Explainability Analysis of V2G Aggregated Capacity Prediction

Francesca Sapuppo[†**§], Luca Patanè [¶**], Antonio Comi[‡††],
Elsiddig Elnour[‡], Giuseppe Napoli[*‡‡], Maria Gabriella Xibilia[‖**]
[**]*Department of Engineering*, *University of Messina*, Messina, Italy
[††]*Department of Enterprise Engineering*, *University of Rome Tor Vergata*, Rome, Italy
[‡‡]*National Research Council of Italy*, *Institute of Advanced Technologies for Energy*, Messina, Italy
ORCID: [¶] 0000-0002-5488-9365, [†] 0000-0001-8772-2759, [‡] 0000-0001-6784-1638, [*] 0000-0003-0730-8641, [‖] 0000-0001-7723-2051

***Abstract*—The integration of Vehicle-to-Grid (V2G) systems in smart grids requires accurate and interpretable predictive models to estimate Aggregated Available Capacity (AAC) and optimize energy flows. This study presents a proof of concept for explainable AI (XAI) methodologies applied to AAC prediction, leveraging a dataset of vehicle mobility patterns, meteorological conditions, and calendar effects in the metropolitan area of Rome. Both linear and nonlinear data-driven models were evaluated, considering FIR and nonlinear FIR using exogenous inputs, and ARX and nonlinear ARX considering autoregressive and exogenous input selections. Black-box nonlinear models were explained using Shapley Additive Explanations (SHAP). Results highlight the importance of exogenous variables in improving AAC prediction for the different intervals of the day. The findings emphasize the need for interpretable models in decision support systems for energy providers and V2G infrastructure managers, ensuring reliable and transparent decision-making.**

***Index Terms*—Vehicle-to-Grid, data-driven predictive model, system identification, machine learning, explainable artificial intelligence, interpretability**

I. Introduction

Explainable Artificial Intelligence (XAI) has emerged as a critical field for enhancing the interpretability and transparency of Machine Learning (ML) models in decision-making systems. In the context of Vehicle-to-Grid (V2G) systems, predictive modeling of Aggregated Available Capacity (AAC) is crucial for optimizing energy flows and ensuring grid reliability. However, while XAI methods, such as SHAP (Shapley Additive Explanations) [1] and LIME (Local Interpretable Model-agnostic Explanations) [2], have been extensively studied in domains like healthcare and finance, their application to V2G AAC prediction remains underexplored.

A review of the literature for the application of XAI for Smart Grid [3] reveals that predictive models for smart energy management often use meteorological data for renewable energy forecasting, such as photovoltaic (PV) energy generation [4], but rarely for user behavior modeling in electric vehicle (EV) charging and driving habits. Similarly, calendar effects, including weekends and holidays, are commonly integrated into energy load forecasting but have not been comprehensively studied for their impact on V2G AAC prediction using fuzzy representations [5]. In addition, literature [6] is present that includes analysis of mobility and fuel consumption analysis not in the V2G environment.

Existing studies have focused on using advanced ML models for energy demand forecasting and smart grid optimization [3], [7]–[10]. However, they often lack interpretability, limiting their adoption in decision support systems (DSS) for energy providers. XAI methodologies like SHAP and LIME provide actionable insights into feature contributions and model decisions, making them highly suitable for complex scenarios like V2G AAC prediction, where analyzing the influence of meteorological and calendar data can provide valuable insights and enhance predictive accuracy [11], [12].

This work presents a proof of concept for the application of XAI techniques in AAC prediction for V2G systems, addressing such a gap in the literature. Additionally, it explores the integration of meteorological and fuzzy holiday data for AAC prediction, extending their application beyond traditional energy forecasting domains. The proposed framework is validated using Floating Car Data (FCD) within the Latium region in Italy, demonstrating its practical applicability and scalability for smart grid management.

II. Data Collection and Analysis

Data pre-processing integrates multiple sources, standardizes formats, and applies cleaning, normalization, aggregation, and feature engineering for AAC prediction.

A. Vehicle Dataset and AAC

1) Data Collection: The data consist of information on car journeys performed by a sample of vehicles driving within the Latium region (that is, at least one survey data within the region on the day of the survey) from the first to the last trip made during the whole day. The data have been analysed to identify travel patterns, thus obtaining indications on the trips performed. The information form includes basic vehicle data for each sampled vehicle, including vehicle class, brand, year,

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 and LA2.12-Analisi dell'offerta territoriale per la realizzazione di modelli di predizione della capacità aggregata fornita da veicoli elettrici a supporto delle esigenze della rete elettrica, CUP E87H23001620005.

