2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) | 979-8-3503-7800-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/METROXRAINE62247.2024.10796291

An Explainable Model Framework for Vehicle-to-Grid Available Aggregated Capacity Prediction

Luca Patanè ¶**, Francesca Sapuppo†**§, Antonio Comi‡††, Giuseppe Napoli*‡‡, Maria Gabriella Xibilia ‖**
**Dipartimento di Ingegneria, Università degli Studi di Messina, Messina, Italy
††Dipartimento di Ingegneria dell'Impresa, Università degli Studi di Roma "Tor Vergata",Roma, Italy
‡‡Consiglio Nazionale delle Ricerche (CNR), Istituto di Tecnologie Avanzate per l'Energia (ITAE), Messina, Italy
ORCID: ¶ 0000-0002-5488-9365, † 0000-0001-8772-2759, ‡ 0000-0001-6784-1638, * 0000-0003-0730-8641, ‖ 0000-0001-7723-2051

***Abstract*—Vehicle-to-grid (V2G) technology has proven to be a promising solution for integrating electric vehicles (EVs) into the electricity grid, offering benefits such as grid stabilization and demand response. Predicting the aggregate available capacity (AAC) of EVs is crucial for effectively utilizing their energy storage capabilities. Here, a comprehensive methodological framework for predicting AAC in V2G systems is presented. It mainly includes data preprocessing and feature selection methods tailored to manage complex datasets with multiple data sources such as GPS, weather, vehicle characteristics, historical data, and calendar information. In addition, data augmentation methods are presented to address the problem of data scarcity that is typical of EV infrastructures. The core of such a framework then focuses on interpretable predictive models based on explainable machine learning or a state-space representation. The discussion on the framework under development aims to highlight the importance of interpretable models in V2G systems and provide insights into future research directions for such a prominent area, considering the evolution of the energy sector.**

***Index Terms*—Vehicle-to-grid, data-driven predictive model, time-series prediction, machine learning, interpretable models, explainable AI**

I. Prominence of the Topic Considering the Evolution of the Energy Sector

At the end of 2021, the number of electric cars on the road exceeded 16.5 million. The global stock of electric cars will expand to nearly 350 million vehicles by 2030 under the International Energy Agency's NZE scenario [1]. This growth will have a profound impact both from the point of view of the management of the electricity grid and from the infrastructural one and future developments will depend on the ability to adapt the whole management system. Just as they draw electricity, electric vehicles (EVs) can potentially supply it to the grid through Vehicle to Grid (V2G) technology which allows, through the bidirectional use of charging devices installed on board the vehicle, to offer services such as balancing supply and demand and participation in flexibility markets. The presence of electric vehicles in the energy markets, as suppliers, both directly and through aggregators, will make it possible to fuel a renewable cycle for more substantial decarbonization. The network manager, based on the needs for frequency regulation or production/consumption balancing, can interact with market operators, who, in turn, act on the resources available, including vehicle charging stations. The growth of the electric vehicle market, and above all the possibilities linked to the flexibility of use and management of the energy content available on-board vehicles, may represent a further step for the definitive paradigm shift towards the evolution of the sector. A significant advantage for the grid cannot be represented by a single vehicle, therefore EVs are generally grouped into fleets or aggregators [2], [3]. The aggregator is an entity that brings together several EVs to generate a significant impact on the electricity grid by acting as an interface between the system operator and the connected fleet. Fundamental to the success of participation in these markets is the ability to predict potential power flows entering or exiting the electricity grid. Assuming a set of vehicles (and therefore batteries) in motion and geographically dispersed, this aspect is of primary importance to guarantee, for example, the quantities of power or energy contracted at a given moment. Managing an aggregation service of this type is complex and places the vehicle user at its center, whose behaviour and decisions dictate a vehicle's willingness to participate. End users can be encouraged to modify their demand profile to help meet the requirements, for example, through Demand Side Management [2], [4].

