IET Electrical Systems in Transportation

The Institution of Engineering and Technology WILEY

CASE STUDY

Short-term load forecasting at electric vehicle charging sites using a multivariate multi-step long short-term memory: A case study from Finland

Tim Unterluggauer[1,2,3] [ID] | Kalle Rauma[2] [ID] | Pertti Järventausta[1] | Christian Rehtanz[2]

[1]Unit of Electrical Engineering, Tampere University, Tampere, Finland

[2]Institute of Energy Systems, Energy Efficiency and Energy Economics, TU Dortmund University, Dortmund, Germany

[3]Department of Electrical Engineering, Technical University of Denmark, Denmark

Correspondence

Kalle Rauma, Institute of Energy Systems, Energy Efficiency and Energy Economics, TU Dortmund University, Emil-Figge-Str. 76, 44227 Dortmund, Germany.
Email: kalle.rauma@tu-dortmund.de

Funding information

The German Federal Ministry of Transportation and Digital Infrastructure

Open access funding enabled and organized by Projekt DEAL.

Abstract

This study assesses the performance of a multivariate multi-step charging load prediction approach based on the long short-term memory (LSTM) and commercial charging data. The major contribution of this study is to provide a comparison of load prediction between various types of charging sites. Real charging data from shopping centres, residential, public, and workplace charging sites are gathered. Altogether, the data consists of 50,504 charging events measured at 37 different charging sites in Finland between January 2019 and January 2020. A forecast of the aggregated charging load is performed in 15-min resolution for each type of charging site. The second contribution of the work is the extended short-term forecast horizon. A multi-step prediction of either four (i.e., one hour) or 96 (i.e., 24 h) time steps is carried out, enabling a comparison of both horizons. The findings reveal that all charging sites exhibit distinct charging characteristics, which affects the forecasting accuracy and suggests a differentiated analysis of the different charging categories. Furthermore, the results indicate that the forecasting accuracy strongly correlates with the forecast horizon. The 4-time step prediction yields considerably superior results compared with the 96-time step forecast.

1 | INTRODUCTION

In 2017, the transport sector was responsible for 27% of all greenhouse gas emissions in the European Union, with passenger cars accounting for 44% of the transport emissions [1]. Electric vehicles (EVs) are one of the solutions to cut carbon emissions in the transport sector and achieve the climate protection goals. Other related concerns, such as urban air pollution and its impact on health, have also encouraged politicians to promote the adoption of EVs [2].

The resulting anticipated large-scale EV rollout represents both a considerable challenge and an opportunity for the power system. On the one hand, the simultaneous charging of a large number of EVs could lead to severe bottlenecks in the distribution network, requiring costly grid upgrades [2–4]. On the other hand, EVs could also benefit the power system by providing ancillary services. Due to their fast-response ability, high degree of flexibility, and energy storage capacity, EVs

could facilitate the seamless integration of intermittent renewable energy sources into the power grid [5–7]. A large number of EVs could also benefit transmission system operators by providing their flexibility in the form of a control reserve to the energy market [8, 9], or assist distribution system operators in terms of voltage regulation, congestion management, peak shaving, and valley filling measures [10–12].

For both cases, an accurate short-term EV charging load forecast is of utmost importance. For grid operators, the aggregated charging load forecast is relevant to detect bottlenecks in the distribution grid at an early stage and initiate appropriate measures [13]. Aggregators, who aggregate a large number of EVs, also rely on accurate capacity predictions to be able to market EV flexibilities in the energy market for ancillary services [4, 14].

With this motivation in mind, this work introduces a novel charging load prediction application based on the long short-term memory. A multivariate multi-step forecasting approach

is presented, which considers important determinants of the load curve as multivariate inputs and allows the forecast of the whole prediction horizon in a single computation.

The subsequent parts of this paper are organised as follows. Section 2 presents the state of the art in recent studies. Section 3 deals with the basics of LSTMs and the origin of the data used in this work. Section 4 introduces the methodology of the charging load forecast. Section 5 presents the results, which are further discussed in Section 6. Conclusions and directions for future research are provided in Section 7.

2 | RELATED RESEARCH

In this section, a literature review of the state of the art in the field of EV short-term charging load forecasting is conducted to identify the shortcomings in previous research and highlight the contributions of this work.

2.1 | Linear methods

Linear forecasting approaches use linear functions to model time series behaviour. Popular methods include autoregressive, moving average, autoregressive moving average, and autoregressive integrated moving average (ARIMA) processes. Concerning the prediction of the EV charging load, ARIMA models in different versions are particularly popular among all linear approaches. A plain ARIMA model is introduced in [15] to forecast the day-ahead EV charging demand in 15-min (min) resolution. The prediction error is lowered with larger aggregation. Another work [16] examines the applicability of an ARIMA model to predict the EV charging demand of the next 24 h (h). The proposed decoupled forecaster, which independently predicts the charging demand and the conventional electrical load, significantly reduces the prediction error compared to an integrated forecaster.

The works in [17–19] present a more sophisticated version of the ARIMA model. While paper [17] introduces a fractional autoregressive integrated moving average (FARIMA) model, article [18] proposes several seasonal autoregressive integrated moving average (SARIMA) approaches for EV charging load forecasting. The FARIMA model outperforms various ARIMA models for all forecast horizons up to 120 min. The SARIMA models achieve superior results compared to a persistence forecast and a modified pattern sequence-based forecast (MPSF). Similar to the work in [18], study [19] introduces two seasonal autoregressive integrated moving average models with exogenous variables (SARIMAX). Compared to a random forest (RF) and gradient boosted regression tree (GBRT) model, the SARIMAX models outperform both machine learning algorithms for different forecast horizons.

2.2 | Non-linear methods

Contrary to the results in [19], several studies indicate a superior performance of non-linear methods for EV charging load predictions. Compared to their linear counterparts, non-linear approaches rely on non-linear functions to capture more complex time series behavior. Study [20] applies four different algorithms, namely the RF, support vector regression (SVR), time weighted dot product based nearest neighbor (TWDPNN), and MPSF for 24-h charging load forecasting. The evaluation is based on two different datasets – charging and station measurements. The most accurate predictions based on the charging measurement are achieved by using the TWDPNN approach, whereas the MPSF provides the most precise results for the station measurement dataset. An approved RF algorithm is applied to predict the EV charging load of single and grouped EV charging stations in [21]. The analysis reveals that in contrast to an aggregated prediction, more precise results are obtained while performing individual forecasts.

Moreover, considerable attention has been paid to artificial neural networks (ANNs) lately. The predictive performance of different neural networks (NN) is frequently addressed in the literature [22–25]. Article [22] performs a 24-h charging load forecast based on historical driving patterns. Three different ANNs – a simple ANN, a rough artificial neural network (R-ANN), and a recurrent rough artificial neural network (RR-ANN) – are deployed and compared to a Monte Carlo simulation (MCS). RR-ANNs generate the most accurate predictions. In [23], a novel reinforcement learning technique is introduced to forecast the EV charging load under three scenarios – uncoordinated, coordinated, and smart charging. The proposed Q-learning technique increases the prediction accuracy compared to a simple ANN and a recurrent neural network (RNN). The work carried out in [24] studies the short-term charging load predictability of four different types of ANNs – a deep neural network (DNN), an RNN, an LSTM, and a gated recurrent unit (GRU). GRUs with one hidden layer achieve the most accurate prediction results. Study [25] provides a further evaluation of different ANN approaches for super short-term charging load forecasting. The predictive performance of a simple ANN, an RNN, an LSTM, a bidirectional LSTM, a GRU, and a stacked auto-encoder (SAE) are analyzed. The LSTM yields the best performance in both case studies – the charging in public and at commercial buildings.

Similar to the results in [25], articles [26–28] demonstrate the superior performance of the LSTMs for EV load forecasting. The study in [26] develops an LSTM for single-step predictions. The analysis demonstrates that the LSTM outperforms a simple ANN, both for a forecast horizon of 15 and 30 min and that the prediction error diminishes with decreasing forecast horizon. Another LSTM forecasting approach is presented in [27]. The work considers four types of EVs – private and commercial EVs, electric busses, and electric taxis. Based on the MCS, the LSTM outperforms a backpropagation (BP) network and SVR. Paper [28] proposes an LSTM-based approach with hybrid classification to forecast EV travel behaviour and their electrical demand. Compared to copula, quasi-MCS, and MCS, the novel approach improves the forecast accuracy, thus, resulting in lower EV aggregator costs. The research in [29] presents a hybrid model of extreme gradient boosting and LSTM for domestic charging station load forecasting. Multivariate inputs are used to predict

the EV consumption in the next period. Study [30] proposes a hybrid approach of a deep belief network (DBN) and the LSTM network for short-term EV load prediction. The LSTM-DBN achieves superior forecasting results compared to a single DBN and single LSTM.

