

# Electric Vehicle Supply Equipment Day-Ahead Power Forecast Based on Deep Learning and the Attention Mechanism

Silvana Matrone<sup>id</sup>, *Student Member, IEEE*, Emanuele Ogliari<sup>id</sup>, *Member, IEEE*,  
Alfredo Nespoli<sup>id</sup>, *Student Member, IEEE*, and Sonia Leva<sup>id</sup>, *Senior Member, IEEE*

**Abstract**—Transports is one of the sectors that produce the highest emissions of CO<sub>2</sub>; in the last ten years, there has been a process of decarbonization which has led to a considerable increase in Electric Vehicles (EVs). However, the sudden introduction of a large number of Electric vehicle supply equipment (EVSE) supplying electrical energy to EVs could cause problems in the management of the electric grid which must cope with the consequent increase in the electrical load demand. In this context, the 24 hour ahead forecast of the power curve associated with the recharge of EVs becomes of vital importance to ensure the reliability of the electric grid. In this paper, different Machine Learning models based on Recurrent Neural Networks (LSTM, GRU) and with different architectures, are compared based on their capability to accurately predict the power curve of an EV charging station one day in advance. A Sequence to Sequence model has been implemented and a thorough analysis of an Attention layer has been detailed. The models are tested on a real world open dataset.

**Index Terms**—Electric vehicles, day-ahead forecast, deep learning.

## NOMENCLATURE

<i>EV</i>	Electric Vehicles.
<i>EVSE</i>	Electric vehicle Supply Equipment.
<i>RES</i>	Renewable Energy Sources.
<i>STLF</i>	Short-Term Load Forecast.
<i>MTLF</i>	Medium-Term Load Forecast.
<i>LTLF</i>	Long-Term Load Forecast.
<i>SSLTF</i>	Super-Short-Term-Load-Forecasting.
<i>ANN</i>	Artificial Neural Networks.
<i>LSTM</i>	Long-Short-Term Memory.
<i>GRU</i>	Gated Recurrent Unit.
<i>S2S</i>	SequencetoSequence.
<i>RNN</i>	Recurrent Neural Networks.
<i>ED</i>	Encoder-Decoder.
<i>EMD</i>	Empirical Mode Decomposition.
<i>MIMO</i>	Multi-Input-Multi-Output.
<i>EMS</i>	Energy Management System.

Manuscript received 26 November 2022; revised 13 April 2023, 16 June 2023, 7 July 2023, and 30 August 2023; accepted 13 April 2024. This work was supported by the MOST—Sustainable Mobility Center through the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)-MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4-D.D. 1033 17/06/2022) under Grant CN00000023. The Associate Editor for this article was N. Papanikolopoulos. (*Corresponding author: Emanuele Ogliari.*)

The authors are with the Department of Energy, Politecnico di Milano, 20156 Milan, Italy (e-mail: emanuelegiovanni.ogliari@polimi.it).  
Digital Object Identifier 10.1109/TITS.2024.3391375

<i>JPL</i>	Jet Propulsion Laboratory.
<i>MAE</i>	Mean Absolute Error.
<i>MSE</i>	Mean Squared Error.
<i>RMSE</i>	Root Mean Squared Error.
<i>SMAPE</i>	Symmetric Mean Absolute Percentage Error.

## I. INTRODUCTION

**E**LECTRIC load forecasting has gained more and more attention during the last ten years along with the increase in global electric car stock that reached 26 million units in 2022 [1]. The main reason behind this phenomenon is found in the growing awareness about climate change that led consumers to adopt more environmentally friendly choices. In 2022, the transport sector alone produced 7.98 Gt of CO<sub>2</sub> [2], the only ways to reduce its impact are decarbonization and electrification of transports, however, this type of evolution still has its own downsides when it comes to the huge increase of EV fleet that will cause a higher need of charging stations, numerous studies have been conducted to model the penetration of electric vehicles and their impact on the national grid [3]. Electric vehicle supply equipment (EVSE) supplies electricity to an electric vehicle (EV) [4]. It is usually called a “charging station” or “charging dock” and it provides electric power to the vehicle for recharging the EV batteries. The new power load demand will cause a significant rise in the peak load and a decrease in reserve margin [5] representing a threat to the security of the national power grids. In this context the forecast of power load demand from EVSEs becomes the turning point to ensure the perfect management of the national grid balance.

EV load forecast is a form of Demand Side Management (DSM) [6], [7], [8] that can be crucial to balance load demand and power production. It is particularly relevant when the vehicles are integrated into Distributed Generation (DG) such as a micro-grid [9]. The management of energy dispatch is the main issue when dealing with DG, the optimization of energy flows is not trivial and is largely addressed in the literature using several different techniques [10], [11]. Renewable Energy Sources (RES) play a key role in DG, a coupling between EV load forecast and RES generation forecast is strategic in this sense, a common example is the coupling with photovoltaic power forecast [12]. In the framework of a micro-grid, an EV could also play a double

role, behaving alternately as a load and as a source of energy (Vehicles-to-grid) or even acting as a decentralized source for energy trading, becoming Connected Electric Vehicles [13], [14].

Based on the time-horizon of the prediction, short-term, medium-term, and long-term forecasting can be defined. Although there is a lack of coherence in this classification in literature, Short-Term Load Forecast (STLF) usually refers to predictions of the load demand between one hour and one week ahead (crucial in the optimization of the grid operation), Medium-Term Load Forecast (MTLF) is used to forecast load from one week to one year in the future and is mainly important in the planning of maintenance operations; lastly, Long-Term Load Forecast (LTLF) has to do with predictions that exceed one year and are mainly involved in the decision process for big investments in new infrastructures and generation units [15]. Another category that has gained importance in the latest years is the Super-Short-Term-Load-Forecasting (SSLTF) [16], also defined as Ultra-Short-Term-Load-Forecasting [17] or Very-Short-Term [18], corresponding to minute predictions or, in general, under one hour.

Electric load forecasting methods can be categorized into traditional statistical models and artificial intelligence models [17]. Traditional models include time series analysis methods, Autoregressive Integrated Moving Average (ARIMA), regression analysis, Kalman filtering, and statistical methods, while artificial intelligence methods include Artificial Neural Networks (ANN), Support Vector Machine, and Deep Learning (DL) models. At the very early stage in EV load forecasting studies, statistical models were the most suitable choice as the lack of real, organized data about EV charging made it necessary to build realistic scenarios data through computational algorithms. With the availability of actual EVSE load data, the approach to EV load forecasting shifted totally from a probabilistic to a data-driven one. Nowadays the main direction of the research in EV power load forecasting is concentrated on ANNs in all their declinations [19], [20].

