediction models of charging behaviour: Among the most popular problems in the EV field are the forecasts of aggregated demand [2]. However, some smart charging schemes require forecasts of individual charging behaviour. The literature proposing prediction models of individual charging behaviour arose together with the availability of open datasets. In [3], [4] authors proposed a methodology to predict the energy consumption and connection duration. In [5] authors predicted connections duration of EVs and evaluated them in the context of smart charging schemes. Although these approaches improve the accuracy compared to benchmark models, they do not consider updating the predictions in time as the charging unfolds.

2) Updates of predictions: Updated predictions are beneficial in many situations, such as for example duration of system outages [6] or incidents duration [7]. Many applications of updated predictions can be found in transport. Methodologies based on Bayesian network were proposed to update predictions of train delays [8], [9]. In [10], train delays predictions are updated using timed event graph with dynamic arc weights. In the context of individual charging behaviour we have not identified any paper exploring updates of predictions.

II. Data and Methods

A. EVnetNL dataset

For numerical experiments we use the EVnetNL dataset maintained by the ElaadNL - knowledge and innovation centre in the field of smart charging and the charging infrastructure [11]. The dataset comprises of two tables, "Transactions" and "Meterreadings". Each charging session in the table "Transactions" is identified by charging point and connector numbers, geographical coordinates, initial and terminal timestamps, and the hashed number of the user RFID cards used to initiate and terminate charging sessions. Table "Meterreadings" documents the energy consumption by meter readings made every 15 minutes. The subset of selected data spans from 01/2016 to 07/2018. In this period the number of charging stations was already stable [12]. The dataset covers 1731 public and semi-public charging stations, about 65k EV drivers, more than 900k charging sessions, and more than 30M meter readings.

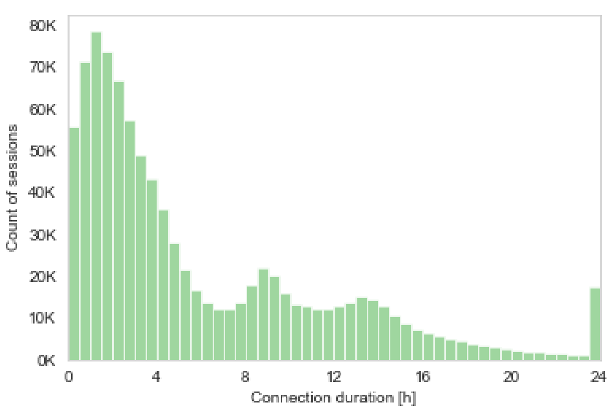

Fig. 1. Histogram of the connection duration for charging sessions in the EVnetNL dataset.

In the dataset, we found a small number of sessions with exceedingly large connection duration, sometimes several days or even weeks. Regarding smart charging, such long duration is not of high relevance and for this reason the connection duration is capped to 24 hours. The distribution of connection duration is shown in Fig. 1. The majority of connections take less than five hours. Only a minority of sessions last more than 20 hours.

B. Feature engineering

To characterise a charging session, we designed a set of features that can be organised in four groups: static features, features describing long-term charging history, features describing short-term charging history and online features. Features included in the first three groups capture developments taking place prior to the start of the charging session upon which the prediction is made and they were designed by considering previous studies [3], [5], [13]. The last group of features captures the progress from the start of the charging session until the time when the prediction is made.

1) Static features: These features take a constant value for all sessions associated with a station or a user:

- first two letters of the station label encoding its type (modelled as a categorical variable and one-hot encoded) (7 features),
- longitude and latitude of the station (2 features),
- maximum charging power estimated as the minimum of the user maximum power and station maximum power (1 feature).

2) Features describing long-term charging history: Features to capture characteristics of charging sessions in a long-term by aggregated statistics:

- mean, minimal and maximal values of the total charged energy, connection duration and charge duration (all values are calculated for both, charging sessions previously made by a user and charging sessions previously taking place at a charging station) ($3 \times 3 \times 2$ features),
- relative frequency of sessions that lasted more than is the current connection duration (the value is calculated for both, for a user and for a charging station) (2 features).