2.3 | Hybrid methods

Similar to studies [29, 30], other hybrid approaches as compounds of different models have gained great popularity in recent studies. Paper [33] combines a deterministic bottom-up approach with a probabilistic RF model for the short-term forecast of the aggregated load of 46 privately owned EVs in Austin, Texas. The proposed model shows superior performance compared to a persistence forecast, but similar to a GBRT model. The study in [31] combines a least squares support vector machine (LSSVM), fuzzy clustering (FC), and a wolf pack algorithm (WPA) to predict the load of electric bus charging stations. The FC-WPA-LSSVM model outperforms a WPA-LSSVM, a regular LSSVM and a BP-NN. Article [32] proposes another hybrid approach, combining a lion algorithm with the niche immune (NILA) and a convolutional neural network (CNN), for short-term EV charging load forecasting. Taking into account the multivariate inputs, the NILA-CNN shows superior performance to a single CNN, a lion algorithm, CNN, and an SVM. The work in [34] introduces a hybrid model, based on wavelet decomposition (WT), convolutional neural networks, and probabilistic queuing, for day-ahead charging load predictions. The WT-CNN outperforms a BP-NN, an SAE, an RNN, an SVR, a time-delayed neural network (TDNN), and a growing DBN.

2.4 | Research gaps and contributions of this work

Table 1 outlines the main findings of the literature review. All in all, three major limitations can be summarised. The studies either lack real charging data [16, 17, 22, 23, 27] or the data employed is outdated, making it difficult to transfer the results to today's strongly changed situation in terms of a larger number of EVs, different charging powers, and the altered utilization of EVs [15, 18–20, 33, 34]. Moreover, most of the works only perform single-step predictions [16, 20, 21, 23–27, 29, 31, 32] which limits the scope of practical applications due to the short forecast horizon. Lastly, although in Ref. [25] it has been indicated that the achievable prediction accuracy may be dependent on the type of charging, no attention has been paid to the analysis and comparison of different charging sites.

This work contributes to the identified research gaps in the following ways:

- Introducing a novel multivariate, multi-step LSTM application, which significantly extends the existing short-term forecast horizon by providing a 4-time step (ts) or 96-ts prediction at once,

- Comparing and assessing the forecasting accuracy at different categories of charging sites which exhibit different charging characteristics, and
- Using a large amount of high-quality charging data from various commercial charging sites.

3 | DESCRIPTION OF THE DATASET AND LSTM

This section deals with both the characteristics of the real charging data used in this work and the LSTM fundamentals.

3.1 | Introduction of the dataset

This study is based on real charging data of 50,504 slow charging events with a maximum capacity of 22 kW, measured at 37 different charging sites in Finland between January 2019 and January 2020. While the smallest charging site contains less than 10, the largest one features nearly 300 charging stations. An overview of the characteristics of the charging data is illustrated in Table 2.

The charging events are clustered according to the four categories of charging sites selected for this work. The data for the first category, the charging at shopping centres (SCc), stems from one charging site. The charging events of 21 sites are assigned to the second category named residential charging (REc), representing the charging at home. The third category, public charging (PUc), covers the charging at car parks that can be used by all EV drivers. The last category, workplace charging (WOc), refers to EV charging at company premises that can only be used by the staff working in the vicinity. Eight and seven sites are assigned to PUc and WOc, respectively.

3.2 | Long short-term memory

The charging load of EVs is typically subject to strong time dependencies as the load corresponds to cyclical and seasonal patterns. The LSTM is able to overcome the vanishing or exploding gradient problem of RNNs, and to learn both long-term and short-term dependencies, thus making it an ideal choice for this work [35]. The ability to learn temporal correlations stems from the unique structure of an LSTM cell, illustrated in Figure 1. The LSTM at a single ts t consists of three gate units, namely the forget $\boldsymbol{f}_t$, input $\boldsymbol{i}_t$, and output gate $\boldsymbol{o}_t$.

The functionality of the LSTM cell can be expressed mathematically by the following formulas [36]

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{f,x}x_t + \boldsymbol{W}_{f,h}h_{t-1} + \boldsymbol{b}_f\right) \quad (1)$$

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{i,x}x_t + \boldsymbol{W}_{i,h}h_{t-1} + \boldsymbol{b}_i\right) \quad (2)$$

TABLE 1 Literature review on studies addressing the issue of short-term EV charging load forecasting

Year	ref.	Proposed model(s)	Baseline model(s)	Superior model	Forecast approach	Reso-Lution	Data characteristics	Data origin [year]
2016	[16]	ARIMA	–	Decoupled forecast	24-h (single-step)	1-h	Travel behaviour	USA
2016	[20]	MPSF, SVR,RF, TWDPNN	–	MPSF, TWDPNN	24-h (single-step)	1-h	Charging &Station records	UCLA (USA) [Dec 2011 - Feb 2014]
2016	[17]	FARIMA	ARIMA	FARIMA	2-h	1.2-min	Travel behaviour	Seattle (USA)
2017	[18]	SARIMA	MPSF,Persistence	SARIMA	1-, 2-, & 24-h	1-h	Charging data (aggregated)	California (USA) [Jan 2011 - Jan 2013]
2018	[31]	FC-WPA-LSSVM	BP-NN, LSSVM, WPA-LSSVM	FC-WPA-LSSVM	24-h (multivariate & single-step)	1-h	Electric bus charging data	Baoding (China) [2017]
2018	[32]	NILA-CNN	CNN, SVM,Lion CNN	NILA-CNN	24-h (multivariate & single-step)	30-min	Public charging (aggregated)	Beijing (China) [Jun 2017 - Nov 2017]
2018	[21]	RF	–	–	multivariate & single-step	15-min	Public charging (single & aggr.)	Shenzhen (China) [2016 - 2018]
2019	[15]	ARIMA	–	–	24-h	15-min	Non-residential charging (aggr.)	California (USA) [2013]
2019	[22]	ANN, R-ANN,RR-ANN	MCS	RR-ANN	24-h	1-h	Travel behaviour	USA [2017]
2019	[26]	LSTM	ANN	LSTM	24-h (Single-step)	15- &30-min	Public charging (aggregated)	Shenzhen (China) [Jul 2017 - Jul 2018]
2019	[25]	ANN, GRU, LSTM, RNN, SAE, DNN,	–	LSTM	1-, 5-, & 15-min (single-step)	1-min	Public & commercial charging (aggr.)	Shenzhen (China) [Jun 2017 - Jul 2018]
2019	[24]	GRU, bidirectional LSTMLSTM, RNN	–	GRU	24-h (multivariate & single-step)	1-h	Public charging (aggregated)	Shenzhen (China) [Apr 2017 - Jun 2018]
2019	[33]	Bottom-up RF	GBRT,Persistence	Bottom-up RF	24-h (multivariate)	–	Residential charging (aggr.)	Austin (USA) [2015]
2019	[19]	GBRT, RF,SARIMAX	Persistence	SARIMAX	7-, 14-, & 28-days	24-h	Charging data (aggregated)	Netherlands [Jan 2012 - Mar 2016]
2019	[27]	LSTM	BP, SVR	LSTM	Univariate & single-step	1-h	Simulated charging data	–
2020	[23]	Q-learning technique	ANN, RNN	Q-learning	24-h (single-step)	1-h	Simulated charging data	–
2020	[34]	WT-CNN & probabilistic queuing model	BP-NN, SVR, SAE, TDNN, RNN, DBN	WT-CNN	24-h (multi-step)	1-h	Traffic flow data	England, U.K [Jan 2014 - Dec 2014]
2020	[30]	LSTM-DBN	DBN, LSTM	LSTM-DBN	Multivariate	15-min	Charging data (aggregated)	Liaoning, China [Jan 2018 - Mar 2018]
2020	[28]	LSTM	Copula, MCS, quasi MCS	LSTM	Day-ahead	1-h	Travel behaviour	USA [2017]
2021	[29]	Extreme gradient boosting LSTM	–	–	Multivariate & single-step	1-h	Charging data (aggregated)	Jiangsu, China [Jan 2017 - Dec 2018]

T A B L E 1 (Continued)

Year	ref.	Proposed model(s)	Baseline model(s)	Superior model	Forecast approach	Reso-Lution	Data characteristics	Data origin [year]
2021	This paper	Novel LSTM implementation with multivariate inputs and multi-step forecast horizon	Comparison of different charging sites & forecast horizons		1-h & 24-h (multivariate & multi-step)	15-min	Charging data (aggregated)	Finland [Jan 2019 – Jan 2020]

Abbreviations: ANN, Artificial neural network; ARIMA, autoregressive integrated moving average; CNN, Convolutional neural network; DBN, Deep belief network; DNN Deep neural network EV, electric vehicle; FARIMA, Fractional autoregressive integrated moving average; FC, Fuzzy clustering; GRU, gated recurrent unit; LSTM, long short-term memory; MCS, Monte carlo simulation; MPSF, modified pattern sequence-based forecast; NILA, Lion algorithm by niche immune; R-ANN, rough artificial neural network; RR-ANN, recurrent rough artificial neural network; RF, Random forest; SAE, stacked auto-encoder; SARIMAX, Seasonal autoregressive integrated moving average (with exogenous variables); TWDPNN, time weighted dot product based nearest neighbour; SARIMA, Seasonal autoregressive integrated moving average; WPA, Wolf pack algorithm.