In this paper S2S (SequencetoSequence) based on Long-Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models are proposed, an attention mechanism inspired by Bahdanau [21] with dot alignment score is added. Our contributions in this paper are:

- 1) a review of the state-of-the-art models for electric and EV load forecast.
- 2) The application of the proposed models to a real world EV charging dataset with the aim to test and compare the performances of EV charging load forecast with 24 hours in advance.
- 3) A detailed analysis on the forecasting accuracy of more complex models compared to simpler RNNs architectures with a focus on the "Attention" mechanism.
- 4) A thorough appraisal of hyperparameters tuning and their influence on the forecasting performances.

The research is based on one of the few available public datasets found in the literature. The remainder of this paper is as follows. Section II describes the related works. Section III

outlines the methodology adopted in the study with a detailed presentation of attention mechanisms, Section IV presents the case study analyzed in the paper, it describes how the dataset is composed, how the data has been treated and reshaped to enter the models introduced and how the results will be evaluated. Section V displays and discusses the results of both study cases while Section VI draws the conclusions of the study.

## II. RELATED WORKS

Although it is a common opinion that Recurrent Neural Networks (RNNs) have outperformed all other techniques, there is not a model that surpasses the others in every task, this leads to a large literature of comparative studies among different ML models, as listed in Table I with increasing time horizon. In the framework of S2S forecasting, Machine Learning and in particular LSTM seems to be highly effective in forecasting the load on minute-level. In [16], the LSTM model results are superior to all other models even on a second, more random, dataset that is less representative of a typical usage. In [17] a novel LSTM-based model is the best-performing model on two different time scales. Authors in [22] present a model based on the Machine Theory of Mind composed of three networks, the first two are built based on LSTM networks. The authors of [23] propose a self-attention-based machine theory of mind assessing its superiority on other 5 state-of-the-art models using a quantile forecast evaluation metric as the loss function.

Moving to slightly longer horizons (24 hours) in the realm of STLF, a larger variety of models and techniques has been analyzed in the literature, RNNs still have high relevance being at the base of more complex architectures that are developed such as Encoder-Decoder(ED) models. Reference [24] presents a performance comparison of four DL-based methods. After tuning the number of hidden layers the GRU model with only one layer results to have the best performance. In [25] the analyzed and compared models are tested on three different synthetically generated time series plus three real-life datasets. The authors in [26] propose an LSTM model coupled with feature engineering Empirical Mode Decomposition (EMD).

In [29], an S2S model is proposed, additionally, two different types of attention (Bahdanau [21] and Luong [43]), are added to the S2S. They concluded that, as the prediction length increases, the accuracy decreases, additional layers do not improve the forecasting accuracy, and that overall S2S models with attention performed better than non-S2S models for all input lengths. A two-stage STLF model based on LSTM and Multi-Level Perceptron (MLP) is presented in [35]. In [31] a model based on Wavenet is described and evaluated through a comparison with state-of-the-art ANN. Bedi and Toshniwal [33] propose a framework (D-FED) to forecast electricity demand, it is based on LSTM with a novel moving window-based Multi-Input-Multi-Output (MIMO) mapping approach technique.

Examples of Reinforcement learning are proposed by [28] and [37]. RL has been applied to EV in literature also to integrate driving behavior in the Energy Management System (EMS) as in [44] and to manage the forecasted charging load in vehicle-based mobility-on-demand systems [45] or

TABLE I  
BIBLIOGRAPHIC REVIEW ON ELECTRIC LOAD FORECAST SORTED BY THE INCREASING TIME HORIZON

Ref	Horizon	Step Ahead	Methods	Metrics Best Value	Dataset
[16]	SSTLF	1 min	ANN, RNN, LSTM, GRU, SAEs, Bi-LSTM	MAE: 0.286 (kW) RMSE: 0.442 (kW)	PEV charging station in Shenzhen(N. A.)
[22]	SSTLF	5 min	MToM-Based Quantile Forecast Network string	personalized loss function	ACN-data [27]
[17]	SSTLF	15/30 min	LSTM	MAE: 12.302 (kWh) RMSE: 5.509 (kWh)	PEV charging station in Shenzhen (N. A.)
[23]	SSTLF	15/30	SAMToM	CRPS: 2.09%, reliability: 1.28%, sharpness: 1.79%	ACN-data [27]
[24]	STLF	1 h	DNN, RNN, LSTM, GRU	NMAE: 0.77% NMRSE: 2.89%	PEV charging station in Shenzhen (N A.)
[28]	STLF	1 h	Q-Learning	MSE: 0.89 (kW)	Simulated Data
[29]	STLF	1 h/ 4 h/ 10 h/ 24 h	S2S with Bahdanau Attention	MAE: 2.705 (kWh), MAPE: 12.155%	Electric Load of a commercial building (N. A.)
[25]	STLF	24 h	ERNN, LSTM, GRU, NARX, ESN	NRMSE: 0.38197/ 0.37678/ 0.87451 (kWh)	Orange dataset (telephonic activity load), ACEA dataset (electricity load), GEFCOM2012 dataset (electricity load) [30]
[31]	STLF	24 h	Wavenet	MAE: 1117.471(MW), MSE: 2.685 (MW), RMSE: 1638.709(MW)	Réseau de Transport d'Électricité (RTE) electric load of the French grid [32]
[33]	STLF	24 h	D-FED	NRMSE: 7.10%	UT Chandigarh energy demand [34]
[35]	STLF	24 h	S2S with Attention + Multi Layer Preceptron	MAPE: 1.96% RMSE: 0.94 (GW)	Tokyo Electric Power Company (TEPCO) data [36]
[37]	STLF	24 h	Reinforcement Learning	MAPE: 1.368 (kW), MAE: 196.314 (kW), RMSE: 0.988 (kW)	ISO-NE dataset, GEFCOM2017 dataset [38]
[39]	MTLF	1 month	S2S	MAE: 9.2 (kWh), RMSE: 12.87 (kWh)	PlugShare [40]
[41]	MTLF	1 month	TDResNet-SD-ISE	MAPE: 1.485%, RMSE: 409.321 (kW)	Independent System Operator-New England (ISO-NE) American Utility dataset [42]

coordinated charging management of EVs based on local information [46]. Prophet-BiLSTM is employed in the day-ahead forecast of EV charging load in [47], it performed better than transformer and DNN. Reference [48] is a review on EV scheduling, clustering, and forecasting. Another review on both load and occupancy day-ahead forecasts is presented in [49], which direct and bottom-up approaches with both statistical and machine learning models. It results that Machine Learning models are mostly preferred over probabilistic models. For example, authors [50] present a novel spatiotemporal Deep Learning network for traffic flow prediction problems, which is capable of modeling the periodicity of traffic flow data with a well-designed time embedding strategy from 15 minutes to 1 hour. Regarding the MTLF forecast, reaching a monthly horizon, more complex model structures are proposed to face the reduced accuracy that results from the increase in the time ahead interval. A medium-term demand forecasting S2S model is presented in [39], while [41] proposes a trilinear deep residual network (TDResNet) structure. Finally, in [51], authors propose an optimization procedure exploiting EV forecast to schedule the recharging processes under dynamic power prices such as real-time pricing, time of use, critical peak pricing, and prime times refunds.