This work was funded by the MASE - **Consiglio Nazionale delle Ricerche** within the project RICERCA DI SISTEMA 22-24 -21.2 Progetto Integrato Tecnologie di accumulo elettrochimico e termico. CUP Master: B53C22008540001, UNIME-DI-RdS22_24: J43C23000670001 Linea 13

§ Corresponding Author

II. V2G Explainable Predictive Model

For V2G systems to be reliable and economically viable for providers, accurate prediction of aggregated available capacity (AAC) associated with specific aggregator hubs is required. Furthermore, the prediction of AAC in V2G energy storage control may not be useful to an energy supply professional as it could be considered unreliable. Therefore, it is of great interest to explain the rationale behind the predicted AAC and provide the energy expert with an understanding of the predicted dynamics. Insights into prediction come primarily from understanding the availability of EVs to deliver energy via a V2G aggregator when needed. This availability can be influenced by various factors: user behavior, location of charging infrastructure [5], and vehicle characteristics. Meteorological data as well as calendar information (weekends and holidays) can also influence the accuracy of AAC predictions. Meteorological factors such as weather conditions and temperature fluctuations can play a decisive role in determining the usage behavior of EVs. For example, unfavorable weather conditions can affect driving behavior [6] and thus the availability of EVs for grid services. In addition, weekends and national holidays lead to different time patterns in EV usage, as travel habits and energy consumption behavior often deviate from typical weekday routines [7].

In the literature, machine learning (ML) models, i.e. Persistence model, generalized linear model, artificial neural networks (ANNs) [8], long short-term memory network (LSTM) [9], [10], K-Means Clustering [11], and Convolutional Neural Networks (CNN) [12], were used to predict AAC based on different types of data sources: EVs fleets with a limited number of vehicles [13], [14], aggregated availability based on charging point data [8], historical vehicle information [15], calendar [11]. Very little literature presents the usage of mobility pattern [16] and weather information [8] oriented to V2G AAC. Most of the ML-based predictive algorithms are black-box, and, to the authors' knowledge, model based on the state space representation have not been used in such a domain. A further obstacle encountered with EVs and EV Infrastructure (EVI) data streams include data scarcity, imbalanced data distribution, insufficient volume of complete datasets, and occurrences of missing or inaccurate values. These challenges are effectively tackled utilizing data augmentation methodologies based on artificial intelligence (AI) [17] and data expansion based on statistics [18].

In this scientific and applicative environment, we propose a comprehensive methodological framework that includes a data integration phase, as the diverse nature of data sources requires standardization for predictive modeling. The data includes real GPS-based mobility of vehicles from available databases, meteorological information as well as weekend and holiday data. The impact of these data on AAC can be preliminarily analyzed using exploratory data analysis. Once data sources are acquired, selected, and validated, data preprocessing is necessary to integrate and standardize the overall data format. Available capacity aggregated in the spatial and temporal domains is based on mobility data, while fuzzification of inputs combining holiday and weekday information is applied. Additionally, data expansion and augmentation are proposed to address the common issue of data scarcity. A set of advanced models classes, including ML and linear and nonlinear identification methods, are proposed for AAC prediction over different time horizons. Emphasis is placed on the interpretation and explanation of the models to gain a more comprehensive understanding of the factors influencing the predictions [19]. Finally the expected outcomes from the proposed methodological framework are presented.

III. Data Collection

A preliminary step in designing predictive models leverages the selection of existing datasets, integrates real supplementary data such as weather and calendar information, and performs exploratory data analysis and feature selection.

