Advances in Science and Technology
Research Journal

Advances in Science and Technology Research Journal, 2026, 20(5), 223–237
https://doi.org/10.12913/22998624/216176
ISSN 2299-8624, License CC-BY 4.0

Received: 2025.12.02
Accepted: 2026.03.15
Published: 2026.04.01

High-resolution electricity load forecasting in a university campus using deep neural networks, Kalman filtering, and self-attention

Adam Krechowicz[1*], Mansueto Rossi[2], Federico Delfino[2], Artur Bartosik[3]

[1] Faculty of Electrical Engineering Automatic Control and Computer Science, Kielce University of Technology, Al.1000-lecia PP. 7, 25-314 Kielce, Poland
[2] Department of Naval, Electrical, Electronic and Telecommunications Engineering, University of Genoa, via alla Opera Pia 11a, Genova, 16145, Italy
[3] Faculty of Management and Computer Modelling, Kielce University of Technology, Al.1000-lecia PP. 7, Kielce, 25-314, Poland

ABSTRACT

Accurate and reliable forecasts of short-term demand for a microgrid and power generation forecasts from photovoltaics are an important part of smart microgrid management, influencing technical, operational, and financial aspects of microgrid operation. Managing university campus microgrids is challenging due to their specificity. This work is a response to the need of university campuses for a reliable deep learning model supporting microgrid energy management that allows for accurate forecasting of both energy demand and photovoltaic electricity generation. In this study, sixteen forecasting models based on deep learning, utilizing various layer combinations and architectures, were developed and evaluated. The combination of a deep neural network, Kalman Filtering, and a Self-Attention model was found as a robust and reliable solution for managing energy within a university campus microgrid. This model enables precise photovoltaic power generation and load forecasting, as indicated by the high determination coefficient values (R2=0.987 for load forecasting, and R2=0.989 for PV power generation forecasting) and the low mean absolute error MAE, root mean square error RMSE, and their standard deviations (4.598 and 5.749 respectively for load forecasting and 0.953, 1.833 for PV generation). Furthermore, it effectively manages conditions that differ from those encountered during training, showcasing its high generalization capacity (R2=0.982 MAE=5.006, RMSE=1.257 for load forecasting and R2=0.992 MAE=0.634, RMSE=1.395 for PV generation). Additionally, this model achieved the lowest standard deviations. Its practical application was validated using actual energy performance data from Savona Campus microgrid in Italy over all seasons, rather than relying on simulation or small-scale research, which helped minimize errors arising from training set imperfections.

Keywords: load forecasting, forecasting electricity generation from photovoltaics, university campus, deep neural network.

INTRODUCTION

Microgrids enable supplying stable electrical power from various intermittent renewable sources. Currently, interest in microgrids is growing due to the increase in energy cost related to international conflicts and to the widespread global trend toward the development of sustainable energy-supply solutions capable of mitigating climate change. Due to the intermittent nature of renewable energy sources, smart microgrids have to face many challenges related to power quality, reliability, and balance between supply and demand [1]. The prediction of short-term demand for a microgrid is an integral part of smart microgrid management. It is a complex task due to its non-smooth and non-linear behavior of the load time series [2]. Precise load forecasting boosts the management of renewable and traditional resources in microgrids, aids in power generation planning, and optimizes economic outcomes when trading energy with the commercial grid [1]. The prediction of energy generation from renewable energy sources (RES), as well as the prediction of load in the microgrid, is vital to ensure optimal use of RES, as well as the perpetual and effective operation of the power grid [3].

The aim of this paper is to propose an efficient model supporting energy management in a university campus microgrid, which will allow accurate prediction of photovoltaic power generation and load forecasting.

The contribution to the body of knowledge of this work includes proposing a reliable deep learning model that:

- is capable of both predicting university campus load and electricity generation from a microgrid with a very high level of compliance of predicted load and electricity generation with actual test results,
- is high resolution, more precise thanks to using 1-minute interval data, which allowed for the enlargement of the data set compared to models based on 1-hour or 15-minute intervals,
- has been tested/validated against real-world university campus microgrid energy performance data, demonstrating its practical application as an effective tool for forecasting both load and microgrid energy generation with good generalization ability (ability to cope with conditions different from those reported during training),
- is trained, tested, and validated on a data set obtained from university campus microgrid in real-world conditions covering all seasons of the year, and not on the simulation results or research on a small scale, which allowed us to reduce the error coming from the imperfections of the training set,
- is dedicated for university campuses, taking into account their specificity: they can be seen as a specific type of urban district, with usually well-defined boundaries, a unified management, and a mix of offices and residential buildings (student residences), as well as research infrastructures; these peculiarities make them particularly suitable for installing and testing new technologies (microgrids, all-electric smart buildings, etc.), as well as for testing new control and management strategies.