In this study, we conduct a comparison of forecasting models employing RNNs for STLF with a prediction horizon of 24 hours. This approach aligns with similar investigations outlined in the literature, as indicated in Table I. The method presented here can be similarly applied to charging stations, mirroring how Machine Learning has been employed

to enhance sensors' intelligence [52]. In our work, the data utilized for training and testing the models consist of publicly available records from EV charging sessions [27]. To enhance the accuracy of the RNN models, we introduce an ED structure and incorporate an attention mechanism. The performance of these enhanced models is then directly compared with the baseline RNN model, considering persistence as a benchmark. Additionally, we conduct a sensitivity analysis on the number of layers and the quantity of RNN neurons in each layer to evaluate their impact on the final forecasting results.

### III. ADOPTED METHODOLOGY

#### A. Encoder-Decoder/Sequence to Sequence

The power load forecast of the charging station is a Many-to-Many prediction as it usually takes in input the time series of power sampled at a certain time-step and tries to predict another sequence of values, it could require the two sequences to have different lengths. A proper way to handle a Many-to-Many prediction is an S2S model. S2S models are based on an ED architecture that consists of 3 components: an encoder, a context vector, and a decoder. Both the encoder and the decoder are constructed as a stack of several RNN units. The encoder takes a sequence as the input and transforms it into a context vector, an element with a fixed shape that acts as a condensed representation of the input time series. The output of the encoder, the hidden state, is usually the state of the last RNN time step. The length of the vector depends on the number of RNN cells in the encoder. The context vector acts

as the initial hidden state of the decoder part of the model as it takes the output from the encoder and feeds it repeatedly as input at each time step to the decoder. The decoder interprets the context vector to make predictions for each time step required in the output.

### B. Attention Mechanisms

Attention mechanisms were introduced for the first time in 1997 by Bahdanau in [21], and a similar, modified version was later presented by Luong [43]. They were developed to solve the main issue that arises when dealing with S2S models: the incapability of the network to remember long sequences given as input due to the limited length of the context vector. In an ED model, at each time step of the decoder, important elements for the creation of the current output could come from any point of the input sequence, but the context vector is not capable of encapsulating the data “seen” at the beginning of the input. Attention mechanisms try to replicate the human brain’s capability of focusing on small portions of the entire information provided at a time, paying attention only to the ones that are relevant to the answer that is required. The encoder functions exactly as in the case of a simple Encoder-Decoder model producing a hidden state  $h$  for each time-step. The context vector is not constant, but changes at each time step of the decoder, calculated as a weighted mean among all the hidden states of the encoder as shown in (1).

$$c_t = \alpha_{t,1}h_1 + \alpha_{t,2}h_2 + \dots + \alpha_{t,m}h_m \quad (1)$$

where  $c_t$  is the context vector at the current time-step  $t$  and  $h_m$  is the hidden state of the last time-step in the input with  $m$  being the length of the input sequence. The weight  $\alpha_{t,i}$  is the *alignment weight* computed at the  $i$ -th time-step. Each  $\alpha$  is the result of the application of a *softmax function* to the attention scores  $e_{t,i}$  to normalize them. The attention scores are computed through a proper alignment model  $a(\cdot)$ , an operation that scores how well the current hidden state of the encoder  $h_i$  matches with the hidden state of the decoder state  $s_{t-1}$  at the previous time-step. Equation (2) and (3) illustrate how alignment scores and alignment weights are calculated. The weights favor the hidden states that are matching with the previous output, in this way the context vector is *paying attention* only to the parts of the input that are considered relevant by the alignment function solving the problem of the limited length.

$$e_{t,i} = a(h_i, s_{t-1}) \quad (2)$$

$$\alpha_{t,i} = \text{softmax}(e_{t,i}) \quad (3)$$

The decoder takes in input the context vector, the previous hidden decoder state and the current output, to compute the final prediction. The attention model described is the one proposed by Bahdanau [21], in this case the alignment function is a single-layer perceptron as shown in (4):

$$a(s_{t-1}, h_i) = v_a^T \tanh(W_a s_{t-1} + U_a h_i) \quad (4)$$

where  $W_a \in R^{n \times n}$ ,  $U_a \in R^{n \times 2n}$  and  $v_a \in R^n$  are weight matrices that are being optimized during the training process.

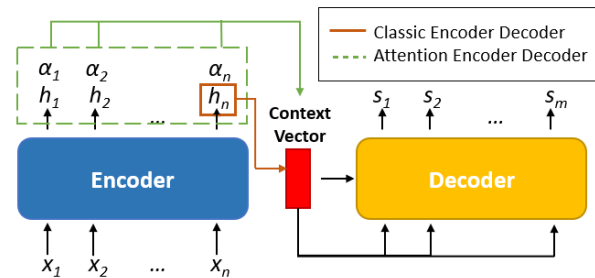


Fig. 1. Context vector comparison between Classic ED and Attention ED.

## IV. CASE STUDY

Despite the growing popularity of EVs, the main issue when dealing with EV charging load forecasting is the scarcity of data. Only a limited number of sources that provide complete data on EV load demand may be found, and most of them are usually not publicly available. The study case proposed in this work is based on one EVSE recording EV charging sessions: the JPL dataset from Adaptive Charging Network (ACN).

### A. JPL Dataset Description

The data collected by ACN have been made available to provide researchers working in the field of electric vehicle charging with a complete set of real data coming from existing facilities. For the purpose of this study, the records from the JPL (Jet Propulsion Laboratory) database will be used. This site currently has 52 EVSEs, or charging stations, and it is only open to employees making it a good representation of a typical workplace schedule. Only a subset of all the available data has been considered suitable due to the anomalies associated with the pandemic, it includes the sessions from 31/12/2018 to 04/01/2020. Some of the values provided for each of them are relevant to the study: session ID, station ID, connection time, disconnection time, done charging time, and time zone. The session ID is a univocal code identifying the single charging event, in the same way, the station ID is a unique code associated with a charging spot. The connection time expresses the date and time at which the vehicle has connected to the station, similarly, the disconnection time is the instant in which the EV has disconnected from the charger, while the done charging time is the date and time at which the charging procedure was actually stopped. The time zone is used to calculate the UTC time. The stations in the JPL parking lot are all near in space and used mainly by the university employees so the power curve of the whole parking lot is analyzed as aggregated power. The multi-step forecast of the power load will be performed using Machine Learning models based on different Recurrent Neural Networks (RNN) and architectures, the time horizon is 24 hours with 15 minutes time-steps for a total of 96 steps, the data of the previous day (96 steps back) is used as input to make predictions on the next day.