T A B L E 2 Characteristics of the EV charging data

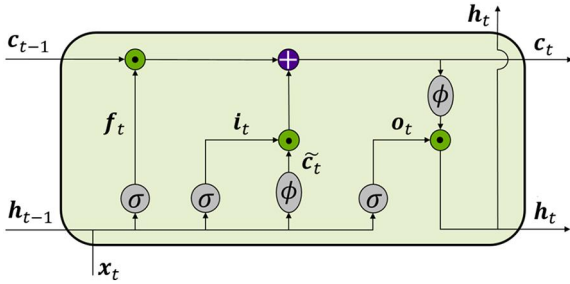| Type of charging	# Charging sites	# Charging events	Measured period
Shopping centre	1	9283	01.01.19 –31.12.19
Residential	21	10,920	21.01.19–20.01.20
Public	8	18,785	31.01.19–30.01.20
Workplace	7	11,516	31.01.19–30.01.20

Abbreviations: EV, electric vehicle.

F I G U R E 1 Architecture of an long short-term memory cell at a single time step t.

$$\tilde{c}_t = \phi\left(W_{\tilde{c},x}x_t + W_{\tilde{c},h}h_{t-1} + b_{\tilde{c}}\right) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (4)$$

$$o_t = \sigma\left(W_{o,x}x_t + W_{o,h}h_{t-1} + b_o\right) \quad (5)$$

$$h_t = o_t \odot \phi(c_t) \quad (6)$$

$W_{f,x}$, $W_{f,h}$, $W_{i,x}$, $W_{i,h}$, $W_{o,x}$, $W_{o,h}$, $W_{\tilde{c},x}$, and $W_{\tilde{c},h}$ label weight matrices, b_f, b_i, b_o, and $b_{\tilde{c}}$ are bias vectors. Operations $+$ and $\odot$ denote the element-wise addition and multiplication, and ϕ depicts the tanh activation function. The forget gate determines which information from the previous cell state, c_{t-1}, is preserved. The cell state represents the long-term memory of the LSTM. The input gate decides which information from the candidate cell state, $\tilde{c}_t$, is to be used to update the previous cell state. The output gate controls which part of the new cell state, c_t, to output and pass as the hidden state, h_t, the short-term memory, to the next LSTM cell. The current input, x_t, and the previous hidden state, h_{t-1}, are multiplied with their respective weights along the bias vectors form the inputs for all three gates. The sigmoid function, σ, introduces non-linear characteristics to the gates and decides which signals pass the gates. While a value of zero causes signals to disappear, a value of one ensures that the signals pass the gate.

4 | METHODOLOGY

The methodology of the proposed multivariate multi-step LSTM is shown in Figure 2 and consists of three main stages – data preprocessing, LSTM implementation and training, and LSTM forecasting. EV, electric vehicle; LSTM, long short-term memory

4.1 | Data preprocessing

The following paragraph details the data preprocessing.

4.1.1 | Aggregated charging load time series generation

In the first step, the event-based charging data is converted into a time series of aggregated charging load values in 15-min resolution under the assumption of a constant charging power.
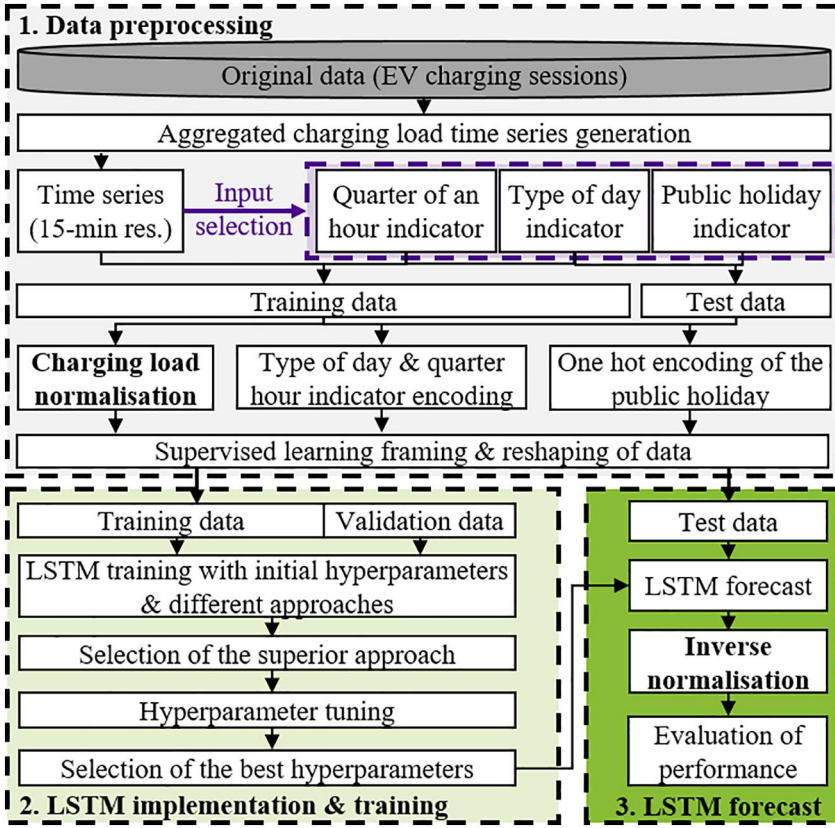
FIGURE 2 Methodology of the LSTM EV charging load forecast

The data contains both the plug-in- and power-to-zero-timestamps and the original amount of energy, $energy_{or}$, charged for each charging event. Consequently, initially, the original charging time, t_{or}, in min is derived by measuring the deviation between both timestamps. In case the power-to-zero-timestamp is missing due to measurement errors, a nominal charging power of 1.8 kW is assumed as most of the charging events indicate charging at low powers. The original charging time is calculated by dividing the original amount of energy charged by the assumed charging power. Subsequently, since the charging times might change due to the analysis interval of 15 min, the original charging time is adjusted subsequently. If the original charging time lasts less than 15 min, it is rounded up to a full 15-min interval. Otherwise, a modulo operation, mod, is performed and the modified charging time, t_{new}, is calculated according to (7).

$$t_{new} = \begin{cases} 15 & t_{or} < 15\text{min} \\ t_{or} - \text{mod}(t_{or}) & \text{mod}(t_{or}) < 7.5\text{min} \\ t_{or} - \text{mod}(t_{or}) + 15 & else \end{cases} \quad (7)$$

Furthermore, the modified charging load, $load_{new}$, of each charging process is calculated according to (8) under the premise of a constant amount of energy charged. Lastly, by aggregating the calculated load of each charging event, the charging load time series is generated, spanning one year with 96 load values for each day.

$$load_{new} = \frac{energy_{or}}{t_{new}} \quad (8)$$

4.1.2 | Multivariate input selection

To determine the important features to be fed to the LSTMas multivariate inputs to support the detection of interrelationships, an analysis of the characteristics of the charging load is carried out. Figure 3 illustrates the development of the aggregated charging load over the course of a year at all four categories of charging sites and reveals two particularities. First, an increase of the aggregated load over the course of the year can be observed for all charging sites. The strong impact of holiday periods on REc, PUc, and WOc marks the second conspicuity. The Finnish summer holidays in July and the winter holidays in December noticeably reduce the load. Both findings support the selection of a month indicator as a feature. However, since the data only covers one year, a feature month would lead to unknown indicators during the data split and is, thus, omitted.