RELATED WORKS

Machine learning models offer promising solutions in many areas, including engineering [4], medicine [5] and art [6]. In [7] the subject of problems with inaccurate prediction of electricity generation from photovoltaic systems was discussed. It was stressed that inaccurate predictions can lead to serious technical, operational, and financial risks, which seriously affect owners and grid operators. It was concluded that precise machine learning (ML) models that take into account a wide time horizon covering all seasons can be the solution to avoid problems with inaccurate predictions. In [8] the main challenges associated with the forecast of electricity generation from photovoltaics were identified. The following problems were discussed: lack of cross-validation, showing the forecast performance for a too short time horizon without considering 4 seasons, enhancement of results by evaluating PV performance also during the night when energy generation is always 0, problems with false readings. Singh et al. [9] used the Support Vector Regression model based on energy generation data, weather patterns, and dynamic grid conditions to predict power generation in a microgrid. The developed model outperformed linear regression models with a mean squared error (MSE) of 2.002 for PV power prediction and 3.059 for wind power prediction. Yuan et al. Yuan et al. [10] proposed a machine learning reinforced model that can forecast the output generation of wind and solar units in the microgrid and a model that uses modified whale optimization that can solve different scheduling plans. The proposed models support the optimization of total operating costs, improve the voltage profile, and reduce power losses.

The forecast of short-term demand for a microgrid is an important part of smart microgrid management. In [11] the prediction of power load consumption with one-hour interval is presented one day ahead based on wavelet transform (WT), simulated annealing (SA) and feedforward artificial neural network (FFANN) on the microgrid in Beijing, China. It achieved a mean absolute percentage error (MAPE) of 2.95%, a root mean square error (RMSE) of 22.15 kW and a forecast skill of 62.73. Indira et al. [12] proposed an adaptive barnacle-mating optimizer and an artificial neural network model for short-term load prediction in one-hour intervals. It outperformed easier basic models such as the regression tree, support vector machine, artificial neural network, and particle swarm optimization-based artificial neural network mean absolute percentage error MAPE=0.9736% and symmetric mean absolute percentage error SMAPE= 0.4968%. Da et al. [13] developed an adaptive convolution neural network long short-term memory for hourly electrical load forecasting taking into consideration parameters from smart solar microgrid and Building Management System. The performance of the model was compared to long short-term memory (LSTM), gated recurrent unit (GRU), convolutional neural network (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM). K-fold cross-validation was utilized to verify deep learning algorithms such as CNN, GRU, LSTM, and Bi-LSTM methods. The findings indicated that the hybrid CNN-LSTM showed certain improvements, which are 20.57%, 29.63%, and 19.06% in MSE, mean absolute error (MAE), and MAPE, and 21.24%, 22.02%, and 3.82% in validation MSE, MAE, and MAPE, respectively. In [14] the authors developed a bidirectional gated recurrent convolution neural network unit to forecast the three-phase load power of the building within the intelligent solar microgrid in the short term. Lan et al. [15] used support vector machine for the prediction and modeling of the charging demand of hybrid electric vehicles. According to the prediction results for the total charging demand of the HEVs. The mean absolute percentage error for prediction was 0.978. In addition, thanks to the use of the proposed approach, a 2.5% decrease in the total operating cost of the system was observed.

Wen et al. [16] developed a deep learning model based on LSTM to forecast combined power load and photovoltaic power output within a community microgrid with hourly resolution. The model outperformed multilayer perception and support vector machine models. Optimization of load dispatch of the grid-connected community microgrid was achieved thanks to the application of particle swarm optimization. Trivedi et al. [17] presented short-term forecasting of PV generation and load demand with half-hourly resolution. The study used random forests models for advanced feature selection and intense hyperparamters tuning for deep learning models. In [18] the authors compared the performance of k-Nearest Neighbors (k-NN) and random forest (RF) developed for micro-grid decision-making and optimization task. RF was more precise for the most important micro-grid components (the fuel cell relay, battery relay, super-capacitor relay, and grid system relay) with F1 score 82%. Ali et al. [19] proposed a hierarchical control approach of an inverter-based microgrid applying cloud-based Internet of Things infrastructure and artificial neural network for island detection. The model was trained with simulated island scenarios in Matlab. Ji et al. [20] developed a real-time scheduling model for a microgrid, taking into account the uncertainty of the load demand, renewable energy and the price of electricity. A Markov Decision Process (MDP) is used for microgrid energy management modelling to minimize the daily operating cost. MDP is solved using a deep reinforcement learning model. In the proposed model, the states of the microgrid are taken as inputs and the real-time generation schedules as outputs. The effectiveness of the proposed solution was checked on the real data set from California. Ashok et al. [21] proposed a deep learning model for the regulation and optimization of power of microgrids, which uses distributed energy resources such as generators and batteries. Controllers enabling monitoring of voltages, phases, and frequencies on both sides of the fixed switch employed several control strategies to stabilize the system. The proposed approach supports effective control mechanisms and a sustainable energy economy. In [22] the authors developed a green energy management system using the XGBoost algorithm and reinforcement learning to predict short-term energy consumption and solar generation. The proposed model has a root mean square error (RMSE) of 14.72, a mean absolute error (MAE) of 12.00, and a mean absolute percentage error (MAPE) of 2.18%.

The problem of managing microgrids on university campuses has not been widely explored in the literature to date. Muqeet et al. [23] reviewed the energy management system of campus microgrids, finding eight challenges facing campus microgrids. They are connecting with increasing green energy sources, lowering operating and running costs, ensuring high reliability, minimizing the use of utility electricity, improving the energy management system, ensuring that the prices of the unit of electricity are efficient, and developing an economic plan to improve the economic benefit of campus microgrid systems. Kim [24] introduced an energy management framework that accounts for both power generation and consumption. This model utilizes IoT sensors to gather data, which includes information from PV generators, among other sources, solar radiation, temperature, dust, etc. The performance of the model was checked using a real data set from the university campus. It was found that due to the application of the model, the electricity cost could be reduced up to 2% and the peak power can be reduced up to 3% compared to the case without considering the uncertainties in a campus microgrid.