### B. Pre-Processing on JPL Data

The data has been organized as a time series, creating a series of 15-minute time steps, each of them associated with the correspondent sum of the average power required by each



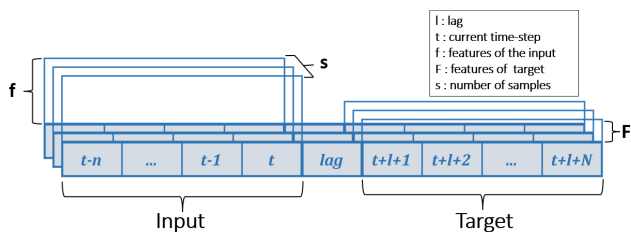


Fig. 2. Input and Target represented as 3D matrices after the reshaping procedure.

vehicle that was in charging mode during that time interval. The time series has been re-framed to enter an ML problem with input patterns (X) and output patterns (Y). The current time is addressed as ( $t$ ), future time steps ( $t + 1, \dots, t + N$ ) are the ones to be forecasted, and past observations ( $t - 1, \dots, t - n$ ) are used to make forecasts. At each time step X and Y are computed shifting the original dataset  $n$  steps back and  $N$  steps forward. After the reshaping, each sample X in the input is a 3D matrix with dimensions ( $samples, n, f$ ), the corresponding output Y is a similar 3D matrix with dimensions ( $samples, N, F$ ) where  $f$  and  $F$  are respectively the features of the input samples and the ones to be forecasted. The lag  $l$  is the number of time-steps between the current time  $t$  and the first forecasted step  $t + l$ . The training samples generation is schematized in Figure 2. After the reshaping the input data is scaled between 0 and 1 using a MinMax scaler to achieve higher stability and shorter training periods. Then the data is split into a training set and a test set, having to do with time-series problems the sets have been split sequentially.

### C. Model Tuning and Evaluation on JPL Dataset

The models used to make predictions are all based on RNNs, an S2S model with LSTM/GRU layers has been implemented, and it is compared to simple LSTM and simple GRU. Additionally, an attention layer inspired by the one described by Bahdanau [21] is introduced in the ED model to check whether or not it can improve its performance. Proper tuning is performed on the models to identify the best architecture for each one. The hyperparameters undergoing the tuning are the number of units in the RNN layers and the number of RNN layers. In the case of S2S structures, both the number of layers in the encoder and the number of layers in the decoder are changing. Three values are considered for the number of units in each layer: 32, 64 and 128; up to three layers are considered in simple models and two layers in the ED ones for a total of 36 different models. The main characteristics of each model are summarized in Table II.

The first architectures to be analyzed in the simulations were the S2S models with LSTM (ED LSTM) and GRU (ED GRU). The built model has an encoder, a context vector, and a decoder. The encoder consists of one or more layers of LSTM, or GRU, based on the different cases. After the analysis of the ED models their corresponding simple models were simulated to understand if, and eventually in what amount the S2S model could improve the performances.

Finally, an attention mechanism is introduced in the S2S models (LSTM Attention) with the aim of assessing the performances with respect to the previously analyzed models.

It has been chosen to build the attention model with LSTM since it has shown better performances and higher stability in previous cases. The employed attention layer is inspired by the one described by Bahdanau [21]. The dot alignment score is used, it is calculated as in (5) where  $h_i^T$  is transposed of the  $i$ -th hidden state in the encoder and  $s_{t-1}$  is the previous hidden state in the decoder.

$$e_{t,i} = \text{dot}(h_i^T, s_{t-1}) \quad (5)$$

The benchmark used to evaluate the models is the simplest forecast baseline model, Naive Persistence (Persistence), it is based on the most trivial assumption that can be made when forecasting, it takes the value at the current time-step ( $t$ ) and assigns it to the predicted value for the next step ( $t + 1$ ), in our case the value of the time-step ( $t$ ) is assigned to the time-step that has to be predicted 24 hours ahead ( $t + 96$ ) as described in (6), where  $y_p$  is the predicted value and  $y$  the actual one.

$$y_p(t + 1) = y(t) \quad (6)$$

### D. Performance Metrics

The performances of all the models are assessed using four different evaluation metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Symmetric Mean Absolute Percentage Error (SMAPE). They are among the most employed performance metrics for time-series forecasting evaluation in the literature.

MAE, MSE, and RMSE are scale-dependent errors, which means that the resulting error is on the same scale as the data, so their values are not capable of expressing the performances of the model in an absolute way but only relative to other models in the same problem, while the adoption of Symmetric Mean Absolute Percentage Error tries to solve this problem giving a share instead of an absolute value leading to a scale-independent error.

Besides, it should be underlined that the training loss function is calculated on the MSE which is normalized between the minimum and maximum values of the time series in order to speed up the training process, therefore the training and test error represented in Figure 6 and 4 is a normalized MSE (nMSE). The employed optimizer in the machine learning models is Adam. Finally, the proposed algorithms have been tested on a workstation at Politecnico di Milano, equipped with an Intel® Core™ i9-10900KF CPU with 10 cores of base frequency of 3.7 GHz.

## V. RESULTS

In Table II the performance comparison of the most accurate models among those analyzed is reported. They are listed according to their: architecture (naive, simple, S2S or mixed with the attention mechanism), model (Encoder-Decoder or not, LSTM, GRU, and Persistence), and number of layers and units per layer. In addition, the elapsing time in the training is reported. The best-performing model includes the attention mechanism (LSTM Attention) and consists of an S2S architecture with a double-layer LSTM both in the encoder and the decoder made by 128 units in each layer. This model scored the best MAE, MSE, and, RMSE slightly higher than

TABLE II  
PERFORMANCE COMPARISON OF THE BEST CASES FOR EACH FORECASTING MODEL

Architecture	Model	Layers	Units	MAE (kW)	MSE (kW)	RMSE (kW)	SMAPE (%)	Training Duration (min)
S2S	ED LSTM	2 – 2	64	11.61	479.63	21.90	49.91	60
S2S	ED GRU	1 – 1	64	12.09	400.59	20.01	50.94	24
Simple	LSTM	1	128	11.35	359.65	18.96	41.56	30
Simple	GRU	1	128	11.69	377.97	19.44	41.68	16
S2S+Attention	LSTM Attention	2 – 2	128	<b>10.33</b>	<b>388.10</b>	<b>19.70</b>	<b>42.27</b>	91
Naive	Persistence	None	None	18.85	1259.29	35.49	84.78	None

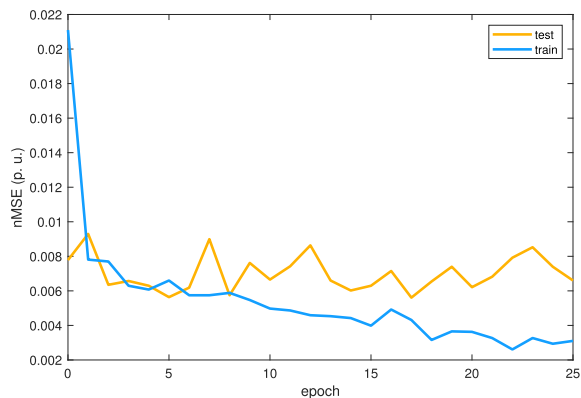


Fig. 3. Trend of the nMSE during the train (blue) and test (orange) of the best performing model (LSTM Attention).