A. Vehicle Dataset

1) Veneto Floating Car Data/Automated Vehicle Monitoring Dataset: The study considers car mobility in the Veneto (Italy) region using Floating Car Data(FCD)/Automated Vehicle Monitoring(AVM) data, which meets certain criteria for being classified as big data, including its enormous volume and by-product nature. From the first to the last automobile journey of the day, information on car trips inside the Veneto area (i.e., at least one survey datum inside the region on the survey day) is included in the data collection. In order to get information on the journeys made, the data were examined to find travel trends. Five working-day observations, distributed during the autumn months (October–November 2018) on various working days, make up the accessible database. The information form includes basic vehicle data for each sampled vehicle, including vehicle class, brand, year, type, fuel type, and gross weight. The daily car operation logs include the following information about every trip the surveyed vehicle makes in chronological order: the vehicle's unique identification number, the date and time the record was logged, the coordinates (latitude and longitude) of the location, the instantaneous speed, the kind of road (urban, ex-tra-urban, motorway), and the direction angle. Following a thorough cleaning and removal of observations containing incomplete data, the remaining data were processed to look into trip patterns and duration. 29,158 cars in total were examined, which translates to around 70,000 trip chains completed in five days.

2) The Open-Access Vehicle Energy Dataset (VED): The Vehicle Energy Dataset (VED) [20] is a freely accessible dataset that is available for this study. It consists of fuel and energy-related information collected from 383 individual vehicles in Ann Arbor, Michigan, USA. This dataset contains GPS records of vehicle routes and time series data on fuel consumption, energy consumption, speed and auxiliary energy use. The dataset covers a wide range of vehicles: 264 internal combustion engines (ICEs), 92 hybrid electric vehicles (HEVs), and 27 plug-in hybrid electric vehicles/electric vehicles (PHEV/EVs), operating in real-world conditions from November 2017 to November 2018. Specifically, the following features are selected: date, vehicle identifier, trip identifier, duration, latitude, and longitude. In addition, the HV Battery State of Charge (SoC[%]) is provided for the PHEVs and EVs. The strength of this dataset lies in its provision of an entire year's worth of data, enabling both interperiod and intraperiod analyses. This enables the analysis of weekly and seasonal periodicity, as well as dependencies on varying vacation and work/school periods.

B. Meteo Dataset

The meteorological data is a freely accessible dataset from the MeteoStat database using the Python API [21] based on GPS coordinates of the geographical area under study. The information on precipitation, temperature, wind speed, humidity, overcast percentage, etc are available on an hourly basis for the period under investigation.

C. National Holidays Dataset

State office closings for state holidays are additional information to be integrated into the input dataset. In particular, they are considered in conjunction with weekend information to obtain a comprehensive holiday rate to be incorporated into the model.

IV. Data Pre-Processing, and Analysis

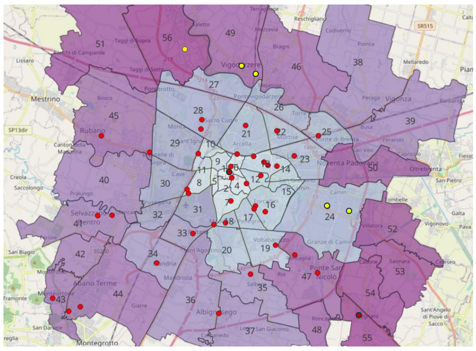
Data pre-processing aims to integrate different data sources and standardize the overall data format. It includes cleaning, normalization, spatial and temporal aggregation and feature engineering to extract relevant information for AAC prediction. Advanced techniques such as feature selection and dimensionality reduction are presented to enhance model performance and interpretability.

A. Vehicle Dataset

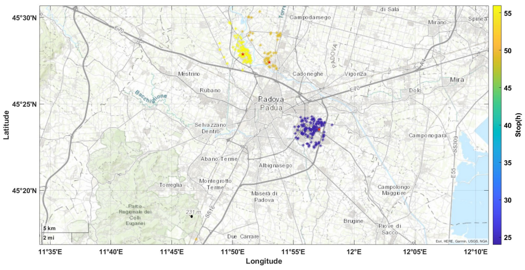
Datasets are processed to obtain the AAC and other aggregate information in space and time domain with a parametric T-hour (Th) interval time. The first step is to identify V2G locations, considering the fact that V2G services rely on the geographical characterisation of trips undertaken by classifying them by place (zone) of destination. Moving from one location (origin) to another (destination) (O-D) to complete one or more tasks is referred to as a trip. The widespread use of telematics allows disclosing car trips with FCD, which enables continuous vehicle tracking over time and location.