3) Features describing short-term charging history: Features calculated for n most recent days or n most recent sessions, to capture the short-term charging history:

- the mean value of the charged energy, the mean value of the connection duration and the count of sessions considering last day (week) for each station (3×2 features),
- the mean values of the energy consumption and connection duration in the last 1, 5, 10 sessions for each user (2×3 features).

4) Online updated features: Features that capture the progress of the charging session since it has started:

- current hour of the day, weekday and month (modelled as categorical variables and one-hot encoded) (43 features),
- total charged energy since the start of the session (1 feature),
- charged energy in the last 15, 30 and 60 minutes (3 features),
- connection duration since the beginning of the session (1 feature).

C. Prediction Methods

1) Naive models: Two naive prediction models are applied to assess the performance of the proposed models:

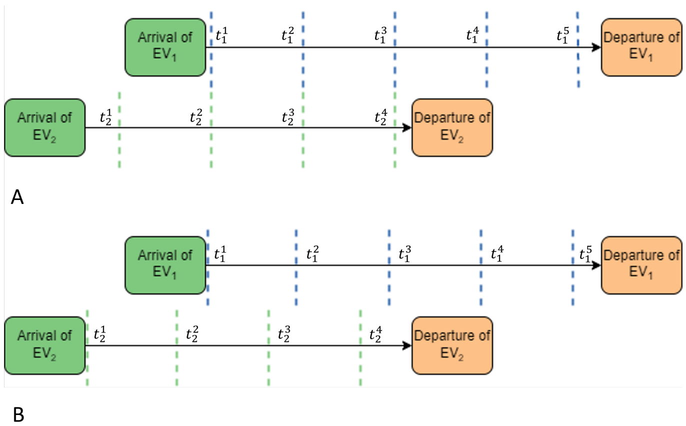

Fig. 2. **A**: Synchronously updated predictions. **B**: Asynchronously updated predictions. We consider two parallel charging sessions initialised at the time of EV arrival and terminating at the departure. Dashed lines indicate times when the predictions are updated.

- Mean-static - the mean connections duration of charging sessions associated with a given user,
- Mean-updated - the mean connection duration of all the user's sessions lasting more than the current duration of the session upon which the prediction is made.

2) LightGBM: We use LightGBM [14] as the prediction method which is a state of the art implementation of the Gradient Boosted Regression Trees (GBRT) [15]. The LightGBM applies two novel techniques: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) speeding up the training process of conventional GBRT by up to over 20 times while achieving almost the same accuracy. For our purposes this is a large advantage as we need to train large number of models.

D. Update strategies

To increase the overall model accuracy, we update the predictions in discrete time steps [8]. There are two possible approaches how to handle the prediction updates (see Fig. 2). The synchronous approach updates all the sessions at the same time, while the asynchronous approach handles each session independently of other sessions. Obviously, synchronous approach is much more resource demanding as it requires a global clock and the online features need to be adjusted to the start time of sessions upon which the predictions are made. For these reasons, we explore further the asynchronous approach.

We can combine the asynchronous approach with either *regular* or *irregular* updates (see Fig. 3). The irregular updates can be used to minimise the overall prediction error when appropriately adjusting the periods between updates. Furthermore, we can train one prediction model for all updates or one prediction model for each update. Considering these options, we compiled the following approaches:

- Static-model - the connection duration is predicted only once at the time when the electric vehicle is plugged in,