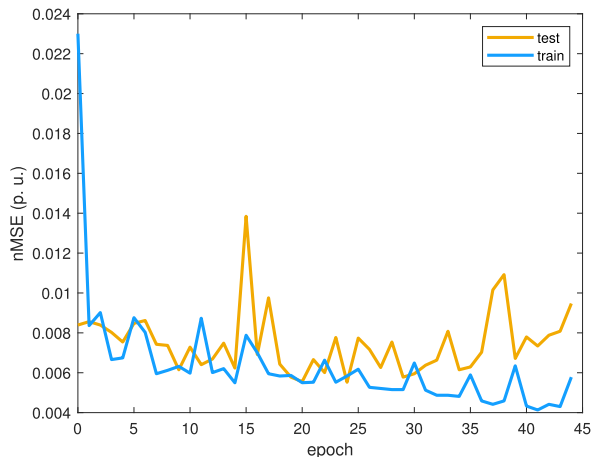


Fig. 4. Trend of the nMSE during the train (blue) and test (orange) of the best performing model (ED LSTM).

the simple GRU and LSTM. However, these simple models suffer from poor stability in terms of error trends in the training process, compared to the relevant models with attention mechanisms. This can be clearly seen in Figure 4 where the training error curve (in blue) of the LSTM with attention shows an almost monotonic descending trend which is smoother than the corresponding S2S LSTM training curve in Figure 3. The LSTM Attention model shows one of the lowest SMAPE in Table II and when counting the time duration it is not fast in the single training epoch when compared to the others. This result is reasonable if considering the higher amount of trainable parameters needed for this model. However, it performed 26 epochs before the training was stopped, which is less than the relevant model without the Attention mechanism. After the tuning of ED LSTM and simple LSTM,

it could be noticed that the simple model had outperformed the correspondent ED one leading to the conclusion that a more complex model was not advantageous. The same observation could be done for the ED GRU and simple GRU. However the introduction of the attention layer, in particular in the LSTM Attention model, completely overturned the results improving the performances of the ED LSTM, with a decrease in MAE of 8%, MSE was lowered by 19%, RMSE by 10%, and SMAPE by 10%.

Some general considerations can be drawn from the analysis of the tuning procedure which is shown in Figure 5. Regarding simple models, their performance metrics MAE, RMSE, and SMAPE tend to lower at first when increasing the complexity of the model and then rise when reaching 128 units, both LSTM and GRU reach the best performances with the model (1 layer and 128 units) as can be observed in Figure 5.

In ED structures, increasing the number of units and layers has different effects on LSTM and GRU, but both of them are performing worse when reaching 128 similar to the simple structures. The higher complexity of the model, when increasing the number of layers in the encoder and the decoder, increases the performances with 32 and 64 units but not in the case of the biggest model with 128 units. In general, the increasing complexity due to the higher number of units is not associated with an increase in the accuracy of the forecast while, on the other hand, it brings to longer training duration. The model with attention is the only one that benefits from an high number of trainable parameters. About the direct comparison between LSTM and GRU, it can be underlined that the second is performing worse on average and this is less evident in the ED architectures but still it consistently has lower training stability making LSTM more advisable.

Figure 6 shows the trend of the actual power and the trend of the forecasts from the LSTM Attention model (in red) and the ED LSTM (in green) during the week spanning from 28/10/2019 to 3/11/2019.

During this week, both models generally underestimated the value of the actual time series giving an under-prediction. However, the LSTM Attention model learned more accurately the daily pattern and is capable of forecasting the starting and the ending of the daily load. Besides, it also learned the weekly pattern being capable of identifying the start and the end of the working days. During the weekend the performances are lower than expected because user behavior is significantly less periodical and then more unpredictable. This worse performance trend is more highlighted during peculiar days such as national holidays.

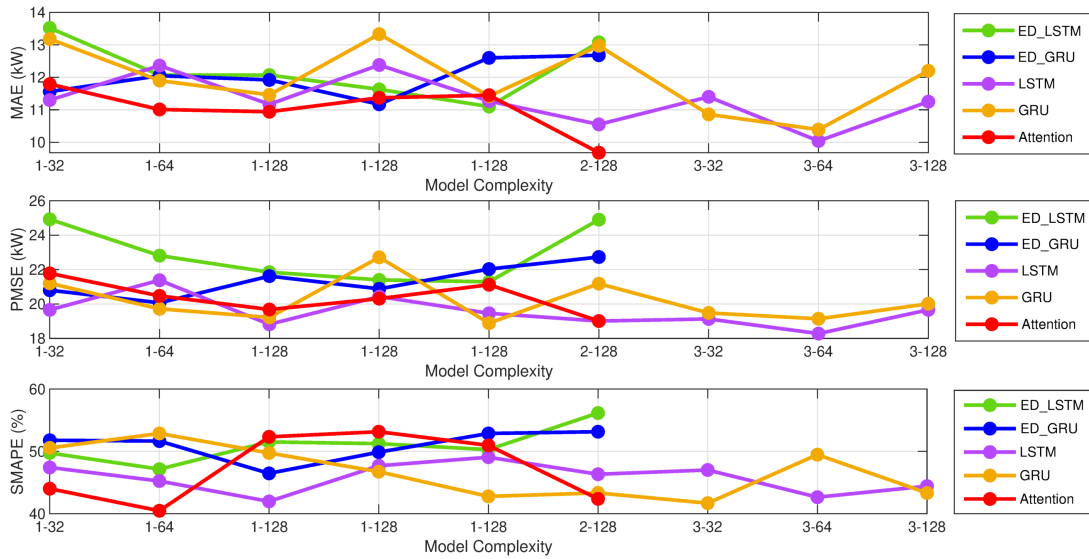


Fig. 5. Trend of the error metrics used to evaluate the models MAE, RMSE and SMAPE as function of the complexity of the models during the tuning procedure, the number of layers and the number of units within the layers is increasing along x axis.