1) Trip Chain and Stops Extraction: Car trip detection phase seeks, based on predetermined guidelines, to identify the activity stops made by every sampled vehicle traversing the research area, especially for fully electric vehicles. As a result, the specific journeys (including O-D) made by resident and non-resident vehicles are acquired. Moreover, the starting and ending times of all travels, O-D locations, battery level at each destination, and additional travels to be completed by each stop are all relevant details for each examined trip. The vehicle position (GPS coordinates and time) and status (travelling or stopping) are disclosed with reference to the prober vehicle datum. Two consecutive data points from a given vehicle are used to identify the start and end of a trip as well as detect major changes in the vehicle's position based on pertinent status information. We define a journey or trip chain as a series of trips that are sequenced so that the destination of one trip coincides with the origin of the next one. As a result, the activity stops must be determined from the fine-grained FCD. Using the EVs as an example, it is possible to forecast the battery level and the additional routine tasks carried out following each stop.

2) Stop Maps and Aggregator Selection: In the land-use analysis, zoning is carried out in accordance with the study area data, and possible sites for the implementation of V2G services are identified, such as parking near movie theatres or retail centres, or parking at places of employment. They are referred to ad points of interest (PoIs) or hubs for V2G aggregators. Figure 1 shows a preliminary analysis of the Veneto mobility data, showing in (a) the identified zones and candidate PoIs for V2G aggregators. Given the density of vehicle stops, three zones (24,49,56) and the included PoIs are selected as the best candidates. Figure 1(b) shows the stop locations in the selected zones on one of the sample days.



(a) Zone division and candidates as PoIs for V2G aggregators: (red) candidates, (yellow) selected PoIs.



(b) Stop maps for three PoIs for V2G aggregators in Zone 24 (blue), Zone 49(orange), Zone 56(yellow) over a sample day under study

Fig. 1. Padua map for V2G aggregators PoIs and zone selection

3) Data Augmentation and Expansion: EV data scarcity is one of the main issues in AVM/FCD data for V2G variable forecasting. Indeed, the issue of estimating O-D demand flows depends on the sampling strategies used. Different approaches are presented to expand available datasets: ML-based data augmentation and statistic-based expansion to the universe of investigation. The first is used to simulate a larger EV fleet mimicking a predefined data distribution to produce novel samples within that distribution. One well-known algorithm is the Generative Adversarial Network (GAN) [22]: it is a neural model rooted in generative architecture involving a competitive interplay between two neural networks (NN). The latter approach is based on the empirical representativeness of the sample both in terms of trip production and attraction [18]. From FCD/AVM dataset, an estimate of the current travel demand can be made once the sample O-D matrices for various survey days or times have been established. Then, it becomes necessary to have knowledge of the sample units, which are cars, as well as the technique for counting the population universe (such as lists of registered cars in a traffic zone or counts of passing cars). It is assumed a stratum sampling procedure, i.e., the total statistical population (vehicles) is divided into K classes (e.g., vehicles registered in a province) or strata. The sample representativeness is then used as a weight to extend the analysis results to the universe of investigation.

4) V2G Activity and State of Charge (SoC) Simulation: Information on electrical features such as the SoC is only available for PHEVs and EVs. As a preliminary step, a simplifying assumption can be made: the ICEs, and HEVs are assumed to be EVs contributing to the V2G activity. For the VED dataset the adopted equivalent model is the Nissan Leaf with 40kWh battery capacity. SoC for ICEs and HEVs is calculated as an indirect measure of distance traveled and charging stop intervals. In particular charge calculation is performed with different methods and under specific hypotheses [12]: $SoC = 100\%$ at the beginning of each simulated day, the minimum state of charge that must be maintained is set as a fixed value ($SoC_{min} = 30\%$) or obtained as the charge necessary to cover the remaining part of the travel chain. Other important parameters are necessary in order to calculate $SoC\%$ at each step of the trip: the energy consumption per kilometer traveled by a vehicle, the discharge rate during the V2G plug in. Other parameters that are taken into account are the rapid- charging hour rating of a vehicle, typically using DC power in the daytime (from 7 a.m. to 7 p.m.), the slow-charging hour rating of a vehicle during the nighttime (between 7 a.m. and 7 p.m.), the efficiency of the charging process taking losses into account, power rating of the export to the grid.