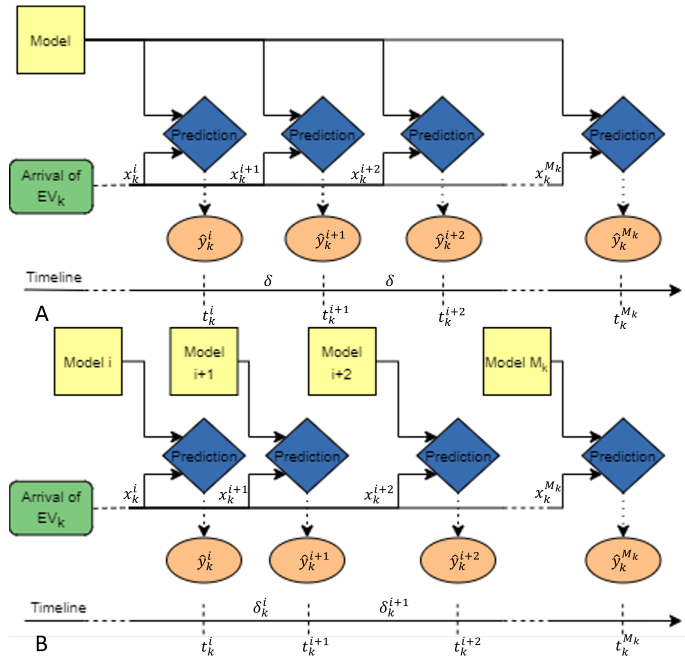

Fig. 3. **A**: A schematic illustrating the single-model approach with regular update strategy. **B**: A schematic illustrating the multi-model approach with irregular update strategy. For every charging session k and the time offsets t_k^i for $i = 1, .., M_k$ since the session start, the predictions $\hat{y}_k^i$ are made using the feature vector $\boldsymbol{x}_k^i$. The length of the period between i-th and $i+1$-th updates is either denoted as δ_k^i (irregular updates) or δ (regular updates).

- Single-model - all prediction updates are made with the same model (Fig. 3A),
- Multi-model - for every prediction update a distinct model is trained (Fig. 3B).

The LightGBM method was used for every approach.

E. Preparation of models

We divided the EVNetNL dataset into the warm-up set (sessions from 2016), training set (sessions from 2017), validation set (the first 50k sessions from 2018) and test set (the remaining sessions from 2018). The warm-up set comprises the history that is required to calculate values of some features. A validation set was allocated to avoid peeking, i.e. using test-set performance to do both to choose a hypothesis and to evaluate it [16]. Thus, the validation set is used to evaluate the values of hyperparameters.

To construct a multi-model with regular updates, each of the models is trained on its own dataset, representing the active sessions with target variable y_k and updated features $\boldsymbol{x}_k$. To tune separately hyperparameters of each model would be computationally very expensive, hence, we identified fit-for-all values of hyperparameters on the validation set. The most suitable values of update times for the irregular multi-model are found as a part of hyperparameter optimisation. To keep the computational burden within a reasonable limits, the features describing long-term history are not updated throughout the process. As a search strategy to find suitable values of hyperparameters, we applied the Bayesian optimisation. To decide which point to evaluate, the expected improvement is estimated [17]. We used the Tree-structured Parzen Estimator Approach (TPE) [18] to model the conditional probability and transform the configuration space to facilitate the expected improvements.

F. Error measures

To take into account that each session can be updated multiple times, we adjusted the standard prediction error measures. We adapted the weighted mean absolute error [19] by averaging the mean weighted absolute error over all updates

$$wMAE = \frac{1}{N}\sum_{k=1}^{N}\frac{\sum_{i=1}^{M_k}|y_k - \hat{y}_k^i| * \delta_k^i}{\sum_{i=1}^{M_k}\delta_k^i}, \qquad (1)$$

where y_k is the connection duration of the charging session k in the test data, $\hat{y}_k^i$ is the i-th update of the connection duration prediction produced at time t_k^i, N is the number of all charging sessions in the test set, M_k is the number of updates and δ_k^i is the period between i-th and $i+1$-th updates. For regular updates the length of all periods is the same and we denote it as δ. When evaluating the models, predictions are made for all t_k^i which precede the end time of the k charging session. The last δ_k^i is selected to share the end point with the the end time of the k charging session.

We evaluate the accuracy of predictions for the time offsets since the sessions start t^i, for $i = 1, \ldots, M$ by

$$MAE_{t^i} = \frac{1}{N}\sum_{k=1}^{N}|y_k - \hat{y}_k^i|. \qquad (2)$$

In Eq. 2 only active sessions are considered. By definition, t^1 is set to value of 0. Thus, the quantity MAE_0 evaluates the accuracy of predictions at the time of EV arrival to the charging station.

III. Results

A. Regular updates

TABLE I
COMPARISON OF NAIVE (MEAN-STATIC, MEAN-UPDATED) AND ADVANCED (STATIC-MODEL, SINGLE-MODEL, MULTI-MODEL) PREDICTION MODELS.