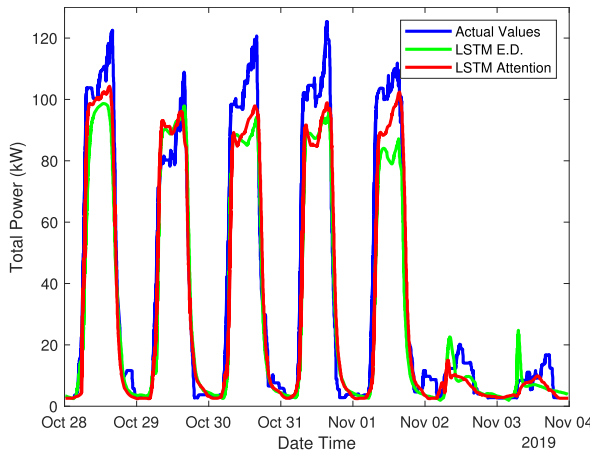


Fig. 6. Trend comparison between the actual total power of all EVSEs in JPL dataset and the total power forecasted by the best model (LSTM Attention) and the direct competitor model (ED LSTM), during a typical week.

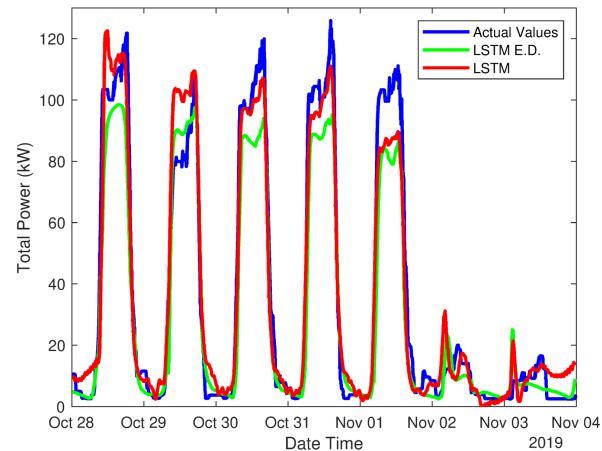


Fig. 8. Comparison between the trend of the actual total power of all EVSEs in JPL dataset and the total power forecasted by simple model LSTM and the ED LSTM on a typical week.

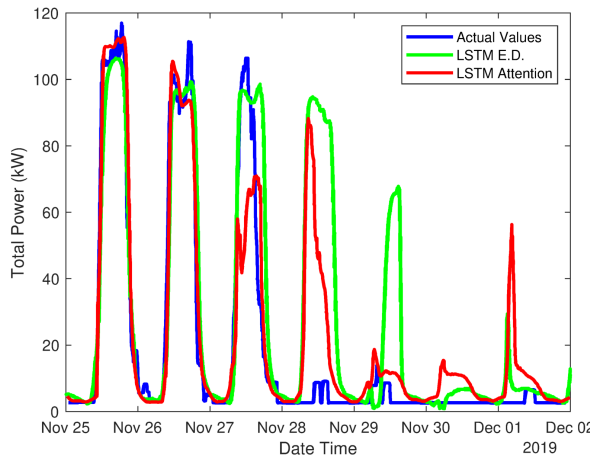


Fig. 7. Trend comparison between the actual total power of all EVSEs in JPL dataset and the total power forecasted by the best model (LSTM Attention) and the direct competitor model (ED LSTM), on the Thanksgiving weekend.

Particularly low performances can be observed during the week from 25/11/2019 to 1/12/2019 (Figure 7), they can be easily addressed to the celebration of Thanksgiving day of

that year, the same is true for the Christmas holidays. This behavior is reasonably due to the lack of holiday “examples” included in the training set, which have a lower statistical significance. Finally, Figure 8 shows the comparison between the best model with just LSTM trend and the best ED LSTM one, it can be noticed how the performances of the two models are comparable, the ED is slightly worse than simple LSTM, on the other side it is not able of properly predicting values close to zero and during the weekend, however, its training stability is a lot lower than the ED model one (ED LSTM).

## VI. CONCLUSION

This paper presents and applies the LSTM model with the attention mechanism feature to EV charging load forecast 24 hours in advance. The model is here applied to a real-world Electric vehicle charging equipment public dataset, to test and compare its performances with other state-of-the-art Deep Learning models.

From the comparison with the benchmark of Naive Persistence, the best-performing model in terms of MAE, MSE, and RMSE, is the proposed LSTM with attention mechanism model. In addition, the obtained SMAPE is comparable to the simple LSTM and GRU models, but it showed better stability in terms of training error trend.

More in detail, the increased effectiveness of higher complexity models compared to simpler RNN architectures and the key role of the attention mechanism in improving S2S results are here inspected. Finally, the tuning of hyperparameters and their influence on forecasting performances is critically analyzed.

The obtained results highlight how the presence of the attention mechanism represents a turning point in EV charging load forecast models justifying the increased complexity of these models.

A notable limitation of this study is its dependence on the availability of limited data, thereby constraining the extent to which validation can be achieved. The lack of accessible open and expansive datasets impedes the progress of scientific research in this domain. However, further studies are directed towards enhancing the precision of forecasting models through the incorporation of supplementary inputs.

In future works, other inputs could be fed to the model to complete the available information. In particular, exogenous variables that could affect the periodical influx of electric vehicles and the normal operation of relevant supply equipment are of particular interest such as weather forecast, ambient temperature, and rainfall. Furthermore, inspecting the influence of peculiar days affecting the EV users' behavior, such as the occurrence of holidays, mass events, maintenance or roadworks, would be extremely useful to better model the EV charging station occupancy trend and the overall associated power supplied. These events are occurring occasionally and, even if they are generally planned for a long time by the municipalities, they are very difficult to detect and predict without prior knowledge and due to the Machine Learning model being data-driven and experience-based.

In conclusion, the presence of additional data would allow us to build a more complex model, with several trainable parameters of different types (numeric, boolean, etc...) improving the final electric vehicle supply equipment day-ahead power forecast accuracy.

#### ACKNOWLEDGMENT

This manuscript reflects only the authors' views and opinions, neither European Union nor European Commission can be considered responsible for them.