5) Data Aggregation: The mapping of trips and stops can be used to create an aggregated dataset in space, related to the V2G aggregator zone, and in time with a parametric Th interval to obtain a time series that feeds into the dynamic prediction model. The real or simulated SoC_v is used to determine the Available Capacity of a vehicle (AC_v). It is defined as the capacity of each vehicle to provide energy to the grid in a half-hour (Th) period.

$$AC_v^{Th} = Max(SoC_v^{Th-1} - SoC_{min}, 0) * BC \qquad (1)$$

where the BC is the battery capacity. Equation 1 is considered to identify available vehicles. The aggregation in space refers to the assumption that the vehicles parked within a specific radius r from a selected V2G hub would connect to it. A vehicle parked within r from the PoI in a T-hour interval, and respecting the SoC_{min} requirement, is considered available (av^r) to feed energy into the V2G system in that interval. The target feature to be predicted is the Aggregated Available Capacity ($AAC_{PoI^r}^{Th}$) in a half-hour interval and within r distance from the hub. It is defined in the following equation:

$$AAC_{PoI^r}^{Th} = \sum_{av^r} AC_v^{Th} \qquad (2)$$

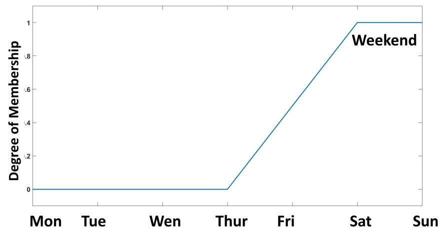
Additional considered aggregated time series are the Mean Hub Distance (MHD_{PoIC}^{Th}) as the mean distance from the hub center C of all vehicles parked in the global area of interest contained in the dataset.

B. Meteo Dataset

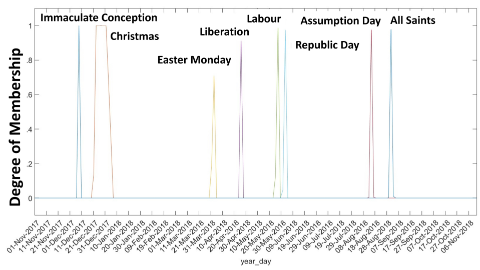
The pre-processing of the meteo information includes the imputing of the missing data: they are replaced by the average of the signal in the same week or by the regression between the neighbouring data. The meteo time series must be resampled to the time interval Th.

C. National Holidays Fuzzy Set

The information about national holidays and weekends is represented by a discontinuous time series unsuitable for dynamic models. The objective of the fuzzification of such inputs is to obtain a continuous time series including both weekend and national holiday information. Fuzzy membership functions are generated by considering the effect that weekends and holidays could have on drivers' habits in the previous and successive days. The national holidays information must be related to the geographical area of interest. An example of the membership functions for the weekend and holidays in Italy related to the Veneto dataset are shown Figures 2(a) and (b) respectively. Applying the membership functions to time interval of interest, the maximum value is taken to obtain a single continuous dynamic value comprehensive for both information.



(a) Weekend



(b) National Holidays

Fig. 2. Membership function used for the fuzzification of the holiday rate when the Veneto dataset is considered.