Methods	**MAE$_0$** [hours]	**wMAE** [hours]
Mean-static	2.814	2.814
Mean-updated ($\delta = 1$)	2.814	2.346
Static-model	2.227	2.227
Single-model ($\delta = 1$)	2.302	1.798
Multi-model ($\delta = 1$)	2.227	1.698

Table I compares the performance of naive models with advanced models updating predictions regularly every hour ($\delta = 1$). The results confirm the superiority of advanced models. The advantage of updating the predictions is already evident from comparing the Mean-static and Mean-updated models. On average, the error drops by almost 0.5 hour due to updated predictions. Similarly, when contrasting the Static-model with Single-model and Multi-model the improvement is also about 0.5 hour. Interestingly, when considering only the value of MAE_0, the Static-model outperforms Single-Model. In Table II, we study the role of the frequency of updates.

TABLE II
IMPACT OF THE FREQUENCY OF UPDATES ON THE ACCURACY OF PREDICTIONS.

Models	wMAE ($\delta = 2$)	wMAE ($\delta = 1$)	wMAE ($\delta = 0.5$)	wMAE ($\delta = 0.25$)
Mean-updated	2.337	2.346	2.366	2.383
Single-models	1.806	1.798	1.794	1.793
Multi-model	1.723	1.698	1.685	1.677

As expected, the higher frequency increases the accuracy of predictions. However, the improvements are relatively minor, suggesting that a few updates are sufficient. It should also be noted that the higher frequency requires more models and data and thus demands more computational and communication resources. Fig. 4 compares the prediction accuracy of models

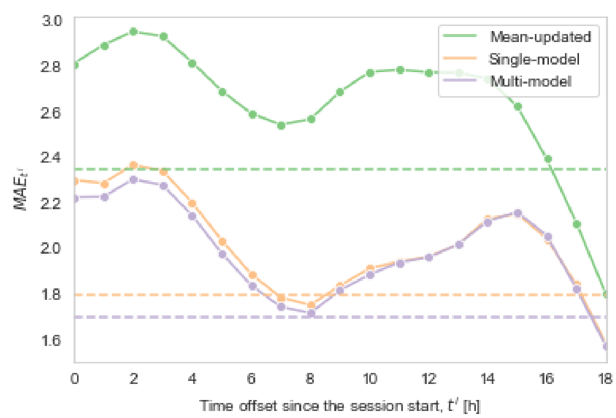

Fig. 4. Accuracy, expressed by MAE_{t^i}, as a function of the prediction time offset t^i. In all models, we used regular updates with $\delta = 1$. Horizontal lines indicate the value of wMAE for each method.

as a function of the time offset t^i since the begging of charging sessions. Due to how we processed the data, the maximum connection duration is 24 hours. Therefore, prediction errors significantly decrease when the time offset exceeds 16 hours. Again, as expected, the naive model (Mean-updated) displays significantly worse performance than advanced models (Single-model and Mutli-model). However, the overall pattern is similar. Initially, the prediction error grows and reaches the maximum value around $t^i = 2$ hours. As it can be seen in Fig. 1, a vast majority of sessions are shorter than four hours. Hence, for small values of t^i a large number of sessions contribute to the error. When short sessions terminate, the error decreases and it starts to grow again from $t^i = 8$ hours. For larger values of t^i, the difference in performance between the Single-model and Mutli-model disappears, indicating that Single-model could be sufficient.

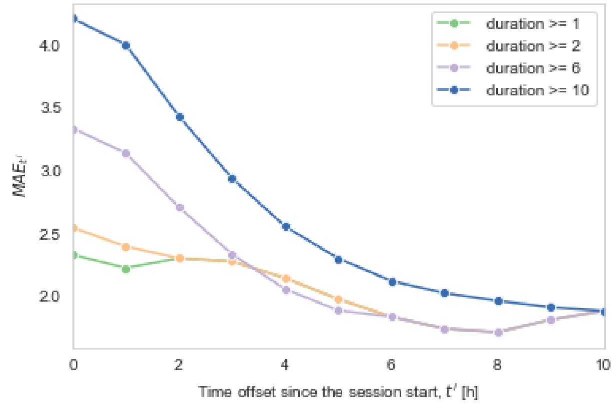

Fig. 5. Accuracy of the Multi-model, expressed by MAE_{t^i}, as a function of the prediction time offset t^i. Updates are made every hour ($\delta = 1$). Only sessions lasting more than 1, 2, 6 and 10 hours are visualised.