A review of the literature revealed that the problem of microgrid management on university campuses using machine learning has not been widely explored so far in the literature. In the analyzed literature, it was stressed that accurate forecasts of both renewable energy production and demand are crucial to the proper operation of microgrids. Although machine learning holds promise in offering new solutions for load forecasting and renewable energy production forecasting, the results obtained in the analyzed works still leave room for improvement. This prompted the authors of this article to address the need for university campuses for a reliable deep learning model that allows accurate forecasting of both energy demand and photovoltaic energy generation.

METHODOLOGY

Data gathering and profiling

The data used in this paper was collected on the Savona Campus (Italy), which hosts the Smart Polygeneration Microgrid (SPM) and other research facilities [25,26]. The SPM is a 3-phase low voltage (400 V line-to-line) distribution system, coupled with a thermal network, composed of electrical/thermal loads and generation units. The SPM electrical grid topology is a ring with one main switchboard (QEG) and five other switchboards (Q01, Q02, Q03, Q04) to which the power plants and the loads are connected; in 2017 a fifth switchboard has been added to feed the smart energy building (SEB), and connected between Q02 and Q03. A dedicated MV/LV transformer links the SPM to the MV busbar of the Savona Campus MV/LV main substation and to the 15 kV utility's distribution network via the Campus point of common coupling (PCC), to which all the other Campus buildings and facilities are also connected.

The SPM hosts:
- two cogeneration microturbines (Capstone C65) fed by natural gas;
- a thermal station equipped with gas boilers fed by natural gas;
- two absorption chillers;
- three photovoltaic (PV) fields: two (PV1, 80 kWp and PV2, 15 kWp) are located on the roof of the Delfino Building, while the third one (PV3, 21 kWp) is located on the roof of the SEB;
- an electrical storage (ES) system (sodium/nickel chloride);
- electric vehicle (EV) charging stations.

The SPM is controlled and managed by field data acquisition and local automation devices and a SCADA (Supervisory Control And Data Acquisition) system. Generation units and loads are constantly monitored and collected data are used to assess the technical performance of power plants. The data collected are stored in the SCADA database. The communication system consists of switches installed in the SPM control room and in each switchboard of the SPM. These switches, compliant with the IEC 61850 protocol, are connected via a double fiber optic ring, the communication backbone of the microgrid. The SPM is also equipped with smart power meters to monitor power fluxes, energies, voltages, and currents in different nodes. Some of the installed devices directly communicate via the IEC 61850 protocol, while others use the simpler Modbus protocol and rely on remote terminal units (RTU), installed on the switchboards, to connect to the SCADA.

In this paper, data sets gathered for Campus load, PV1 generation, and PV3 generations were used. The time spans and the number of data samples are presented in Table 1.

Data preprocessing

Several preprocessing techniques were applied to prepare the data for deep learning. First, seasonal features were created that represent the time of day, week, and year. Those features were created on the basis of the Equations 1, 2, and 3.

In the next step, the entire data sets were checked for invalid values. In case of the dataset concerning campus load some negative values were present due to unknown problem during the data aggregation process. The missing values

Table 1. Time spans of used datasets

Dataset	Start date	End date	# Samples
Campus load	2018-01-07	2023-09-07	2.710.547
PV1 generation	2017-02-23	2023-09-07	3.067.407
PV3 generation	2019-10-11	2023-09-07	1.933.264

$$\left(day_{sin}(t), day_{cos}(t)\right) = \left(sin\left(t \times \frac{2\pi}{24 \times 60 \times 60}\right), cos\left(t \times \frac{2\pi}{24 \times 60 \times 60}\right)\right) \quad (1)$$

$$\left(week_{sin}(t), week_{cos}(t)\right) = (sin\left(t \times \frac{2\pi}{7 \times 24 \times 60 \times 60}\right), cos\left(t \times \frac{2\pi}{7 \times 24 \times 60 \times 60}\right)) \quad (2)$$

$$\left(year_{sin}(t), year_{cos}(t)\right) = (sin\left(t \times \frac{2\pi}{365.2425 \times 24 \times 60 \times 60}\right) cos\left(t \times \frac{2\pi}{365.2425 \times 24 \times 60 \times 60}\right)) \quad (3)$$

were present in the 89 samples, which consists of 0.3% of the total samples. Those cases were replaced by zeros. In the dataset containing campus load there are visible random fluctuations that are typically caused by human behavior which can not be exactly predicted. Those fluctuations can seriously affect the forecasting. Because of that, we used Kalman filtering to remove them while preserving the overall profile of the load. We set the parameters experimentally in such a way that it allows us to eliminate random noises while in the same time not interfere with the natural trend in the data. We used the following parameters: Q = matrix for the Discrete Constant White Noise Model, dimensions = 2, time step = 0.05, variance of the noise = 0.1, P = I·1000, R=5.

To ensure that all features will be treated in the same way by machine learning models, their values have been standardized. We used z-score method to perform this task.