D. Feature Selection

ML-based feature selection represents a crucial step in developing efficient ML predictive models by identifying the most relevant features from a dataset. In addition to explorative data analysis based on the concepts of correlation and mutual information, examples of algorithms are: minimum redundancy maximum relevance (MRMR) [23] using mutual information for computing relevance and redundancy among variables (features), Relieff [24] estimating the quality of attributes according to their capability of distinguishing between instances that are near to each other, F-Test examining the importance of each predictor individually using an F-test, and then ranking features using the p-values of the F-test statistics. Further studies on feature selection are conducted to provide insights not only on features that have a greater impact on prediction but also on optimizing data preprocessing methods and the structure of applied predictive models.

V. Models

The proposed methodological framework, based on preprocessed and standardized data structures, is designed to integrate different classes of models for the prediction of AAC. Emphasis is then placed on the interpretability of the models in order to gain a more comprehensive understanding of the factors influencing the predictions. The regression models are based on the theory of dynamic system identification: they use the inputs as time series sampled by Th and AAC_{PoI^r} as the sequence output.

A. Machine Learning Models

Black-box machine learning techniques are considered suitable for the AAC prediction, in particular neural networks, long short-term memory (LSTM) networks, ensemble trees, Gaussian process regression and trees. To increase the explainability of the prediction, it is necessary to add a further layer based on explainable artificial intelligence (XAI) [25], [26]. The optimization of the model and the selection of hyperparameters are performed using optimization algorithms such as Bayesian algorithms or grid search methods. In addition, k-fold validation or regularization techniques are considered to avoid overfitting in the learning process.

B. Linear Regression and State Space Representation Models

Models such as Autoregressive Models with eXogenous Inputs (ARX) and Finite Impulse Response (FIR) models are considered. Furthermore, we investigate exploring the potential application of Hankel dynamic mode decomposition with control (HDMDc) [27], a data-driven technique to extract spatiotemporal patterns and dynamics from AAC data. HDMDc can be applied to V2G datasets by considering the AAC as the state of the system, the other aggregated data, the meteo and the calendar information as exogenous inputs. The advantage of a state space representation lies in the clear relationships between the state, the output and the exogenous inputs.

C. Nonlinear Modeling

We consider the utilization of block modeling to account for nonlinearities, including methods based on Wiener-Hammerstein models. It would allows us to highlight both static/dynamic, and linear/nonlinear contributes involved in predicting the AAC.

D. Model Interpretability

Recent literature and applications have introduced interpretation techniques for complex machine learning algorithms, particularly deep learning (DL), to improve model interpretability while maintaining model complexity. Examples of such methods, referred to as XAI, include local interpretable model-agnostic explanation (LIME) [28] and Shapley additive explanations(SHAP) [29]. The resulting rating values, provided as output of these methods, represent the contribution of each feature to the result with respect to the output of the baseline model $E[f(x)]$.

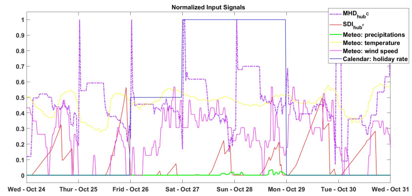
VI. Model Identification and Expected Outcomes

According to the dataset, different approaches need to be be used for selecting the training, validation and test sets. In the case of the VED dataset, which covers an annual interval, 3 weeks per month can be selected for the training/validation phase and 1 week for the test phase to avoid seasonal distortions and unbalanced datasets, whereby holiday weeks must be included in both the training/validation and the test phase. The Veneto dataset contains 5 days for three different PoIs, cross-validation techniques can be considered for model hyperparameters optimization. The concept of model transferability can be applied by training models on selected PoIs and subsequently deploying them on different PoIs. Predictions for various time horizons can be implemented using two distinct approaches. The first approach involves identifying and optimizing the model for one-step-ahead prediction and then iterating this process over the desired time horizon. The second approach entails identifying and optimizing the model by incorporating a delay in the target value corresponding to the prediction horizon.