[§] Corresponding Author

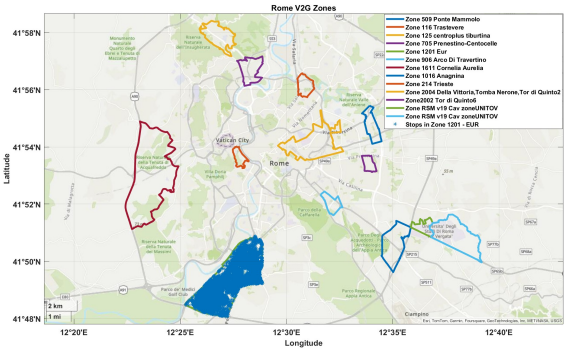
type, fuel type, and gross weight. The vehicle identifier, date (the day the record is logged), timestamp (the time the record is logged), coordinates (the geographic location: latitude and longitude), instantaneous speed, type of road (urban, extra-urban, freeway), and direction angle are all included in the daily car operation logs, which list all trips the surveyed vehicle made in chronological order. There is no information available regarding the nature of the activity performed or the trip purpose (e.g., work or study) of the surveyed cars. A subset of 661 internal combustion engine vehicles, active under real-world conditions during 2023 and circulating in a specific area of interest (EUR, Rome), was analyzed. Vehicles were classified following established procedures [13] and virtually electrified by assigning the determined values for battery pack energy and energy consumption. This enabled the analysis of virtual Battery Electric Vehicles (BEV) equivalents. The available database consists of 58 days spread over four time intervals (February 15–28, May 31–June 15, July 12–25, and September 27–October 10) capturing seasonal and calendar-related variations.

2) Metropolitan Area Zoning and Hubs Selection: An initial analysis of the mobility data was carried out by plotting the stops in the selected time intervals on the global map to identify geographical points suitable as V2G hubs. Figure 1 shows a preliminary analysis of the mobility data and the density of stops. The classification of areas of interest was based on their primary functions or designated land use. The EUR zone in Rome was selected: it is primarily designated as a business district and public office area, with commercial land use being the predominant category.

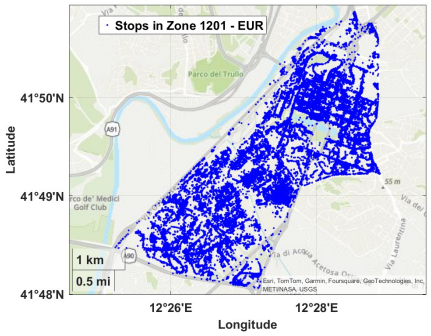
3) Trip Chain and Stops Extraction: The car trip detection phase identifies activity stops for each sampled vehicle based on predefined guidelines, particularly for BEV equivalents. It extracts origin-destination (O-D) trips for resident and non-resident vehicles, including travel times, O-D locations, battery charge at stops, and subsequent trips. GPS coordinates and vehicle status (traveling or stopped) determine trip start and end points by detecting significant position changes. A trip chain is defined as a sequence of trips where each destination aligns with the next trip's origin. This fine-grained FCD analysis enables determining the battery levels and subsequent vehicle activities after each stop.

4) V2G Activity and State of Charge (SoC) Simulation: Based on the virtual electrification of the fleet, SoC is estimated on distance traveled and charging stops, assuming an initial 100% SoC and a minimum threshold SoC_{min}, based on a constant (30%) or on the charge required to complete the trip chain [14]. Key parameters, provided by the virtual electrification process, include energy consumption per kilometer, V2G discharge rates, charging efficiency, and power export ratings, with DC fast charging during the day and slow charging at night.