Figure 4 provides a detailed analysis of the charging sites by visualising the charging load for three different weeks of the year. When comparing the load curve of the first week of February (green) with that of the end of November (grey), the increase in charging load over the course of a year is evident. Moreover, a clear impact of the type of day and time of the day
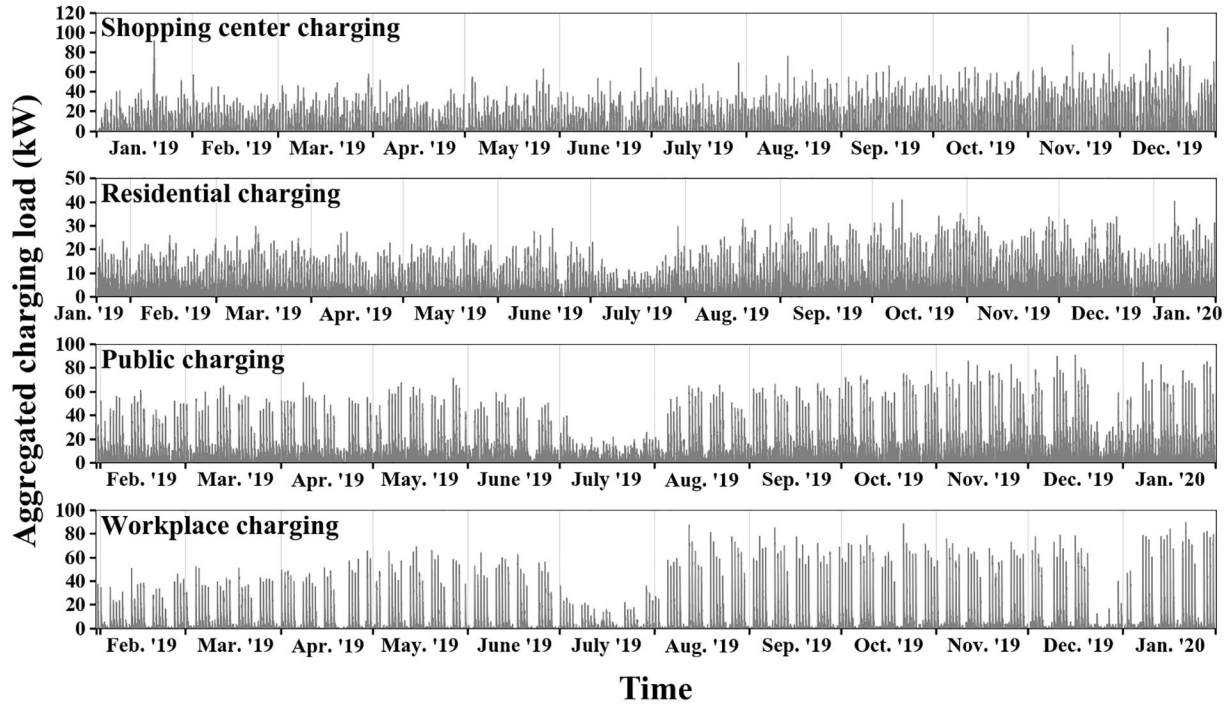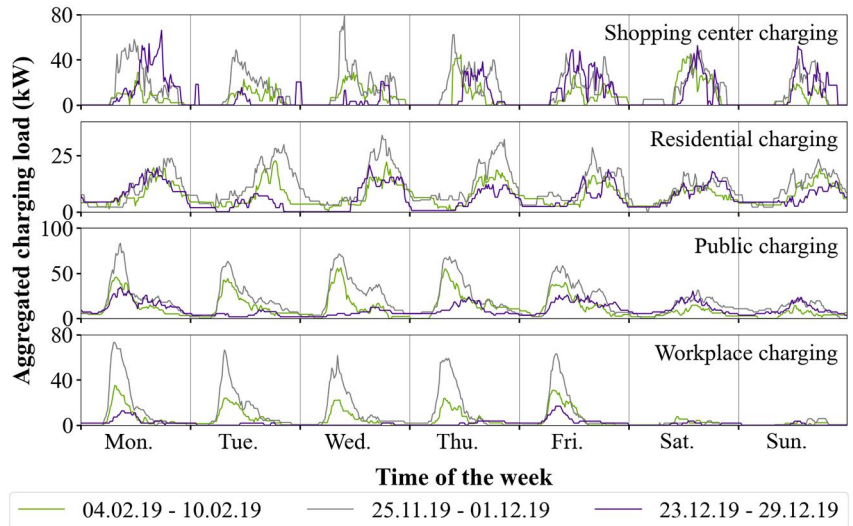
FIGURE 3 Aggregated electric vehicle charging load over the course of one year illustrated for each category of charging site

FIGURE 4 Aggregated charging at each category of charging site exemplified for three different weeks

can be seen for REc, PUc, and WOc. The charging load varies significantly between weekdays and weekends, and the time of the peak load follows a clear pattern. SCc is the only charging site that exhibits a highly random load pattern. When looking at the purple plot, the influence of public holidays on the load is evident as well. For all charging sites, the load is considerably reduced during the three Christmas holidays. Consequently, to help the LSTM to detect interrelationships, each load value is linked to a quarter hour indicator, type of day indicator, and public holiday indicator, collectively representing the multivariate inputs to the LSTM.

4.1.3 | Scaling and encoding of data

To avoid data leakage, the time series is split into training and test sets with a 90% to 10% ratio prior to the scaling and encoding. Furthermore, 20% of the training data is used to validate the model performance during training, resulting in a 72% training, 18% validation, and 10% test split.

To accelerate network convergence and allow the LSTM to process categorical features, the load is normalised, and the indicators are encoded. The commonly used min-max scaling method is used to normalise the load values [37]. The popular

one hot encoding technique is used to encode the binary character of the public holiday indicator [38]. For the quarter hour and type of day indicators, sine/cosine encoding is tested as well to allow a better representation of the cyclic nature of both indicators [39, 40].

4.1.4 | Supervised learning framing and reshaping of data

Finally, the data is converted into a supervised learning problem and formatted such that the shape meets the specific LSTM requirements. In the first step, the normalised and scaled data is concatenated into a single array for every 15-min ts. Using one hot encoding, each ts possesses 106 dimensions of feature space, this number is reduced to seven while using the sine/cosine encoding.

Subsequently, the data is framed in a supervised learning manner by separating the data into input and output sequences. In this study, two approaches are examined – the stateless and stateful mode. The sliding window approach is applied for the stateless mode. In the stateful mode, the data is prepared in such a way that the fixed input sequences adjoin each other without overlapping. For both cases, the input data contains the load value and all indicators for each ts, the output only contains the load to be compared with the predicted load.

In the last step, the data is reshaped in the required 3-dimensional shape, defining the number of samples, the number of input ts, and the number of features for each ts.

4.2 | Long short-term memory implementation and training

The most suitable LSTM configuration is selected in two steps. First, the LSTM is trained with initially selected hyperparameters (hp) on different approaches: the two encoding techniques, the two modes, and a varying number of input ts.

In stateless mode 4, 8, 16, and 32 ts are tested as inputs for the 4-ts forecast, 96, and 192 ts for the 96-ts prediction. The equal amount of input and output ts are applied in the stateful mode. Subsequently, based on the minimal validation loss, the most suitable approach undergoes tuning.

4.2.1 | Training and hyperparameter tuning implementation

The LSTM is implemented as illustrated in Figure 5, using the Sequential Model in Keras. While indices n denote the number of input features, indices tin depict the amount of input ts for each sample. Indices $tout$ define the number of predicted load values subsequent to the input sequence.

Table 3 provides an overview of the hp selected for initial training and hp tuning. The tuning, based on a random search, is performed in hyperopt, a popular library for conducting hp optimisation in Python. A detailed description of the hyperopt is given in [40]. The number of evaluations is set to 50 and the seeding is set to one to ensure the comparability between the different charging sites and reproducibility of the results.