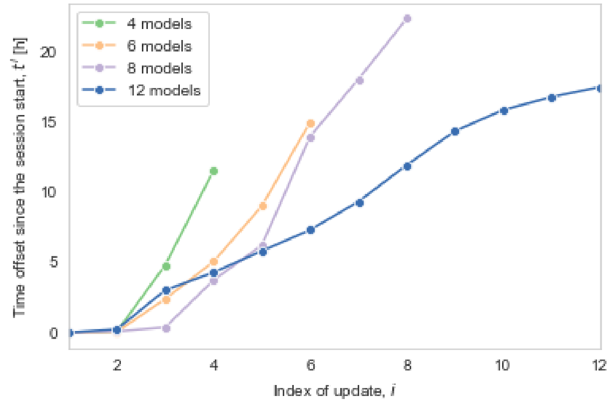

Fig. 6. The most favourable time offsets t^i at each prediction update as identified for Multi-model by the used search strategy. We considered M = 4, 6, 8 and 12 irregular updates.

In Fig. 4 we observed an increase of the prediction error for small and intermediate values of t^i. We divided the sessions into overlapping subsets based on their duration to analyse the prediction error. In Fig. 5 we analysed the values of MAE_{t^i} separately for sessions taking longer than 1, 2, 6 and 10 hours. On average, shorter sessions reach lower MAE_{t^i} values. The error mostly decreases with the time offset t^i. The decrease in MAE_{t^i} becomes smaller when the time offset t^i is larger. It is caused by sessions with longer duration that are harder to predict. Only such sessions are left when the time offset t^i is larger.

B. Irregular updates

TABLE III
COMPARISON OF REGULAR AND IRREGULAR UPDATES FOR VARIOUS NUMBERS OF USED PREDICTION MODELS M.

	Regular Multi-model	**Irregular Multi-model**
wMAE (M=4)	1.858	1.753
wMAE (M=6)	1.786	1.672
wMAE (M=8)	1.752	1.685
wMAE (M=12)	1.723	1.667

Table III compares the performance of regular and irregular updates when using the Multi-model. With the same number of models, irregular updates achieve higher accuracy of predictions. As expected, the marginal accuracy improvement vanishes by growing the number of models. By comparing the results with Table II, we observe that the regular Multi-model with 96 models ($\delta = 0.25$) gives a similar level of accuracy as irregular Multi-model just with $M = 6$ models. The obtained time offsets t^i for the irregular Multi-models from Table III are presented in Fig. 6. The frequency of prediction updates is higher for the early stages of charging sessions. Later on, the updates are approximately uniformly distributed.

C. Comparison of models

Table IV sums up how the models utilise the update strategies and pros and cons of each model.

TABLE IV
COMPARISON OF THE MODELS.

Model	**Update strategy**	**Pros and cons**
Static-model	Prediction at the time of EV's arrival.	The prediction remains unchanged for the whole duration of a charging session.
Single-model	An individual model predicts connection duration for each time offset.	Single-model improves the accuracy compared to the static-model. The initial prediction at the time of EV's arrival is less accurate.
Multi-model regular	For every time offset a different model is used. Time offsets are regularly spaced in time.	Multi-model eliminates the single-model drawback and improves the accuracy of predictions at the time of EV's arrival and also improves the overall accuracy.
Multi-model irregular	For every updating time offset a distinct model is used. Time offsets are irregularly spaced in time.	Irregular approach finds the most suitable time offsets. As a result, irregular multi-model reaches similar prediction accuracy than multi-model with regular time offsets with significantly lower number of models.

IV. CONCLUSIONS

We explored how the updates of the connection duration predictions can improve the accuracy. We prepared two naive and three advanced prediction models. From the comparison of their performance, we derived the following main conclusions:

- Regular updates significantly improve the accuracy of predictions.
- The accuracy of predictions varies when they are done with a different time offset since the beginning of the

charging session. This seems to be linked to the way how the properties of charging sessions change with their duration.
- Irregular updates further improve the accuracy of predictions. Higher frequency of prediction updates is beneficial in the early stages of charging sessions. Later on, approximately uniformly distributed updates are sufficient.