5) Spatial and Temporal Aggregation: Data Trips and stops are mapped to generate a spatiotemporal dataset for V2G aggregation, using a parametric Th interval to create time series for dynamic prediction. The available capacity of each



(a) Rome Zones



(b) Rome Zones

Fig. 1. Rome V2G zone candidates: (a) entire metropolitan area, (b) stops in the EUR zone in Rome.

vehicle (AC_v) in a half-hour period is derived from real or simulated SoC_v:

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

where BC is battery capacity. Vehicles within the zone of interest z of a V2G hub and meeting SoC_{min} criteria are considered available (av^z) for energy transfer. The $AAC_{PoI^z}^{Th}$ within z is:

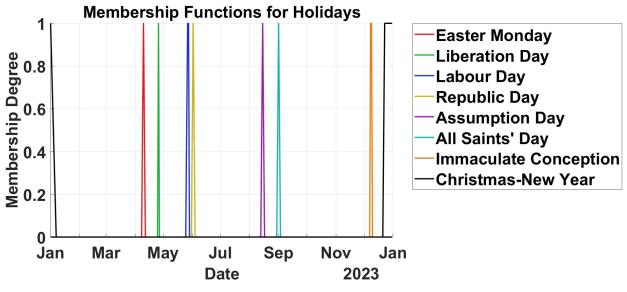
$$AAC_{PoI^z}^{Th} = \sum_{av^z} AC_v^{Th} \qquad (2)$$

B. Meteo Dataset

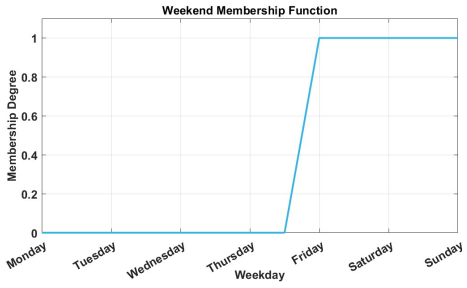
The meteorological data was extracted from the Visual Crossing Weather Service database [15]. The information on precipitation, precipitation probability, temperature, feels-like temperature, humidity, sea level pressure, visibility, cloud cover, conditions, wind speed and wind direction was extracted on an hourly basis for the period under investigation. The pre-processing involved the imputing of the missing data and the resampling to half hour period. The missing data were replaced using a moving average filter.

C. National Holidays Fuzzy Set

The information about Italian national holidays and week-ends is represented by a discontinuous time series unsuitable



(a) Holiday



(b) Weekend

Fig. 2. Membership function for the fuzzification of the holiday rate. (a) Weekend membership. (b) National holiday membership functions.

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 were generated by considering the effect that holidays and weekends could have on drivers' habits in the previous and successive days. The membership functions for the holidays and weekends, Figures 2(a) and (b) respectively, were both applied to the original dates, and the maximum value was taken to obtain a single continuous dynamic feature comprehensive for both information.

D. Explorative Data Analysis

An initial feature selection was conducted by eliminating exogenous input variables with low variability, high correlation to the other input variables or low correlation to the predicted variable. Specifically, feels-like temperature and humidity were excluded due to their high correlation with temperature, sea level pressure was removed due to its weak correlation with AAC, and visibility was discarded because of its minimal variability in the Latium area.

1) Zero-Lag Correlation: Figure 3 presents the correlation coefficients between the input features and the predicted AAC variable, offering a static overview of their relationships.

2) Dynamic Model Order Selection: Given the dynamic nature of the system, the model order was determined through correlation curve analysis at various lags between each selected input and the predicted AAC variable. This analysis guided the selection of input and output regressors, ensuring that relevant temporal dependencies were captured. AAC demonstrated correlations with past temperature values, as depicted in Figure 4, justifying the inclusion of lagged temperature values as regressors. A lag of 24 samples (equivalent to 12 hours) was selected based on the observed persistence in temperature, wind speed, and wind direction correlations, which also appeared reasonable for the holiday index. Conversely, variables exhibiting the highest zero-lag correlation (e.g., cloud cover and datetime) were included without additional regressors. The 24-sample lag was also applied to the autoregressive components to maintain consistency in capturing temporal dependencies.

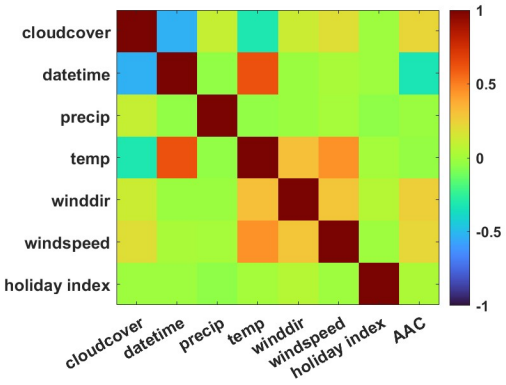


Fig. 3. Correlation Matrix

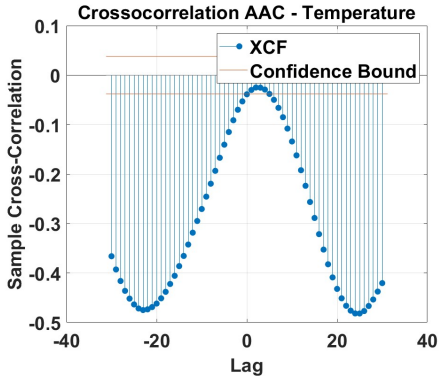


Fig. 4. Correlation curve between the AAC target variable and the exogenous inputs