4.2.2 | LSTM training process

The training proceeds as follows. The training and validation data is shown to the LSTM in the input layer. Both data sets consist of multiple samples which are successively processed by the hidden layer(s), each comprising the input and output sequences. For each sample, the concatenated features of each ts form the current input for the LSTM cell. In the stateless mode, the final state of each processed batch is removed. In the stateful mode, however, the final state is provided as the initial state for each sample in the next batch. The state is manually reset after each epoch as each epoch contains the same time series data. After the input data has been processed by either one or two hidden layers, the final state of the last cell
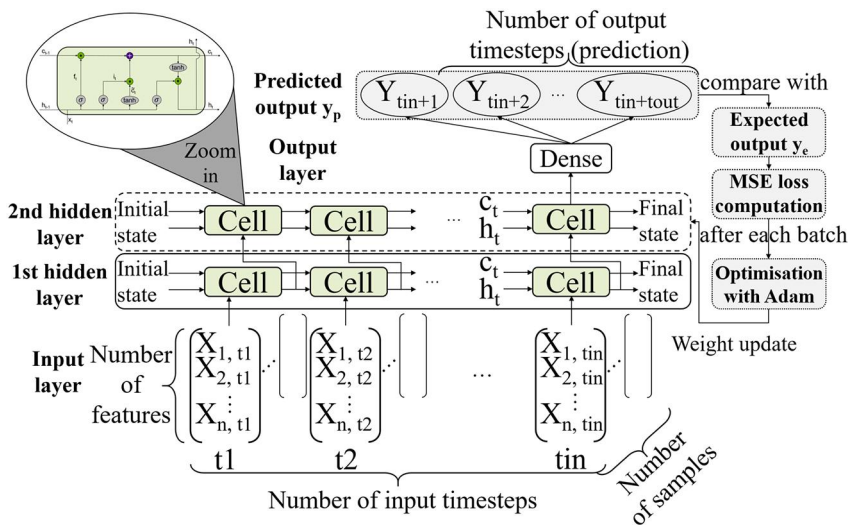

FIGURE 5 Architecture and training process of the long short-term memory. MSE, mean squared error

T A B L E 3 Initial hyperparameters and search space

Type of hp	Initial Training	Hp Tuning
Optimiser	Adam	No tuning
Loss function	MSE	No tuning
Activation function	Tanh & hard sigmoid (cell), ReLU (dense layer)	No tuning
# Epochs	500	1000 (no tuning)
# Hiddenlayers	1	1, 2
# Units	2, 8, 32, 128	2, 4, 8, 16, 32, 64, 128
Learning rate	0.001	0.01, 0.001, 0.0001
Batch size	1 (stateful), 32 (stateless)	16, 32, 64, 96
Dropout [%]	0	0, 20, 50

Abbreviations: MSE, mean squared error; ReLU, Rectified Linear Unit activation function.

in the last hidden layer is passed to a dense layer which outputs the number of ts to be predicted for each input sequence. Once a certain number of samples, defined by the batch size, is processed, the predicted values are compared with the desired real load values to calculate the training and validation loss. Based on the mean squared error (MSE) during training, the weights are then modified after each batch size by the Adam optimiser using backpropagation through time. This process is repeated until the maximum number of epochs is met or the training is terminated prematurely by the implemented early stopping callback.

4.3 | LSTM forecasting

After the completion of the random search, the LSTM with the optimal hp configuration performs the load forecast.

4.3.1 | Implementation of the forecast

The forecast is carried out with a batch size of one, to avoid errors in the prediction process caused by the varying batch sizes during training. For every input sequence of the test data, the model predicts the subsequent four or 96 ts. Afterwards, both the normalised real and predicted load are converted to load values in kW. The inverse normalisation is performed with the initial parameters of the min-max scaler that have been saved during the preprocessing stage for this purpose.

4.3.2 | Metrics

Six different error metrics are employed to evaluate the model performance. The mean absolute error (MAE) in kW, calculated according to (9), measures the precision of the forecast by averaging the error between the predicted and real load, and is selected due to its simple comprehensibility. While N depicts the number of forecasts, $load_t$ indicates the real charging load at time t, and $lo\hat{a}d_t$ the predicted one. However, the MAE is

scale-dependent, which implies the need for additional error metrics to allow a better comparison of the different charging sites. Given its scale-independence, the popular mean absolute percentage error (MAPE) would provide an easily interpretative error metric. However, the charging load time series exhibit a charging load of zero at numerous points in time. For this reason, MAPE cannot be used for overall comparison, as the calculation is based on the division of the error by the true load value at each ts. To overcome these difficulties, two variants of the normalised mean absolute error (NMAE) are used as shown in Equations (10) and (11). The MAE is normalised by the mean charging load and the difference between the maximum $load_{max}$ and minimum load $load_{min}$, respectively.

$$MAE = \frac{1}{N} \sum_{t=1}^{N} |load_t - lo\hat{a}d_t| \qquad (9)$$

$$NMAE1 = MAE \bigg/ \frac{1}{N} \sum_{t=1}^{N} load_t \qquad (10)$$

$$NMAE2 = \frac{MAE}{load_{max} - load_{min}} \qquad (11)$$

For the 96-ts forecast, three additional metrics are used to better assess how well the daily peak load can be predicted, both in terms of temporal occurrence and magnitude of the peak load. The timing and magnitude of the peak load are especially important for assessing possible congestions in the distribution network. The average peak deviation, $pdev$, in kW between the actual daily peak load, $load_{dm}$, and predicted daily peak load, $lo\hat{a}d_{dm}$, is calculated according to Equation (12). While D specifies the number of predicted days, dm denotes the daily peak. Likewise, the MAPE in % is calculated according to Equation (13). Finally, Equation (14) specifies the average time deviation, $tdev$, in min separating the real and predicted peak load. Variable $slot$ corresponds to the respective 15-min interval of the real and the predicted peak load.

$$pdev = \frac{1}{D}\sum_{d=1}^{D}|load_{dm} - lo\hat{a}d_{dm}| \qquad (12)$$

$$MAPE = \frac{1}{D}\sum_{d=1}^{D}(load_{dm} - lo\hat{a}d_{dm})\Big/load_{dm} \times 100 \quad (13)$$

$$tdev = \frac{1}{D}\sum_{d=1}^{D}|slot(load_{dm}) - slot(lo\hat{a}d_{dm})| \times 15 \qquad (14)$$

5 | RESULTS

Having covered the methodology of this work, this chapter addresses the training, tuning, and prediction results.

5.1 | Initial training results

Table 4 illustrates the selected approach for each charging site and forecast horizon based on the minimum validation loss obtained during initial training. In all cases the minimum loss is recorded while training the LSTM with 128 units.

The small epoch number for most of the charging sites indicates that the LSTM suffers from early overfitting. Moreover, the different modes, encoding techniques, and input ts only yield small differences in the validation loss. However, in all the cases the lowest MSE is obtained while using the stateless mode. The sine/cosine encoding shows a slightly superior performance in most cases. Finally, using the same amount of input ts yields the lowest MSE for the 96-ts forecast. For the 4-ts prediction, 16 input ts results in the lowest validation loss for SCc, Rec, and WOc, whereas four input ts achieve the best fit for PUc.

5.2 | Hyperparameter tuning results

The maximum and minimum validation loss obtained during tuning, and the chosen superior hp combination is given in Table 5. The maximum and minimum validation loss differ considerably. The highest MSE values are obtained for the 4-ts prediction while using a high dropout or learning rate. For the 96-ts forecast, a small number of units lead to the highest validation losses. While comparing the minimum loss between random search and initial training, it is evident that the hp tuning only yields improvements for half of the forecasts. The lowest losses are obtained with a high number of units (64 or 128). In most cases, an LSTM with only one hidden layer seems sufficient. Only for the REc 96-ts forecast and WOc 4-ts prediction the minimum loss is recorded while using two hidden layers. All batch sizes are applied for the different forecasts. Dropout generally does not have a positive effect on

TABLE 4 Summary of the best approach after initial training

Site horizon	SCc		REc		PUc		WOc	
	4-ts	*96-ts*	*4-ts*	*96-ts*	*4-ts*	*96-ts*	*4-ts*	*96-ts*
Min. Val. loss[a]	5.29	8.94	2.83	8.35	2.02	9.79	1.85	6.95
# Epochs	17	7	49	31	14	147	221	130
# Input ts	16	96	16	96	4	96	16	96
Mode[b]	SL	SL	SL	SL	SL	SL	SL	SL
Encoding[c]	S/C	OH	S/C	S/C	S/C	S/C	S/C	OH

Abbreviations: PUc, public charging; REc, residential charging; SCc, shopping center charging; WOc, workplace charging.

[a]All MSE values (validation loss) are given in units of 10^{-3}.

[b]SL = Stateless mode.

[c]S/C = Sine/Cosine encoding, OH = One hot encoding.

training, except for the 96-ts REc forecast. Lastly, in most of the cases, the learning rate of 0.001 yields the lowest loss.