The weather data we used have a resolution of 30 minutes. To integrate them with the campus load data, we performed a linear interpolation.

Figure 1 presents the division of the whole data set into different parts that were used during cross-validation and testing. We used 20% of the dataset for testing and the rest for training models. The cross-validation was performed on the training part of the whole dataset. For the purpose of the machine learning methods, data were organized into batches that are simultaneously processed by ML model. The size of the batch was 32.

Developing deep neural network

Figure 2 presents the final architecture of the neural network used to predict campus load. The proposed architecture was chosen experimentally. Since we operated on relatively large dataset in this paper, we focused on deep learning methods which can be hardware accelerated. The final architecture consists of the following types of layers:

- dense layers – layers that are fully connected;
- 1 dimensional convolutional layers – layers that use sliding filter to detect patterns the data series;
- long short term memory layers – recurrent layers that allow to remember previous state;
- self-attention layers – layers that detect important part of the data series;
- dropout layers – layers for rejecting part of the input for improving generalization.

To assess the quality of the architecture, a k–fold cross–validation with 5 folds was used. In addition to that, we also used a standard validation mechanism that allowed us to quickly evaluate the quality of the examined model. Different numbers and combinations of layers were analyzed. In particular, we examined the variants of the structure without each complex type of the layers to asses their role in the effectiveness of the model. We analyzed 16 forecasting models

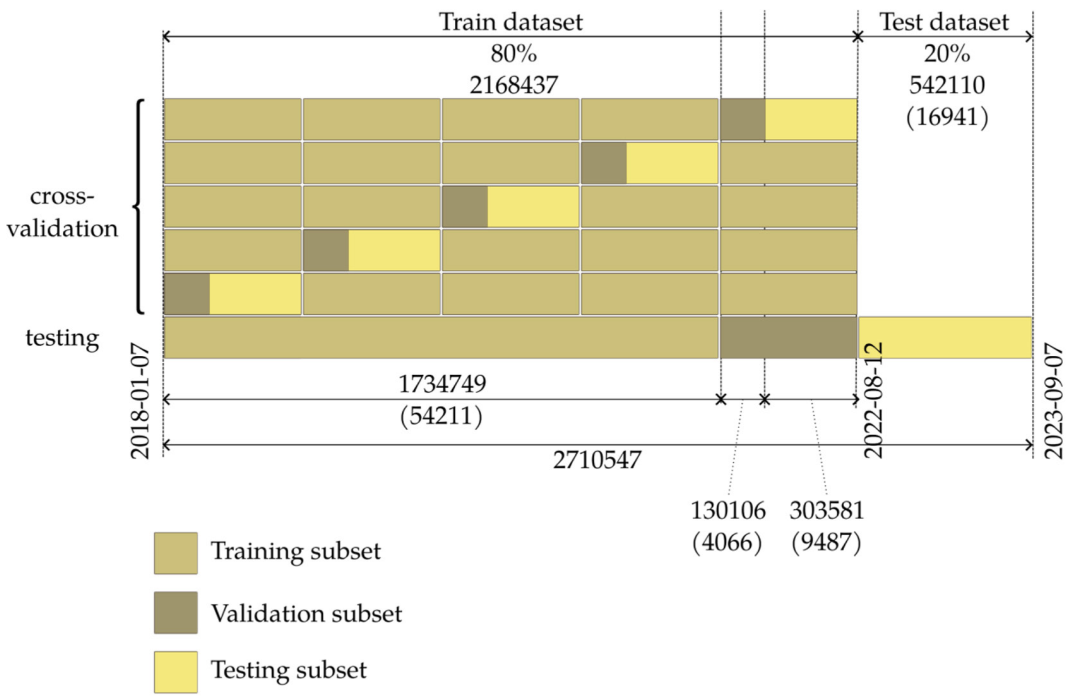


Figure 1. Division of the dataset into parts. Values are presented in fractions and number of the single data items. Values in parentheses represent the number of batches

which are summarized in Table 3. For analyzing all of the combinations, we used the campus load dataset. In addition, we examined the best models on PV generation datasets.

Evaluation

The developed models were evaluated using standard regression metrics: Mean squared error (Equation 4), coefficient of determination (Equation 5), mean average error (Equation 6) and root mean squared error (Equation 7). Additionally, to enable easy comparison with other models, we also used the normalized version of the root mean squared error Equation 8).

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{4}$$

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\acute{y})^2} \tag{5}$$

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i) \tag{6}$$

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2} \tag{7}$$

$$\text{NRMSE}=\frac{\text{RMSE}}{y_{max}-y_{min}} \tag{8}$$

RESULTS

The visualization of the preprocessing techniques is shown in Figure 3. The occurrence of the missing values that were replaced by zeros is presented in sub Figure 3a. The example of timestamp generated signals is shown in sub Figure 3b. The example of application of Kalman filtering on campus load signal is presented in sub Figure 3c. The interpolation of the weather data is shown in sub Figure 3d.

The basic description of the features in the datasets is described in Table 2 while the visualization of the distribution of these features is shown in Figure 4a. Analysis of Figure 4a reveals the significant difference in the distribution of different features. This is the main reason behind using feature standardization during the preprocessing. The correlation matrix for data features is presented in Figure 4b. As it can be seen there is no strong correlation of campus load with any of input features. The most correlated features are: global radiation (0.3), day cos (-0.49) and week cos (0.26).