III. Explainable Predictive Models

Data-driven modeling techniques were employed, with a focus on interpretability and explainability to support energy providers in V2G applications. Linear models offer inherent interpretability, as they are represented by equations where coefficients indicate the relative importance of system variables in the prediction. In contrast, ML-based nonlinear models necessitates post-hoc explanation methods.

A. Models

Modeling approaches for time series prediction depend on the selection of the appropriate model class, whether linear or nonlinear. Models can incorporate only autoregressive components, leveraging past values of the predicted variable (AR for linear and NAR for nonlinear models), only exogenous

inputs, capturing external influencing factors (FIR for linear and NFIR for nonlinear models), or a combination of both autoregressive and exogenous inputs (ARX for linear and NARX for nonlinear models).

1) Linear Models: A linear ARX model set is determined by two polynomials whose degrees are n_a and n_b, respectively:

$$\begin{aligned} A(z^{-1},\theta) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a} \\ B(z^{-1},\theta) &= b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b} \end{aligned} \tag{3}$$

where z^{-1} represents the time delay operator and θ is the set of parameters:

$$\theta := [a_1 \;\; a_2 \;\; \cdots \; a_{n_a} \;\; b_1 \;\; b_2 \;\; \cdots \; b_{n_b}]^T \tag{4}$$

The acronym ARX can be explained in the model equation form for the calculation of $y(t)$, the predicted output at the time instant t:

$$A(z^{-1},\theta)y(t) = B(z^{-1},\theta)u(t) + e(t) \tag{5}$$

where $e(t)$ is a zero-mean white noise process and $u(t)$ is the exogenous input vector [16].

AR refers to the AutoRegressive part $A(z^{-1},\theta)y(t)$ in the model, while X refers to an eXogenous term $B(z^{-1},\theta)u(t)$. The model set is completely determined once the integers n_a, n_b, and the parameter set θ have been specified. The AR model is a special case of Eq. (3) with $n_b = 0$. Alternatively, Finite Impulse Response (FIR) models can be used when output regressors are omitted. FIR is a special case of Eq. (3) with $n_a = 0$.

2) Nonlinear Models: Different ML-based nonlinear models [17] were applied with different input sets.

- Tree Ensemble (TE): Ensemble models combine results from many weak learners into one high-quality ensemble model.The hyperparameters to be optimized are ensemble method (Bag, LSBoost), Number of learners (10-500), Learning rate (0.001-1), Minimum leaf size (1-1080), Number of predictors to sample (1-146)
- Gaussian process regression (GPR): The hyperparameters to be optimized are: Basis function (zero, constant, linear), Kernel function (Rational Quadratic, Squared Exponential, Matern 5/2, Matern 3/2, and Exponential), Sigma.

B. Model Interpretation and Explaination

The explanation of the model is crucial for two main reasons. Firstly, it aids in the design and validation of the predictive model by providing insights into input selection. Secondly, it ensures reliability for the end user, specifically the energy provider, by justifying the model's predictions. Black-box nonlinear models require explainability techniques such as SHAP [1] to interpret their decision-making process. It applies cooperative game theory principles to attribute feature contributions in predictive models. Each model input feature is treated as a game player, and the model function defines the game rules. The computed Shapley values quantify each feature's contribution to the model's prediction by assessing its impact across all possible coalitions. Given a model $f(x)$ and an input instance x, the Shapley value for each feature i is computed as:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{i\}) - f(S) \right] \tag{6}$$

where S represents all possible feature subsets excluding i, and N is the total set of features. SHAP ensures a fair feature attribution by considering all possible feature coalitions. The sum of all Shapley values satisfies:

$$f(x) = \sum_{i=1}^{N} \phi_i \tag{7}$$

which decomposes the model output into additive feature contributions. This property enables SHAP to perform feature importance analysis, thereby enhancing trust in AI-driven decision-making.