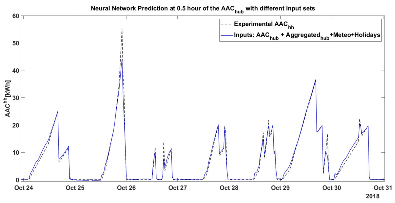
A proof of concept for predictions on the VED dataset for the test week of Wednesday, October 24 to Wednesday, October 31, 2018, is presented in Figure 3. The figure includes: (a) the input time series normalized to a range of 0-1, and (b) an example of 0.5-hour ahead prediction of AAC using neural networks (NNs). The holiday rate exhibits positive values coinciding with the weekend, which strongly correlates with a decrease in AAC. An explanation of a specific query point (October 25, 2018, 22:00) using the SHAP technique is shown in Figure 3(c).

VII. Conclusions

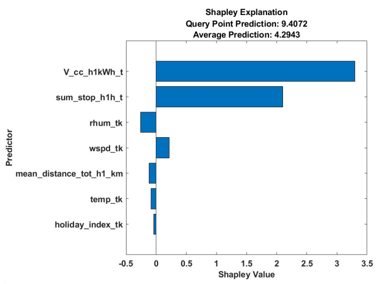
Predicting AAC in V2G systems is a critical task with significant implications for grid management and energy optimization. As final goal of the research, the development of a methodological framework with integrated datasets and interpretable models is presented, which offers a promising way to improve the transparency and trustworthiness of AAC prediction in the energy managment field. The versatility of our forecasting framework extends to different mobility datasets, as demonstrated by the VED and Veneto datasets, and enables the strategic positioning of aggregator hubs. This optimization facilitates the efficient management and utilization of geographically distributed EVs and improves the scalability and adaptability of the V2G infrastructure. An innovative contribution of the proposed framework to the current literature is the integration of geographical, meteorological and calendar variables through a fuzzy input set that incorporates national holidays and weekends. By integrating these contextual components, our framework provides a more comprehensive explanation of the intricate relationships within the V2G system. The proposed methodological framework emphasizes the importance of data pre-processing and the potential application of advanced techniques such as linear regression, HDMDc and ML models, with a focus on the interpretability of the provided prediction.



(a) Predictors



(b) 0.5-hour ahead prediction with NN



(c) Shapley Values explanation of time-series query points (October 25th 2018, 22:00)

Fig. 3. Prediction and SHAP explanation for AAC_{PoI_r} for the week in the test dataset Oct. 24th to Oct. 31st, 2018