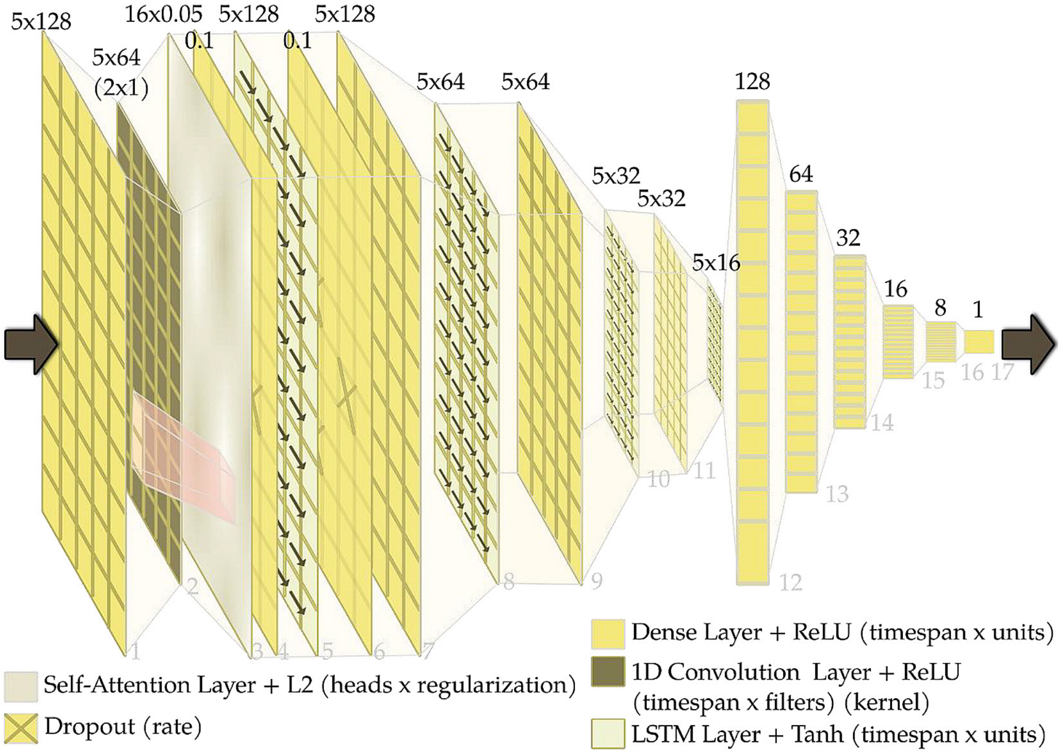


Figure 2. The proposed neural network architecture. Network consists of three types of layers: Dense, Recurrent and Convolutional. The first part of the network works on temporal data with time span of 5 min. Temporal data are flatten in the last recurrent layer with do not return each sequences. Second part of the network consists of only dense layers (time span = 1). The Self-Attention layer consists of Multi Head Attention with 16 heads, addition layer and normalization layer

The results of the model prediction load are presented in Table 3. The table contains the average and standard deviation of the cross-validation, as well as results from the test data set. The best results obtained were achieved for model number 3 (with Kalman filtering and LSTM), for model number 5 (with Kalman filtering, Self-Attention, LSTM and CNN) and for model number 2 (with Kalman filtering, LSTM and CNN). In cases of models without Kalman filtering the best achieved results were obtained for model number 13 (with Self-Attention, LSTM and CNN) and for model number 11 (only LSTM).

We performed the paired t-test to try to assess the statistical significance between models. While assuming p=0.05 we failed to reject the hypothesis of different distributions between models 5, 3 and 2. However, differences in the standard deviations indicate that these models may perform differently with the data gathered in the future.

The learning histories of loss value (mean squared error) are presented in Figure 5 for model number 3. Similar histories were obtained for other models. Figure 5a shows the learning history for the cross–validation and Figure 5b presents history for testing. In all cases, the learning process was performed in the appropriate manner without signs of overfitting. In the case of Figure 5a the curves represent the mean values from all the epochs, while brighter areas represent the standard deviation.

The example results obtained for the month are shown in Figure 6 for the campus load and Figure 7 for photovoltaic generation. Figures 6 and 7a present the results produced with Model 3. Figures 6b and 7b present the results produced with Model 5. The interesting fact can be seen in Figure 6 which shows how the models correctly predicted the decrease in campus load in the period of the last week of December caused by the Christmas holidays. The example results obtained for a single day by Model 3 are presented in Figure 8. Figure 8a presents the campus load prediction while Figure 8b presents the PV generation.