5.3 | Forecasting results

In the following sections, the forecasting results are analysed both graphically and numerically.

5.3.1 | Graphical results

Figure 6 illustrates the forecasting results for the 36 days of test data for each category of charging site for both the 4-ts (green) and 96-ts (purple) prediction. While comparing the outcomes with the real charging load (grey), it becomes apparent that the 4-ts prediction achieves superior results compared to the 96-ts forecast. The 4-ts forecast yields a relatively precise picture of the real load curve in grey, with no major outliers to be seen.

Considering the 96-ts prediction results, three short-comings can be identified. First, the LSTM is not capable of predicting the charging load on public holidays. The projected load almost exclusively exceeds the real load by a substantial margin and exhibits a course similar to that of non-holidays. Second, the LSTM is not able to accurately predict the level of real peak load for most days, frequently exceeding the predicted peak load substantially. Finally, the LSTM struggles to anticipate the impact of holiday periods. This is particularly evident for REc, where the forecast significantly exceeds the reduced real charging load during the winter holiday period.

5.3.2 | Numerical results

Table 6 summarises the numerical results. As previously evidenced during the graphical analysis, reducing the forecast horizon from 96 ts to 4 ts considerably reduces the prediction error. The MAE for REc, PUc, and WOc can be more than

TABLE 5 Summary of tuning results and selected hyperparameters

Site horizon	SCc		REc		PUc		WOc	
	4-ts	96-ts	4-ts	96-ts	4-ts	96-ts	4-ts	96-ts
Min. loss[a]	33.68	13.51	10.17	21.04	12.08	32.34	47.24	14.55
Max. loss[a]	5.33	8.78	2.83	8.00	2.03	9.05	1.59	7.25
Improvement	No	Yes	No	Yes	No	Yes	Yes	No
# Hidden layer	1	1	1	2	1	1	2	1
# Units	128	128	128	128	128	64	64	128
Batch size	32	64	32	16	32	96	96	32
Dropout [%]	0	0	0	20	0	0	0	0
Learning rate	0.001	0.01	0.001	0.001	0.001	0.0001	0.0001	0.001

Abbreviations: PUc, public charging; REc, residential charging; SCc, shopping center charging; WOc, workplace charging.
[a]All MSE values are given in units of 10^{-3}.
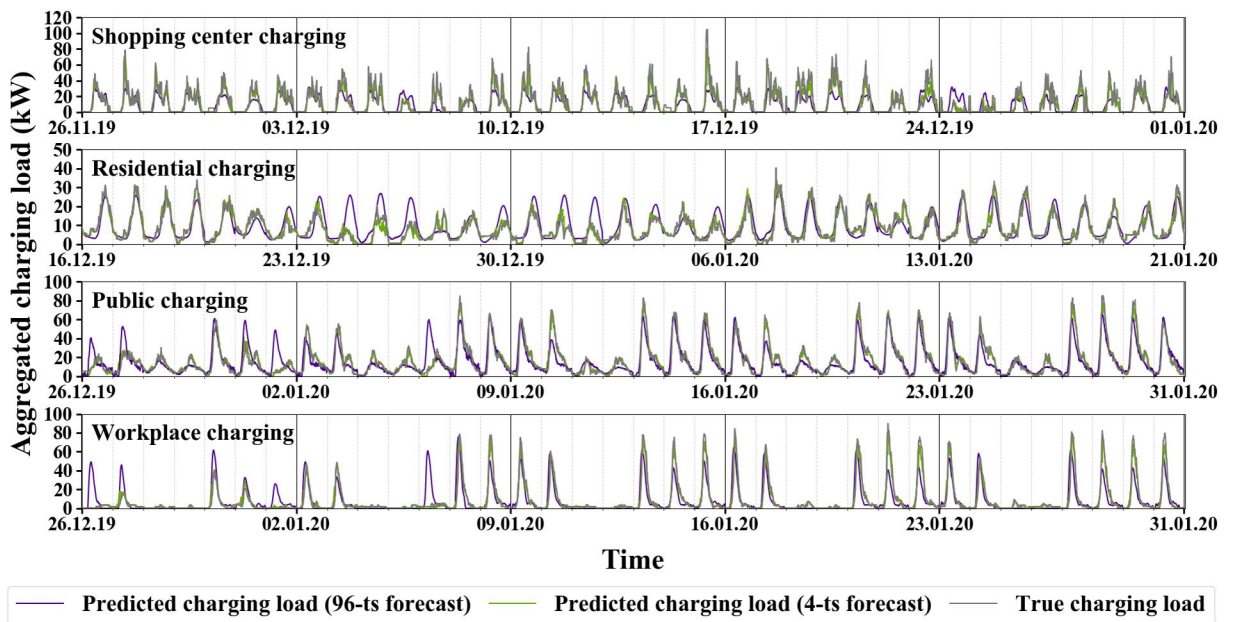
FIGURE 6 Forecasting results for the 36 days of test data illustrated for all categories of charging sites

halved in most of the cases. For SCc, the MAE is reduced in all cases by more than 2 kW. The lowest overall MAE for both the forecast horizons is recorded for REc. WOc generates the second lowest MAE values, followed by PUc. The highest MAE values are seen for SCc.

A different picture emerges while looking at the NMAE1 and NMAE2 results. For the 4-ts prediction, the lowest NMAE1 is obtained for PUc, followed by REc, WOc, and SCc. Conversely, using the NMAE2, the most precise results are seen for WOc, followed by PUc, Rec, and SCc. For the 96-ts forecast the lowest NMAE1 again is obtained for PUc and REc, however, the least accurate prognosis is given for WOc this time. In contrast, the lowest NMAE2 is seen for WOc, followed by SCc and PUc. REc ranks last.

While comparing the clustered NMAE1 results, it can be seen that the 4-ts forecast yields the most accurate forecast for PUc and REc. SCc and WOc score the poorest results.

Conversely, while using the NMAE2 as the evaluation criterion, varying results are visible depending on the type of day. For weekdays the most accurate predictions are seen for WOc followed by PUc, Rec, and SCc. While REc yields the most exact forecast for weekends, WOc, PUc, and SCc take the second, third, and fourth place. For public holidays, REc ranks first again, followed by SCc, PUc, and WOc. Similar findings to the NMAE1 results of the 4-ts forecast can be seen for the NMAE1 results of the 96-ts forecast. REc and PUc achieve the most precise scores in most of the cases. However, for public holidays, SCc outperforms PUc and comes in the second place, behind REc. Looking at the NMAE2 results, the poorest performance for weekdays can be observed for Rec, whereas SCc yields the most accurate results. On weekends SCc ranks first again, and WOc ranks the last. For public holidays, WOc performs particularly poor again. REc obtains the most favourable score.

Table 7 quantifies the LSTM's difficulties in predicting peak loads with a forecast horizon of 96 ts in numerical terms. The pdev, the MAPE, and the tdev are given for each charging site. For all forecasts, the average deviation between true and predicted peak load amounts to 29.16 kW for SCc, 6.37 kW for REc, 14.17 kW for PUc, and 17.85 kW for WOc. The MAPE amounts to 53.19% for SCc, 34.38% for REc, 37.16% for PUc, and 128.97% for WOc. The average deviation between real and predicted time of the peak load in minutes are 153, 88, 104, and 285 for SCc, REc, PUc, and WOc, respectively.

Examining the scores grouped by weekdays, weekends, and public holidays, several discrepancies between the charging sites can be highlighted. For SCc, on weekends and public holidays, the absolute and relative errors can be reduced relative to weekdays. In addition, the time deviation between true and predicted occurrence of the peak load is significantly lowered as well. For REc, in contrast, the highest relative and

absolute errors are seen for public holidays, although the time deviation is minimised. For PUc and WOc, according to the MAPE, the most precise results are achieved during the week. For WOc predictions on weekends, the absolute error amounts to 4.73 kW and the MAPE to 100% due to a constant prediction of 0 kW. Therefore, no deviation between the time of the true and predicted peak load can be calculated. The highest error scores between the real and predicted peak load are recorded on public holidays for both charging sites and the time deviation between true and predicted peak load amounts to a multiple of the deviation during the week.

6 | DISCUSSION

In this section, the initial training, hp tuning, and forecasting results are discussed in more detail.