IV. Results

Different model identification and prediction were performed for the selected model class, input set and the regressors:

- *FIR/NFIR*
- *ARX/NARX*

Data was divided into training and validation (85%) and testing (15%) datasets. In particular, to avoid seasonal bias and unbalanced datasets, each period of the year for which data is extracted was divided in training/validation and test, and holiday weeks were included both in the training and test phases. A k-fold validation technique with $k = 5$ was applied to avoid overfitting. Data was standardized using z-score normalization. The model optimization for hyperparameter determination was implemented via a Bayesian algorithm on the training/validation data for each input set. The optimization metric was based on the minimization of the root mean square error (RMSE), and global performances were compared using also mean absolute error (MAE) and determination coefficient (R^2).

A. FIR/NFIR Identification

The optimized hyperparameters for the nonlinear ML-based models were determined as follows:

- TE: ensemble method (LSBoost), Number of learners (484), Learning rate (0.17), Minimum leaf size (29), Number of predictors to sample (6)
- Gaussian process regression (GPR): The hyperparameters to be optimized are: Basis function (constant), Kernel function (Exponential), Sigma (Automatic).

B. ARX/NARX Identification

The optimized hyperparameters for the nonlinear ML-based models were determined as follows:

- TE: ensemble method (LSBoost), Number of learners (181), Learning rate (0.12), Minimum leaf size (9), Number of predictors to sample (34)

- Gaussian process regression (GPR): The hyperparameters to be optimized are: Basis function (constant), Kernel function (Exponential), Sigma (Automatic).

C. Model Comparison

The models were applied for a 30-minute (1 sample) ahead prediction and compared based on global key performance indicators (KPIs) and their time prediction response. Table I highlights that ARX/NARX models generally outperform FIR/NFIR models, indicating that exogenous inputs alone do not sufficiently capture the system dynamics without incorporating historical information of the target AAC variable. This suggests that autoregressive components play a fundamental role in capturing temporal dependencies in AAC prediction.

Within the FIR/NFIR models, the Exponential GPR exhibits slightly superior performance in the training/validation phase, whereas the TE achieves the best results in the test dataset. This suggests that exogenous inputs exhibit a nonlinear relationships with the involved variables, making nonlinear models more suitable for capturing their influence. A similar pattern is observed in ARX/NARX models: the Exponential GPR performs better in training/validation, while the TE achieves the best correlation with actual AAC values in the test set. The superior generalization of the TE in the test phase suggests its robustness to unseen data.

To further investigate model performance, the time response analysis is illustrated in Figure 5. Both the Exponential GPR and the TE exhibit strong alignment with the true AAC dynamics, accurately capturing peak variations, particularly around 12:00 AM, which is crucial for V2G ancillary service provision.

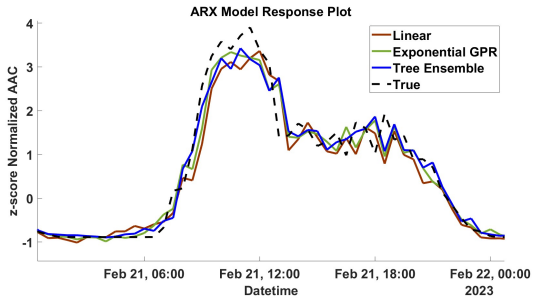
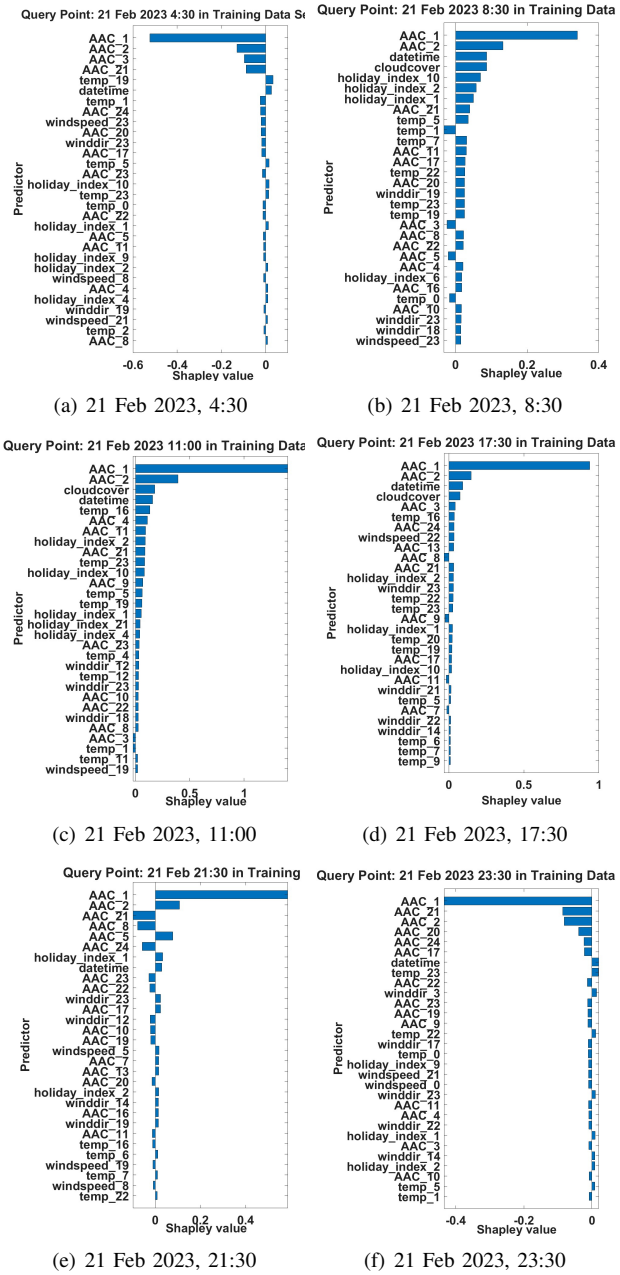