When the Kalman filter was not used, the best results were achieved for Model 13 (with Kalman Self-attention, LSTM and CNN). It should also be noted that when data from photovoltaic electricity

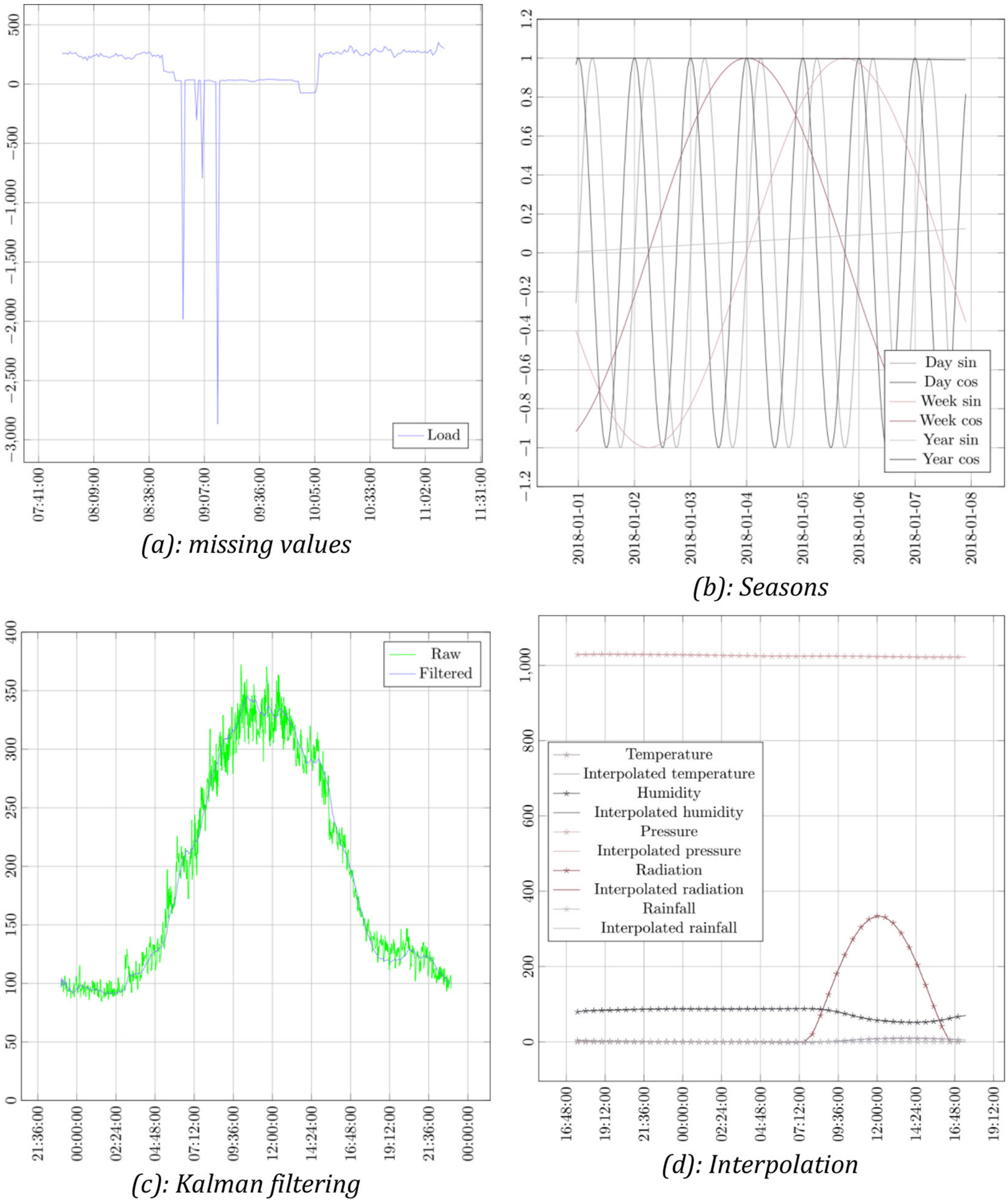
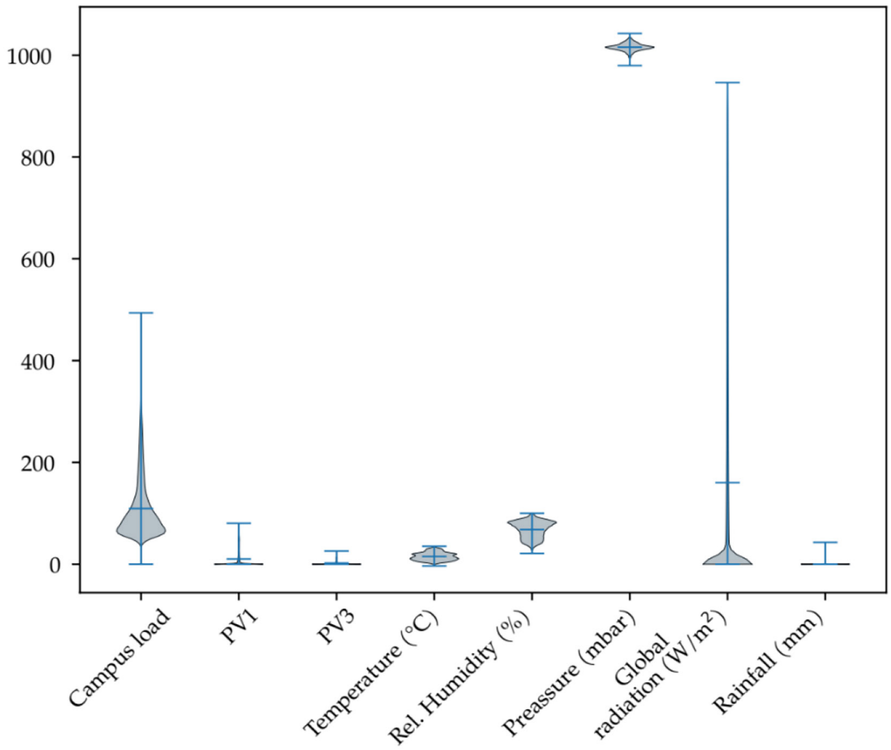