6.1 | Initial training results

Four key findings can be summarised. To begin with, the inclusion of multivariate inputs shows a positive influence on the training results in terms of validation loss. However, a detailed analysis of the impact on the training results and forecasting results was not part of the analysis. Since the training data used in this work spans less than one year for each category of charging site, the inclusion of multivariate inputs is selected to help the LSTM identify patterns in the data. With multiyear data available, the importance of including multivariate inputs may decrease or become obsolete.

Next, the early overfitting indicates that the complexity of the model, relative to the size of the available data, is too high, and due to changing load characteristics throughout the year, the validation data is not fully representative for the entire dataset. The access to charging data measured over a longer period is, thus, of great interest to enhance the performance of the model in future.

Furthermore, the different encoding techniques and input ts variants only cause negligible differences in the validation loss. Thus, the extra time needed to implement and train the LSTM with various variants cannot be justified. Hence, the selection of the same number of input ts and encoding choice for all sites

TABLE 6 Evaluation of forecasting results

Charging site horizon	SCc		REc		PUc		WOc	
	4-ts	96-ts	4-ts	96-ts	4-ts	96-ts	4-ts	96-ts
MAE[a] (all)	4.35	6.78	1.53	3.37	2.7	5.84	1.85	5.35
NMAE1 (all)	0.352	0.549	0.178	0.391	0.166	0.359	0.204	0.59
NMAE2 (all)	0.041	0.065	0.038	0.083	0.032	0.068	0.021	0.06
MAE[a] (wd)	4.41	6.75	1.52	3.33	3.1	6.37	2.56	6.56
NMAE1 (wd)	0.369	0.564	0.175	0.381	0.152	0.312	0.191	0.489
NMAE2 (wd)	0.042	0.064	0.038	0.082	0.036	0.075	0.029	0.073
MAE[a] (we)	4.39	7.15	1.6	3.62	1.83	3.2	0.49	1.31
NMAE1 (we)	0.329	0.536	0.189	0.428	0.208	0.363	0.377	1.00
NMAE2 (we)	0.06	0.097	0.047	0.106	0.056	0.098	0.054	0.142
MAE[a] (ph)	3.71	5.74	1.36	2.89	2.48	10.64	0.88	9.51
NMAE1 (ph)	0.312	0.484	0.162	0.345	0.262	1.124	0.595	6.454
NMAE2 (ph)	0.073	0.113	0.047	0.099	0.092	0.393	0.16	1.739

Abbreviations: MAE, mean absolute error; NMAE, normalised mean absolute error; ph = public holiday; PUc, public charging; REc, residential charging; SCc, shopping center charging; WOc, workplace charging; wd = weekday, we = weekend.
[a]All MAE values are given in kW

TABLE 7 Summary of the peak deviation, MAPE and time deviation results of the 96-ts prediction

Charging site metric	SCc			REc			PUc			WOc		
	Pdev (kW)	MAPE (%)	Tdev (min)	Pdev (kW)	MAPE (%)	Tdev (min)	Pdev (kW)	MAPE (%)	Tdev (min)	Pdev (kW)	MAPE (%)	Tdev (min)
All days	29.16	53.19	153	6.37	34.38	88	14.17	37.16	104	17.85	128.97	285
Weekdays	30.96	53.65	187	6.24	34.71	86	13.75	23.50	52	20.54	35.50	35
Weekends	26.37	53.11	95	6.29	32.00	99	11.39	43.28	116	4.73	100.00	-
Public holidays	24.65	49.90	85	7.64	39.80	65	26.62	121.46	615	41.01	942.13	435

Abbreviations: MAPE, mean absolute percentage error; Pdev, average deviation between real and predicted peak load; PUc, public charging; REc, residential charging; SCc, shopping center charging; Tdev, average time deviation of predicted peak load.

seems to be more rational. Whereas the sine/cosine encoding offers a viable encoding variant, 16 input ts for the 4-ts and 96 input ts for the 96-ts prediction seem to be appropriate.

Lastly, the stateful mode does not provide benefits compared to the stateless mode under the given conditions. However, when multi-year data becomes available in the future, the stateful mode may reveal its strengths, as seasonal relationships can be identified by the LSTM. Consequently, with the availability of data over several years, a further comparison of the different modes is advisable.

6.2 | Hyperparameter tuning results

Three relevant conclusions can be drawn. First, the significant variation between the minimum and maximum loss during tuning reveals that the different hp combinations exert a major impact on the generalisability of the LSTM and, ultimately, on its predictive power. Thus, the importance of hp tuning for selecting the most suitable LSTM model is outlined.

Moreover, the varying combinations of hp exert a fundamentally different impact on the validation loss for each charging site category and forecast horizon. Thus, it is essential to perform the hp tuning separately for each use case.

Finally, random search shows only minor or no enhancements compared with the initial training. This might be due to the limited amount of evaluations and search space. A higher number of executions, the inclusion of other hp in the search space, or the use of a more subtle method, like the bayesian hp tuning, could lead to further improvements.

6.3 | Forecasting results

The results of the forecast are discussed below, followed by implications for the practical application of the forecast.

6.3.1 | Differences between the different categories of charging sites

The comparison of the forecasting results of the different charging sites revealed that the evaluation is highly dependent on the metric. However, SCc almost solely yields the poorest results for the 4-ts prediction, indicating that the fluctuating load hinders the predictive power of the LSTM.

Further, evident discrepancies involve the varying results for weekdays, weekends, and public holidays. While the lowest NMAE1 is obtained for SCc and REc on public holidays, the public holiday load predictions for PUc and WOc are significantly less accurate than the predictions on weekdays and weekends. This finding is attributable to the distinct load profile characteristics rather than to the different forecasting ability of the LSTM. While the load for PUc and WOc shifts radically on public holidays, the change for REc is less pronounced. With SCc, the load profile also varies

noticeably, but the entire load series shows strongly fluctuating load characteristics. Thus, the impact of public holidays remains small.

Similar findings can be seen while looking at the peak deviation and MAPE results. While for SCc and Rec, the MAPE difference between weekdays and public holidays amounts to only 3.75 and 5.09 percentage points, respectively, the figure for public and workplace charging increases by 97.96 and 906.63 percentage points, respectively. Once again, the discrepancies can be explained by the impact of public holidays on the load pattern. Analysing the weekday results, the significant discrepancies in the MAPE and time deviation, likewise, indicate that the characteristics of the load profile decisively impact the accuracy of the forecast. The highly irregular load profile for SCc results in the poorest accuracy for both, the predicted peak load level and the time of the peak load. WOc, on the other hand, exhibits the steadiest pattern concerning the time of the peak load, which is why the average time deviation is the lowest at 35 min.

6.3.2 | Overall findings

The superior performance of the 4-ts prediction compared to the 96-ts forecast is attributable to two factors. Due to the shorter forecast horizon, the preceding load values are shown to the LSTM more frequently and assist the LSTM in mapping the height and course of the load more precisely. Moreover, the poor prediction results for the 96-ts forecast might be traced back to the strongly altered load profile of the test data, caused by the increase of charging load throughout the year and the winter holiday period. Therefore, the training and validation data might have not been fully representative of the test data, limiting the LSTM's predictive power.

6.3.3 | Implications for aggregators and network operators

There are two courses of action on how aggregators and network operators can implement the LSTM to achieve optimised forecasts for their respective use case. First, it is beneficial to limit the load forecast to weekdays. Weekends and public holidays are accompanied by changes in the user behaviour that are difficult to predict and exhibit a much lower aggregation potential and risk of bottlenecks due to the reduced charging load. By focussing the forecast on weekdays, a higher prediction accuracy can be obtained.

Second, the forecast should only be carried out for the periods of a day with the highest charging load, where the aggregation potential is the highest, and bottlenecks in the distribution network are most likely. By shortening the forecast period, the accuracy is improved as shown in this work. The forecast will be further enhanced by focussing the LSTM on a specific period of the day.

7 | CONCLUSIONS AND FUTURE WORK

This study proposes a novel multivariate long short-term memory approach for multi-step EV charging load forecasting with two different prediction horizons. The performance of the forecasting approach is evaluated and compared between four different categories of charging sites. The results show that the distinct characteristics of the different charging sites influence the predictability of the charging load, but that the evaluation is highly dependent on the chosen metric. It is also demonstrated that with an increasing forecast horizon, the accuracy diminishes as well. Reducing the forecast horizon from 96 to four time steps, the MAE is more than halved in most cases, and amounts to 4.35 kW for shopping centre, 1.53 kW for residential, 2.7 kW for public, and 1.85 kW for workplace charging. In general, the forecasting accuracy tends to be the best at residential and workplace charging sites. The weakest accuracy is found at shopping centres.