Fig. 5. 30-minute ahead prediction for AAC_{EUR} using different Tree Ensamble NARX for a period validation dataset time interval

D. Model Explanation

A more detailed analysis of model performance and prediction behavior is conducted through local SHAP-based explanations for a selected day (February 15, 2023) from the training and validation dataset at different time intervals. The first query point analyzed is at 4:30 AM, when the AAC is close to zero. The local explanation indicates that this low prediction is primarily driven by autoregressive components with negative Shapley values, while the previous day's temperature and datetime provide a slight positive contribution. This pattern likely reflects overnight parking trends in the EUR commercial district, with early departures influenced by residual temperature effects from the previous day.

The second query point, 8:30 AM, falls within the rising phase of the AAC curve, where drivers' parking decisions may be affected by meteorological and calendar-related factors. The SHAP analysis reveals that datetime, cloud cover, holiday index, and temperature, as well as their past values, positively contribute to AAC predictions, suggesting that these variables play a significant role in morning parking dynamics in the EUR zone.



(a) 21 Feb 2023, 4:30 (b) 21 Feb 2023, 8:30

(c) 21 Feb 2023, 11:00 (d) 21 Feb 2023, 17:30

(e) 21 Feb 2023, 21:30 (f) 21 Feb 2023, 23:30

Fig. 6. Local SHAP-based prediction explanation for the NARX Tree Ensemble model. Explanation for different query points on February 21, belonging to the training/validation dataset: (a) 04:30, (b) 08:30, (c) 11:00, (d) 17:30, (e) 21:30, (f) 23:30.

TABLE I
PERFORMANCE METRICS FOR DIFFERENT MODEL TYPES GROUPED BY MODEL CLASS.

Model Class	Model Type	RMSE (Valid.)	R^2 (Valid.)	MAE (Valid.)	RMSE (Test)	R^2 (Test)	MAE (Test)
FIR/NFIR	**Linear Regression**	0.62	0.41	0.46	0.59	0.20	0.45
	Exponential GPR	0.38	0.77	0.27	0.61	0.14	0.47
	Tree Ensemble	0.32	0.84	0.23	0.58	0.23	0.46
ARX/NARX	**Linear Regression**	0.31	0.85	0.23	**0.29**	0.71	0.23
	Exponential GPR	**0.30**	**0.86**	**0.21**	0.34	0.74	0.26
	Tree Ensemble	0.31	0.85	0.22	0.32	**0.77**	**0.23**

The subsequent query points at 11:00 AM and 5:30 PM occur near AAC peaks and are predominantly influenced by the most recent autoregressive components (lags 1 and 2), combined with meteorological and holiday information. These factors indicate that both short-term historical patterns and external conditions strongly impact peak-hour parking availability.

The later query points at 9:30 PM and 11:30 PM show a stronger dependence on daytime effects and autoregressive components, likely reflecting habitual driver behaviors related to evening commutes and overnight parking. The holiday index has only a marginal influence, suggesting that special calendar events might slightly alter these recurring patterns.