References

[1] I. E. Agency, "Net zero by 2050. a roadmap for the global energy sector." in *IEA, Paris*, 2021.
[2] K. D. Afentoulis, Z. N. Bampos, S. I. Vagropoulos, S. D. Keranidis, and P. N. Biskas, "Smart charging business model framework for electric vehicle aggregators," *Applied Energy*, vol. 328, p. 120179, 2022.
[3] M. Barbero, C. Corchero, L. Canals Casals, L. Igualada, and F.-J. Heredia, "Critical evaluation of european balancing markets to enable the participation of demand aggregators," *Applied Energy*, vol. 264, p. 114707, 2020.
[4] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Trans. Ind. Inform*, vol. 7, no. 3, pp. 381–388, 2011.
[5] J. Dixon, W. Bukhsh, K. Bell, and C. Brand, "Vehicle to grid: driver plug-in patterns, their impact on the cost and carbon of charging, and implications for system flexibility," *eTransportation*, vol. 13, 2022.
[6] V. Bakhshi, K. Aghabayk, N. Parishad, and N. Shiwakoti, "Evaluating rainy weather effects on driving behaviour dimensions of driving behaviour questionnaire," *J. Adv. Transp.*, pp. 1–10, 3 2022.
[7] T.-H. T. Gim, "Sem application to the household travel survey on weekends versus weekdays: the case of seoul, south korea," *European Transport Research Review*, vol. 10, p. 11, 3 2018.
[8] J. Graham and F. Teng, "Vehicle-to-grid plug-in forecasting for participation in ancillary services markets," in *2023 IEEE Belgrade PowerTech*, jun 2023.
[9] F. Curreri, L. Patanè, and M. G. Xibilia, "Rnn- and lstm-based soft sensors transferability for an industrial process," *Sensors*, vol. 21, no. 3, 2021.
[10] L. Patanè, F. Sapuppo, , and M. G. Xibilia, "A comprehensive data analysis for aggregate capacity forecasting in vehicle-to-grid applications," in *2024 ICCAD*, Paris (France), Oct. 2024, p. 6.
[11] D. Perry, N. Wang, and S.-S. Ho, "Energy demand prediction with optimized clustering-based federated learning," in *2021 IEEE GLOBECOM*, 2021, pp. 1–6.
[12] R. Shipman, R. Roberts, J. Waldron, C. Rimmer, L. Rodrigues, and M. Gillott, "Online machine learning of available capacity for vehicle-to-grid services during the coronavirus pandemic," *Energies*, vol. 14, no. 21, 2021.
[13] R. Shipman, R. Roberts, J. Waldron, S. Naylor, J. Pinchin, L. Rodrigues, and M. Gillott, "We got the power: Predicting available capacity for vehicle-to-grid services using a deep recurrent neural network," *Energy*, vol. 221, p. 119813, 2021.
[14] H. S. Nogay, "Estimating the aggregated available capacity for vehicle to grid services using deep learning and nonlinear autoregressive neural network," *SEGAN*, vol. 29, p. 100590, 2022.
[15] S. Li, C. Gu, J. Li, H. Wang, and Q. Yang, "Boosting grid efficiency and resiliency by releasing V2G potentiality through a novel rolling prediction-decision framework and deep-lstm algorithm," *IEEE JSYST*, vol. 15, no. 2, pp. 2562–2570, 2021.
[16] T. Zeng, S. Moura, and Z. Zhou, "Joint mobility and vehicle-to-grid coordination in rebalancing shared mobility-on-demand systems," vol. 56, no. 2, pp. 6642–6647, 2023, 22nd IFAC World Congress.
[17] V. Sumanasena, L. Gunasekara, S. Kahawala, N. Mills, D. De Silva, M. Jalili, S. Sierla, and A. Jennings, "Artificial intelligence for electric vehicle infrastructure: Demand profiling, data augmentation, demand forecasting, demand explainability and charge optimisation," *Energies*, vol. 16, no. 5, 2023.
[18] A. Comi, A. Rossolov, A. Polimeni, and A. Nuzzolo, "Private car O-D flow estimation based on automated vehicle monitoring data: Theoretical issues and empirical evidence," *Information*, vol. 12, no. 12, 2021.
[19] L. Patanè, F. Sapuppo, G. Scipilliti, and M. G. Xibilia, "Model-agnostic soft sensor interpretability," in *2023 MetroXRAINE*, 2023, pp. 828–833.
[20] G. Oh, D. J. Leblanc, and H. Peng, "Vehicle energy dataset (VED), a large-scale dataset for vehicle energy consumption research," *IEEE Trans Intel Trans Sys*, vol. 23, no. 4, pp. 3302–3312, 2022.
[21] "Meteo stat phyton api," https://dev.meteostat.net/python/.
[22] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 6 2014.
[23] Z. Zhao, R. Anand, and M. Wang, "Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform," in *2019 IEEE DSAA*, 2019, pp. 442–452.
[24] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of relieff and rrelieff," *Mach. Learn.*, vol. 53, pp. 23–69, 2003.
[25] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," 2019.
[26] G. V. Aiosa, M. Palesi, and F. Sapuppo, "Explainable ai for decision support to obesity comorbidities diagnosis," *IEEE Access*, vol. 11, pp. 107 767–107 782, 2023.
[27] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic mode decomposition with control," *SIAD*, vol. 15, pp. 142–161, 1 2016.
[28] D. Garreau and U. von Luxburg, "Explaining the explainer: A first theoretical analysis of lime," in *Proceedings of the Twenty Third AISTATS*, 26–28 Aug 2020.
[29] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Adv Neural Inf Process Syst*, vol. 30, 2017.