The findings of this work benefit mainly two stakeholders. Aggregators have great interest in a reliable load forecast to sell EV flexibilities to the energy market. Network operators, in contrast, are keen on forecasting the charging load to identify possible bottlenecks in the distribution network. To apply the proposed LSTM model most effectively for both use cases, future studies should consider two aspects.

To address the outlined weaknesses of the 96-time step forecast, the LSTM will be trained on data collected over a longer time period to ensure that the training and validation data is truly representative of the test data. With data available over several years, the LSTM will be extended by further indicators, such as the month, the stateful mode will be investigated again, and time series cross validation will be performed to increase the robustness of the model and overcome early overfitting.

Furthermore, a future research objective involves the development of a customised LSTM, tailored to each charging site. The forecast will be targeted to weekdays, since user behaviour on weekends and holidays is often difficult to predict and, in many cases, as with workplace charging, the aggregation potential is not sufficient enough. Additionally, the prediction will be narrowed down to the most suitable time periods of the day at each charging site when a sufficient volume of EVs is available for pooling the EV flexibility and bottlenecks caused by simultaneous charging, such as 4:00 PM to 8:00 PM for residential charging, 8:00 AM to 12:00 PM for public charging, or 6:00 AM to 10:00 AM for workplace charging.

ORCID
Tim Unterluggauer 🔟 https://orcid.org/0000-0002-6461-7291
Kalle Rauma 🔟 https://orcid.org/0000-0002-5553-8751

REFERENCES

1. Greenhouse Gas Emissions From Transport in Europe. https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12. (2020). Accessed 10 Dec 2020
2. Diana, N., Sivakumar, A., Polak, J.W.: Modelling electric vehicles use: a survey on the methods. Renew. Sust. Energ. Rev. 68, 447–460 (2017)
3. Wu, X., et al.: Stochastic optimal energy management of smart home with PEV energy storage. IEEE Trans. Smart Grid. 9(3), 2065–2075 (2018)
4. Toquica, D., De Oliveira-De Jesus, P.M., Cadena, A.I.: Power market equilibrium considering an EV storage aggregator exposed to marginal prices - a bilevel optimization approach. J. Energy Storage. 28, 101267 (2020)
5. Wang, D., et al.: Integrated energy exchange scheduling for multi-microgrid system with electric vehicles. IEEE Trans. Smart Grid. 7(4), 1762–1774 (2016)
6. Xia, S., et al.: An autonomous real-time charging strategy for plug-in electric vehicles to regulate frequency of distribution system with fluctuating wind generation. IEEE Trans. Sustain. Energy. 9(2), 511–524 (2018)
7. Appino, R.R., et al.: Reliable dispatch of renewable generation via charging of time-varying PEV populations. IEEE Trans. Power Syst. 34(2), 1558–1568 (2019)
8. Janjic, A., et al.: Commercial electric vehicle fleet scheduling for secondary frequency control. Elec. Power Syst. Res. 147, 31–41 (2017)
9. Knezovic, K., et al.: Enhancing the role of electric vehicles in the power grid: field validation of multiple ancillary services. IEEE Trans. Transp. Electrific. 3(1), 201–209 (2017)
10. Zeraati, M., Hamedani Golshan, M.E., Guerrero, J.M.: A consensus-based cooperative control of PEV battery and PV active power curtailment for voltage regulation in distribution networks. IEEE Trans. Smart Grid. 10(1), 670–680 (2019)
11. Asrari, A., et al.: A market framework for decentralised congestion management in smart distribution grids considering collaboration among electric vehicle aggregators. IEEE Trans. Smart Grid. 11(2), 1147–1158 (2020)
12. Liang, H., et al.: Dynamic economic/emission dispatch including PEVs for peak shaving and valley filling. IEEE Trans. Ind. Electron. 66(4), 2880–2890 (2019)
13. Duan, Z., Gutierrez, B., Wang, L.: Forecasting plug-in electric vehicle sales and the diurnal recharging load curve. IEEE Trans. Smart Grid. 5(1), 527–535 (2014)
14. Zhang, H., et al.: Evaluation of achievable vehicle-to-grid capacity using aggregate PEV model. IEEE Trans. Power Syst. 32(1), 784–794 (2017)
15. Pertl, M., et al.: An equivalent time-variant storage model to harness EV flexibility: forecast and aggregation. IEEE Trans. Ind. Inf. 15(4), 1899–1910 (2019)
16. Amini, M.H., Kargarian, A., Karabasoglu, O.: ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation. Elec. Power Syst. Res. 140, 378–390 (2016)
17. Korolko, N., Sahinoglu, Z., Nikovski, D.: Modelling and forecasting self-similar power load due to EV fast chargers. IEEE Trans. Smart Grid. 7(3), 1620–1629 (2016)
18. Louie, H.M.: Time-series modelling of aggregated electric vehicle charging station load. Elec. Power Compon. Syst. 45(14), 1498–1511 (2017)
19. Buzna, L., et al.: Electric vehicle load forecasting: a comparison between time series and machine learning approaches. In: 2019 1st International

Conference on Energy Transition in the Mediterranean Area (SyNERGY MED), pp. 1–5. Cagliary (2019). May

20. Majidpour, M., et al.: Forecasting the EV charging load based on customer profile or station measurement? Appl. Energy. 163, 134–141 (2016)
21. Lu, Y., et al.: The application of improved random forest algorithm on the prediction of electric vehicle charging load. Energies. 11(11), 1–16 (2018)
22. Jahangir, H., et al.: Charging demand of plug-in electric vehicles: forecasting travel behaviour based on a novel rough artificial neural network approach. J. Clean. Prod. 229, 1029–1044 (2019)
23. Dabbaghjamanesh, M., Moeini, A., Kavousi-Fard, A.: Reinforcement learning-based load forecasting of electric vehicle charging station using Q-learning technique. IEEE Trans. Ind. Inf. 17(6), 4229–4237 (2021)
24. Zhu, J., et al.: Short-term load forecasting for electric vehicle charging stations based on deep learning approaches. Appl. Sci. 9(9), 1–12 (2019)
25. Zhu, J., et al.: Electric vehicle charging load forecasting: a comparative study of deep learning approaches. Energies. 12(14) (2019)
26. Zhu, J., et al.: A novel LSTM based deep learning approach for multi-time scale electric vehicles charging load prediction, In: IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), pp. 3531–3536. Chengdu (2019)
27. Gao, Q., et al.: Charging load forecasting of electric vehicle based on Monte Carlo and deep Learning. IEEE Sustainable Power and Energy Conference (iSPEC), pp. 1309–1314, Beijing (2019)
28. Jahangir, H., et al.: Plug-in electric vehicle behaviour modelling in energy market: a novel deep learning-based approach with clustering technique. IEEE Trans. Smart Grid. 11(6), 4738–4748 (2020)
29. Xue, M., et al.: Research on load forecasting of charging station based on XGBoost and LSTM model. J. Phys. Conf. 1757(1), 1–10 (2021)
30. Na, Z., et al.: Short-term load forecasting algorithm based on LSTM-DBN considering the flexibility of electric vehicle. IOP Conf. Ser. Earth Environ. Sci. 546(4), 1–8 (2020)
31. Zhang, X.: Short-term load forecasting for electric bus charging stations based on fuzzy clustering and least squares support vector machine optimised by wolf pack algorithm. Energies. 11(6), 1–18 (2018)
32. Li, Y., Huang, Y., Zhang, M.: Short-term load forecasting for electric vehicle charging station based on niche immunity lion algorithm and convolutional neural network. Energies. 11(5), 1–18 (2018)
33. Gerossier, A., Girard, R., Kariniotakis, G.: Modelling and forecasting electric vehicle consumption profiles. Energies. 12(7), 1–14 (2019)
34. Zhang, X., et al.: Deep- learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model. IEEE Trans. Cybernetics. 51(6), 3157–3170 (2020)
35. Bianchi, F.M., et al.: Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis. Springer, Berlin (2017)
36. Qing, X., Niu, Y.: Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy. 148, 461–468 (2018)
37. Kong, W., et al.: Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans. Smart Grid. 10(1), 841–851 (2019)
38. Potdar, K., Pardawala, S.T., Pai, D.C.: A comparative study of categorical variable encoding techniques for neural network classifiers. Ijca. 175(4), 7–9 (2017)
39. Hernandez, L., et al.: Short-term load forecasting for microgrids based on artificial neural networks. Energies. 6(3), 1385–1408 (2013)
40. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13(1), 281–305 (2012)