V. CONCLUSIONS

A key challenge in V2G management is balancing the need for accurate predictions with model explainability, enabling energy providers to make informed decisions. To address this challenge, we applied both linear and nonlinear data-driven models, comparing their performance using global KPIs and time-series analysis. In particular ARX/NARX models integrated both autoregressive and exogenous components, with nonlinear models achieving the best overall performance by dynamically combining both inputs. Additionally, we employed explainability methods to have insight on the model outputs for different hours of the day, ensuring transparency in the decision-making process. From a practical perspective, these findings have significant implications for decision support systems in V2G applications. XAI techniques, such as SHAP, and LIME in future work, enhance the transparency of nonlinear models, thereby fostering trust and facilitating their adoption in real-world applications.

Future work will focus on extending the dataset to include additional areas with the same land use designation, as well as diverse zones characterized by different land use types, mobility behaviors, and energy consumption patterns. This will enable either the generalization and the transfer of the model or the development of multiple models to analyze how different input variables influence various V2G infrastructure dynamics enabling more precise decision-making for V2G infrastructure planning. Additionally, future work will address key deployment aspects, including integration with operational energy management systems, scalability to large-scale V2G networks, and real-time performance, to ensure practical feasibility in real-world applications.

VI. ACKNOWLEDGMENTS

The authors thank the Department of Enterprise Engineering "Mario Lucertini" of the University of Rome Tor Vergata for the FCD analysis sharing.

REFERENCES

[1] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," p. 4765–4774, 2017.
[2] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should i trust you?" explaining the predictions of any classifier," p. 1135–1144, 2016.
[3] C. Xu, Z. Liao, C. Li, X. Zhou, and R. Xie, "Review on interpretable machine learning in smart grid," *Energies*, vol. 15, no. 4427, 2022.
[4] D. Perez and L. Solar, "The role of meteorological data in pv system performance forecasting," *Rene. Energy*, vol. 134, p. 620–633, 2019.
[5] A. Holiday and E. Thompson, "Fuzzy logic applications in calendar effects for energy load forecasting," *Energy and AI*, vol. 1, p. 100013, 2020.
[6] A. Barbado and Óscar Corcho, "Interpretable machine learning models for predicting and explaining vehicle fuel consumption anomalies," *Eng. Appl. Artif. Intell.*, vol. 115, p. 105222, 2022.
[7] L. Patanè, F. Sapuppo, M. G. Xibilia, and G. Napoli, "A comprehensive data analysis for aggregate capacity forecasting in vehicle-to-grid applications," in *2024 International Conference on Control, Automation and Diagnosis (ICCAD)*. IEEE, 2024, pp. 1–6.
[8] L. Patanè, F. Sapuppo, G. Napoli, and M. G. Xibilia, "Predictive models for aggregate available capacity prediction in vehicle-to-grid applications," *J. of Sens. and Act. Net.*, vol. 13, no. 5, p. 49, 2024.
[9] L. Patanè, F. Sapuppo, A. Comi, G. Napoli, and M. G. Xibilia, "An explainable model framework for vehicle-to-grid available aggregated capacity prediction," in *MetroXRAINE 2024*. IEEE, 2024.
[10] L. Patanè, F. Sapuppo, G. Rinaldi, A. Comi, G. Napoli, M. Xibilia *et al.*, "Model identification and transferability analysis for vehicle-to-grid aggregate available capacity prediction based on origin–destination mobility data," *Energies*, vol. 17, no. 24, 2024.
[11] A. Barredo Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, p. 82–115, 2020.
[12] M. Sundararajan and A. Najmi, "The many shapley values for model explanation," *Proceedings of the 38th International Conference on Machine Learning*, vol. 139, p. 9269–9278, 2021.
[13] Regulation (EEC) No 4064/89 Merger Procedure. [Online]. Available: https://ec.europa.eu/competition/mergers/cases/decisions/m1572_en.pdf
[14] 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.
[15] "Visual crossing weather data service," https://www.visualcrossing.com/weather/weather-data-services.
[16] L. Ljung, *System Identification: Theory for the User*. USA: Prentice-Hall, Inc., 1986.
[17] A. Lindholm, N. Wahlström, F. Lindsten, and T. B. Schön, *Machine Learning: A First Course for Engineers and Scientists*. Cambridge University Press, 2022.