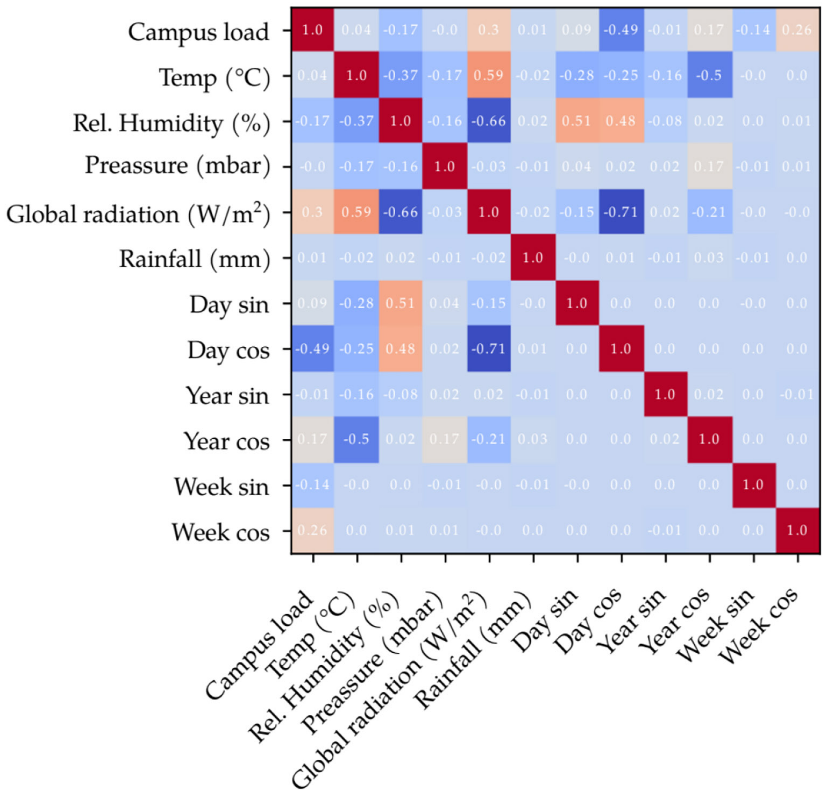
Figure 3. Preprocessing tasks

Table 2. Description of the analyzed features

Feature	mean	std	min	max
Campus load	109.469	57.988	0.000	493.856
PV1	10.327	16.732	0.000	80.280
PV3	2.220	3.970	0.000	25.690
Temperature (C)	15.121	7.628	-3.600	35.100
Rel. humidity (%)	67.989	16.080	21.200	100.000
Pressure (mbar)	1015.942	7.367	979.683	1042.669
Global radiation (W/m^2)	160.136	239.968	0.000	946.400
Rainfall (mm)	0.047	0.638	0.000	42.950

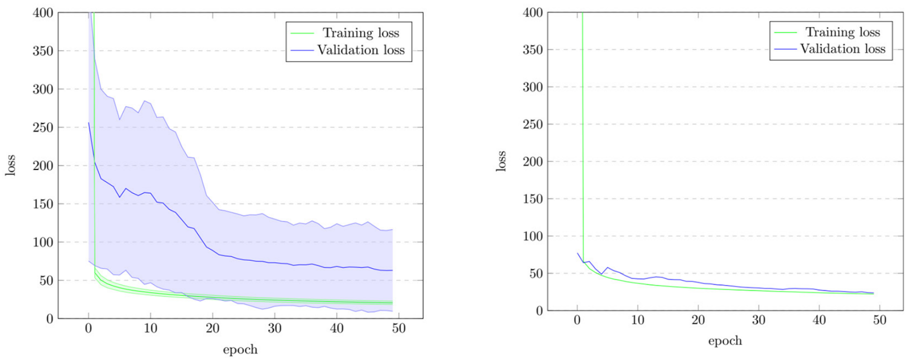


(a): Distribution of analyzed features



(b): Correlation matrix of analyzed features

Figure 4. Data presentation



(a): cross–validation history *(b): testing history*

Figure 5. Loss curve for training the model number 3 with Kalman filter and LSTM layers

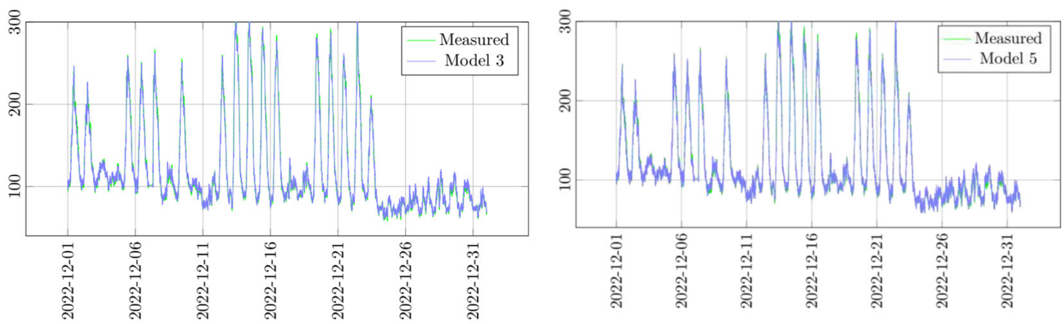
Table 3. The performance results of different configuration of the network for forecasting campus load for 15 minutes ahead based into consideration the window span of 5 minutes. The best achieved result is marked with bold face. The best achieved result without Kalman Filtering is marked with underline