#### REFERENCES

- [1] (2021). *Trends and Developments in Electric Vehicle Markets—Global EV Outlook 2021—Analysis*. [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2021/trends-and-developments-in-electric-vehicle-markets>
- [2] IEA. *Emissions by Sector—Greenhouse Gas Emissions From Energy: Overview—Analysis*. Accessed: Nov. 2022. [Online]. Available: <https://www.iea.org/reports/greenhouse-gas-emissions-from-energy-overview/emissions-by-sector>
- [3] S. Yadav, P. K. Dubey, M. Sudman, R. Srinivas, R. Srisainath, and V. Devi, "Development of an GA-RBF based model for penetration of electric vehicles and its projections," in *Proc. Int. Conf. Self Sustain. Artif. Intell. Syst. (ICSSAS)*, Oct. 2023, pp. 1–6.
- [4] P. K. Dubey, B. Singh, D. K. Patel, and D. Singh, "Distributed generation current scenario in the world," pp. 1–25, 2023. Accessed: Apr. 30, 2024. [Online]. Available: <https://www.ijfmr.com>
- [5] S. Deb, K. Kalita, and P. Mahanta, "Review of impact of electric vehicle charging station on the power grid," in *Proc. Int. Conf. Technol. Advancements Power Energy (TAP Energy)*, Dec. 2017, pp. 1–6.
- [6] D. Said, "A survey on information communication technologies in modern demand-side management for smart grids: Challenges, solutions, and opportunities," *IEEE Eng. Manag. Rev.*, vol. 51, no. 1, pp. 76–107, 1st Quart., 2023.
- [7] A. Nespoli, E. Ogliari, and S. Leva, "User behavior clustering based method for EV charging forecast," *IEEE Access*, vol. 11, pp. 6273–6283, 2023.
- [8] S. Matrone, E. G. C. Ogliari, A. Nespoli, G. Gruosso, and A. Gandelli, "Electric vehicles charging sessions classification technique for optimized battery charge based on machine learning," *IEEE Access*, vol. 11, pp. 52444–52451, 2023.
- [9] D. K. Patel, D. Singh, and B. Singh, "A comparative analysis for impact of distributed generations with electric vehicles planning," *Sustain. Energy Technol. Assessments*, vol. 52, Aug. 2022, Art. no. 101840. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S221338821008547>
- [10] D. K. Patel, D. Singh, and B. Singh, "Impact assessment of distributed generations with electric vehicles planning: A review," *J. Energy Storage*, vol. 43, Nov. 2021, Art. no. 103092. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352152X21007970>
- [11] B. Singh and P. K. Dubey, "Distributed power generation planning for distribution networks using electric vehicles: Systematic attention to challenges and opportunities," *J. Energy Storage*, vol. 48, Apr. 2022, Art. no. 104030. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352152X22000731>
- [12] D. Said, "Intelligent photovoltaic power forecasting methods for a sustainable electricity market of smart micro-grid," *IEEE Commun. Mag.*, vol. 59, no. 7, pp. 122–128, Jul. 2021.
- [13] D. Said, "A decentralized electricity trading framework (DETF) for connected EVs: A blockchain and machine learning for profit margin optimization," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 6594–6602, Oct. 2021.
- [14] P. K. Dubey, B. Singh, V. Kumar, and D. Singh, "A novel approach for comparative analysis of distributed generations and electric vehicles in distribution systems," *Electr. Eng.*, pp. 1–20, Oct. 2023, doi: [10.1007/S00202-023-02072-2](https://doi.org/10.1007/S00202-023-02072-2).
- [15] E. S. Xydias, C. E. Marmaras, L. M. Cipcigan, A. S. Hassan, and N. Jenkins, "Forecasting electric vehicle charging demand using support vector machines," in *Proc. 48th Int. Universities Power Eng. Conf. (UPEC)*. Dublin, Ireland: IEEE, Sep. 2013, pp. 1–6.
- [16] J. Zhu et al., "Electric vehicle charging load forecasting: A comparative study of deep learning approaches," *Energies*, vol. 12, no. 14, p. 2692, Jul. 2019.
- [17] J. Zhu, Z. Yang, Y. Chang, Y. Guo, K. Zhu, and J. Zhang, "A novel LSTM based deep learning approach for multi-time scale electric vehicles charging load prediction," in *Proc. IEEE Innov. Smart Grid Technol.-Asia (ISGT Asia)*. Chengdu, China: IEEE, May 2019, pp. 3531–3536.
- [18] J. W. Taylor, "An evaluation of methods for very short-term load forecasting using minute-by-minute British data," *Int. J. Forecasting*, vol. 24, no. 4, pp. 645–658, Oct. 2008.
- [19] L. Calearo, M. Marinelli, and C. Ziras, "A review of data sources for electric vehicle integration studies," *Renew. Sustain. Energy Rev.*, vol. 151, Nov. 2021, Art. no. 111518.
- [20] Q. Gao et al., "Charging load forecasting of electric vehicle based on Monte Carlo and deep learning," in *Proc. IEEE Sustain. Power Energy Conf. (iSPEC)*. Beijing, China: IEEE, Nov. 2019, pp. 1309–1314.
- [21] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
- [22] T. Hu, K. Liu, and H. Ma, "Probabilistic electric vehicle charging demand forecast based on deep learning and machine theory of mind," in *Proc. IEEE Transp. Electrific. Conf. Expo (ITEC)*. Chicago, IL, USA: IEEE, Jun. 2021, pp. 795–799.