A. Limitations and future outlooks

The presented research suffers from several limitations that we will address in future research:

- We used only a single method. By utilising some other methods, e.g., neural networks, some further improvements could be achieved.
- We used only a single dataset. This dataset covers charging behaviour of EV-drivers at public slow charging locations. For DC fast chargers, home or workplace chargers the results might differ.
- The prediction updates should also be investigated with other characteristics of charging sessions than the connection duration.
- The performance of prediction updates was evaluated by accuracy measures. It could be interesting to evaluate the benefits of updated predictions in the context of smart charging schemes.

REFERENCES

[1] V. Masson-Delmotte, et al, Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC, 2021: Summary for Policymakers, 2021

[2] L. Buzna, P. De Falco, G. Ferruzzi, S. Khormali, D. Proto, N. Refa, M. Straka and G. V. D. Poel. An ensemble methodology for hierarchical probabilistic electric vehicle load forecasting at regular charging stations. Applied Energy, 283: 116337,2020.

[3] Y. Chung, B. Khaki, T. Li, C. Chu and R. Gadh. Ensemble machine learning-based algorithm for electric vehicle user behavior prediction. Applied Energy, 254:113732, 2019.

[4] Z. J. Lee, T. Li and S. H. Low, ACN-Data: Analysis and applications of an open EV charging dataset. Proceedings of the Tenth ACM International Conference on Future Energy Systems. 2019.

[5] O. Frendo, N. Gaertner and H. Stuckenschmidt, Improving smart charging prioritization by predicting electric vehicle departure time.IEEE Transactions on Intelligent Transportation Systems, 2020.

[6] A. Jaech, B. Zhang, M. Ostendorf and D. S. Kirschen, ”Real-Time Prediction of the Duration of Distribution System Outages,” in IEEE Transactions on Power Systems, vol. 34, no. 1, pp. 773-781, Jan. 2019.

[7] R. Li, F.C. Pereira and M.E. Ben-Akiva, Overview of traffic incident duration analysis and prediction. Eur. Transp. Res. Rev. 10, 22 (2018).

[8] F. Corman and L. Meng, A review of online dynamic models and algorithms for railway traffic management. IEEE Transactions on Intelligent Transportation Systems, 16(3):1274–1284, 2015.

[9] J. Lessan, L. Fu and Ch. Wen, A hybrid Bayesian network model for predicting delays in train operations, Computers & Industrial Engineering, (Vol. 127, pp. 1214-1222), 2019.

[10] P. Kecman and R. Goverde, ”Online data-driven adaptive prediction of train event times.” IEEE Transactions on Intelligent Transportation Systems 16.1 (2014): 465-474.

[11] ElaadNL, https://www.elaad.nl/, Accessed: 2021-06-25, 2021.

[12] M. Straka, R. Carvalho, G. V. D. Poel and L. Buzna, Analysis of Energy Consumption at Slow Charging Infrastructure for Electric Vehicles, IEEE Access, vol. 9, pp. 53885-53901, 2021.

[13] A. Lucas, R. Barranco and N. Refa, EV idle time estimation on charging infrastructure, comparing supervised machine learning regressions.Energies, 12(2):269,2019.

[14] G. Ke, et al, ”Lightgbm: A highly efficient gradient boosting decision tree.” Advances in neural information processing systems 30 (2017).

[15] T. Hastie, R. Tibshirani and J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 359-361). New York: springer, 2009.

[16] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach. Pearson,2009.

[17] F. Hutter, L. Kotthoff and J. Vanschoren, Automated machine learning: methods, systems, challenges. Springer Nature, 2019.

[18] J. Bergstra, et al, ”Algorithms for hyper-parameter optimisation.” Advances in neural information processing systems 24 (2011).

[19] S. Cleger-Tamayo, J.M. Fernández-Luna and J.F. Huete, On the use of Weighted Mean Absolute Error in Recommender Systems, 2012.