- [23] T. Hu, H. Ma, H. Liu, H. Sun, and K. Liu, "Self-attention-based machine theory of mind for electric vehicle charging demand forecast," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 8191–8202, Nov. 2022.
- [24] J. Zhu, Z. Yang, Y. Guo, J. Zhang, and H. Yang, "Short-term load forecasting for electric vehicle charging stations based on deep learning approaches," *Appl. Sci.*, vol. 9, no. 9, p. 1723, Apr. 2019, doi: [10.3390/app9091723](https://doi.org/10.3390/app9091723).
- [25] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting," 2017, *arXiv:1705.04378*.
- [26] M. Wood, E. Ogliari, A. Nespoli, T. Simpkins, and S. Leva, "Day ahead electric load forecast: A comprehensive LSTM-EMD methodology and several diverse case studies," *Forecasting*, vol. 5, no. 1, pp. 297–314, Mar. 2023.
- [27] *ACN-Data—A Public EV Charging Dataset*. Accessed: Nov. 2022. [Online]. Available: <https://ev.caltech.edu/dataset>
- [28] M. Dabbaghjamesh, A. Moeini, and A. Kavousi-Fard, "Reinforcement learning-based load forecasting of electric vehicle charging station using Q-learning technique," *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 4229–4237, Jun. 2021.
- [29] L. Sehovac and K. Grolinger, "Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention," *IEEE Access*, vol. 8, pp. 36411–36426, 2020.
- [30] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *Int. J. Forecasting*, vol. 30, no. 2, pp. 357–363, Apr. 2014.
- [31] F. D. Rueda, J. D. Suárez, and A. del Real Torres, "Short-term load forecasting using encoder–decoder WaveNet: Application to the French grid," *Energies*, vol. 14, no. 9, p. 2524, Apr. 2021.
- [32] *éCO<sub>2</sub>mix—La Production D'électricité Par Filière*. Accessed: Nov. 2022. [Online]. Available: <https://www.rte-france.com/eco2mix/la-production-deletricit-e-par-filiere>
- [33] J. Bedi and D. Toshniwal, "Deep learning framework to forecast electricity demand," *Appl. Energy*, vol. 238, pp. 1312–1326, Mar. 2019.
- [34] *Industry Data Briefs*. Accessed: Nov. 2022. [Online]. Available: [https://knoema.com/data-brief/Energy?utm\\_source=promo&utm\\_campaign=Download+our+latest+ENERGY+Data+Brief](https://knoema.com/data-brief/Energy?utm_source=promo&utm_campaign=Download+our+latest+ENERGY+Data+Brief)
- [35] Y. Xie, Y. Ueda, and M. Sugiyama, "A two-stage short-term load forecasting method using long short-term memory and multilayer perceptron," *Energies*, vol. 14, no. 18, p. 5873, Sep. 2021, doi: [10.3390/en14185873](https://doi.org/10.3390/en14185873).
- [36] *TEPCO: Electricity Forecast | Download Past Electricity Demand Data*. Accessed: Nov. 2022. [Online]. Available: <https://www4.tepco.co.jp/en/forecast/html/download-e.html>
- [37] W. Zhang, Q. Chen, J. Yan, S. Zhang, and J. Xu, "A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting," *Energy*, vol. 236, Dec. 2021, Art. no. 121492.
- [38] *Data.Zip*. Accessed: Nov. 2022. [Online]. Available: <https://www.dropbox.com/s/no9r15m9nshl63/Data.zip?dl%2C%BC0=>
- [39] Z. Yi, X. C. Liu, R. Wei, X. Chen, and J. Dai, "Electric vehicle charging demand forecasting using deep learning model," *J. Intell. Transp. Syst.*, vol. 26, no. 6, pp. 690–703, Aug. 2021.
- [40] *PlugShare—Mappa Delle Stazioni Di Ricarica Dei Veicoli Elettrici. Trova Un Luogo Dove Caricare*. Accessed: Nov. 2022. [Online]. Available: <https://www.plugshare.com/>
- [41] Q. Chen, W. Zhang, K. Zhu, D. Zhou, H. Dai, and Q. Wu, "A novel trilinear deep residual network with self-adaptive dropout method for short-term load forecasting," *Expert Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115272.
- [42] *Loaddata—Fkeynia*. Accessed: Nov. 2022. [Online]. Available: <https://sites.google.com/site/fkeynia/loaddata>
- [43] M.-T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," 2015, *arXiv:1508.04025*.
- [44] H. Sun, F. Tao, Z. Fu, A. Gao, and L. Jiao, "Driving-behavior-aware optimal energy management strategy for multi-source fuel cell hybrid electric vehicles based on adaptive soft deep-reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4127–4146, Apr. 2023.
- [45] L. Ni, B. Sun, X. Tan, and D. H. K. Tsang, "Mobility and energy management in electric vehicle based mobility-on-demand systems: Models and solutions," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 3702–3713, Apr. 2023.
- [46] S. Li et al., "A multiagent deep reinforcement learning based approach for the optimization of transformer life using coordinated electric vehicles," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 7639–7652, Nov. 2022.
- [47] C. Li, Y. Liao, L. Zou, R. Diao, R. Sun, and H. Xie, "Short-term forecasting of EV charging load using prophet-BiLSTM," in *Proc. IEEE Transp. Electr. Conf. Expo (ITEC Asia-Pacific)*, Oct. 2022, pp. 1–4, doi: [10.1109/ITECAsia-Pacific56316.2022.9942039](https://doi.org/10.1109/ITECAsia-Pacific56316.2022.9942039).
- [48] A. S. Al-Ogaili et al., "Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: Challenges and recommendations," *IEEE Access*, vol. 7, pp. 128353–128371, 2019.
- [49] Y. Amara-Ouali, Y. Goude, B. Hamrouche, and M. Bishara, "A benchmark of electric vehicle load and occupancy models for day-ahead forecasting on open charging session data," in *Proc. 13th ACM Int. Conf. Future Energy Syst.* New York, NY, USA: Association for Computing Machinery, Jun. 2022, pp. 193–207, doi: [10.1145/3538637.3538850](https://doi.org/10.1145/3538637.3538850).
- [50] R. Xue, S. Zhao, and F. Han, "An embedding-driven multi-hop spatio-temporal attention network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 13192–13207, Nov. 2023, doi: [10.1109/TITS.2022.3220915](https://doi.org/10.1109/TITS.2022.3220915).
- [51] B. Singh, P. K. Dubey, and S. N. Singh, "Recent optimization techniques for coordinated control of electric vehicles in super smart power grids network: A state of the art," in *Proc. IEEE 9th Uttar Pradesh Sect. Int. Conf. Electr., Electron. Comput. Eng. (UPCON)*, Dec. 2022, pp. 1–7, doi: [10.1109/UPCON56432.2022.9986471](https://doi.org/10.1109/UPCON56432.2022.9986471).
- [52] M. Palo, B. Schubert, E. Ogliari, J. Wei, K. Gu, and W. Liu, "Detection, features extraction and classification of radio-frequency pulses in a high-voltage power substation: Results from a measurement campaign," in *Proc. IEEE 3rd Int. Conf. Dielectrics (ICD)*, Jun. 2020, pp. 677–680. [Online]. Available: <https://ieeexplore.ieee.org/document/9342022>



**Silvana Matrone** (Student Member, IEEE) received the B.Sc. degree in mechanical engineering from the University of Naples Federico II, Italy, in 2019, and the M.Sc. degree in energy engineering from Politecnico di Milano, Italy, in 2022, where she is currently pursuing the Ph.D. degree with the Electrical Division, Department of Energy. Her research interests include machine learning techniques with applications in energy dispatching systems.



**Emanuele Ogliari** (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Politecnico di Milano, Milan, Italy, in 2016. He is currently an Associate Professor with the Department of Energy, Politecnico di Milano, where he teaches electrical engineering and photovoltaic-based systems. He is also a Staff Member with the SolarTech<sup>Lab</sup> and the Laboratory of MicroGrids (MG<sup>2</sup>LAB), Politecnico di Milano.



**Alfredo Nespoli** (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in energy engineering from Politecnico di Milano, Italy, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree. Since then, he has been a Temporary Researcher with the Electrical Division, Department of Energy, Politecnico di Milano. His research interests include AI and renewable energy sources integration in energy systems.



**Sonia Leva** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Politecnico di Milano, Italy, in 2001. She is currently a Full Professor of electrical engineering with the Department of Energy, Politecnico di Milano, where she is also the Director of the Solar Tech Laboratory (SolarTech Lab) and the Laboratory of MicroGrids (MG Lab). She is a member of the IEEE Working Group "Distributed Resources: Modeling and Analysis" and a Senior Member of the IEEE Power and Energy Society.