Network configuration					Cross-validation results										Testing results				
NO	Kal-man	Atten-tion	LSTM	CNN	mean R2	std R2	mean MSE	std MSE	mean RMSE	std RMSE	mean nRMSE	std nRMSE	mean MAE	std MAE	R2	MSE	RMSE	nRMSE	MAE
Campus load prediction																			
1	+	X	X	X	0.933	0.011	198.885	135.852	13.443	4.766	0.031	0.011	10.456	3.828	0.944	148.336	12.179	0.028	9.510
2	+	X	+	+	0.984	0.008	**40.519**	18.916	6.241	1.399	0.014	0.003	4.630	**0.705**	0.987	33.960	5.827	0.013	4.209
3	+	X	+	X	**0.986**	**0.003**	42.628	28.533	**6.235**	2.167	**0.014**	0.005	**4.335**	1.218	**0.989**	**27.834**	**5.276**	**0.012**	**3.881**
4	+	X	X	+	0.942	0.015	180.976	140.358	12.632	5.173	0.029	0.012	9.659	4.155	0.950	130.241	11.412	0.026	9.048
5	+	+	+	+	0.982	0.011	42.038	**17.371**	6.385	**1.257**	0.015	**0.003**	5.006	0.754	0.987	33.050	5.749	0.013	4.598
6	+	+	X	X	0.858	0.073	453.920	415.792	19.147	10.447	0.044	0.024	14.483	8.503	0.862	361.449	19.012	0.043	13.249
7	+	+	+	X	0.953	0.011	134.392	77.594	11.184	3.409	0.025	0.008	7.815	2.476	0.838	424.881	20.613	0.047	18.388
8	+	+	X	+	0.946	0.022	174.715	150.518	12.215	5.648	0.028	0.013	9.736	5.363	0.945	143.319	11.972	0.027	9.326
9	X	X	X	X	0.854	0.014	418.814	262.793	19.675	6.297	0.040	0.013	13.597	4.875	0.864	355.796	18.863	0.038	13.098
10	X	X	+	+	0.862	0.065	329.248	120.117	17.927	3.137	0.036	0.006	13.489	1.554	0.873	331.974	18.220	0.037	13.525
11	X	X	+	X	0.867	0.034	337.363	126.464	18.113	3.409	0.037	0.007	13.365	1.487	0.879	315.610	17.765	0.036	13.288
12	X	X	X	+	0.215	1.437	1517.762	2392.807	31.455	25.699	0.064	0.052	23.803	22.822	0.878	318.115	17.836	0.036	12.118
13	X	+	+	+	0.866	0.051	322.619	98.399	17.805	2.643	0.036	0.005	13.653	1.209	0.868	345.267	18.581	0.038	13.645
14	X	+	X	X	0.742	0.088	817.347	681.955	26.523	11.930	0.054	0.024	18.566	9.203	0.798	528.143	22.981	0.047	17.868
15	X	+	+	X	0.140	1.593	1224.105	1639.845	30.156	19.834	0.061	0.040	23.236	17.987	0.773	592.161	24.334	0.049	20.153
16	X	+	X	+	0.828	0.023	501.293	325.728	21.423	7.277	0.043	0.015	15.147	5.051	0.843	410.444	20.259	0.041	14.107
PV1 installation generation prediction																			
2	+	X	+	+	0.993	0.001	1.910	0.622	1.368	0.223	0.017	0.003	0.603	0.142	0.990	2.959	1.720	0.021	0.838
3	+	X	+	X	**0.993**	0.001	**1.808**	0.513	**1.333**	0.196	**0.017**	0.002	**0.570**	0.123	**0.991**	**2.716**	**1.648**	**0.021**	**0.825**
5	+	+	+	+	0.992	**0.0003**	1.969	**0.462**	1.395	**0.163**	0.017	**0.002**	0.634	**0.098**	0.989	3.361	1.833	0.023	0.953
10	X	X	+	+	0.919	0.014	19.871	7.654	4.391	0.858	0.055	0.011	1.763	0.526	0.923	22.481	4.741	0.059	2.053
12	X	X	+	X	0.920	0.011	19.437	6.770	4.355	0.766	0.054	0.010	1.767	0.469	0.921	23.223	4.819	0.060	2.086
13	X	+	+	+	0.919	0.009	19.604	6.229	4.383	0.701	0.055	0.009	1.804	0.430	0.914	25.120	5.012	0.062	2.267
PV3 installation generation prediction																			
2	+	X	+	+	0.987	**0.003**	**0.194**	**0.074**	0.435	**0.083**	0.017	**0.003**	0.209	**0.052**	0.978	0.407	0.638	0.025	0.348
3	+	X	+	X	**0.988**	0.005	0.197	0.161	**0.421**	0.157	**0.016**	0.006	**0.207**	0.107	**0.989**	0.199	**0.446**	**0.017**	**0.200**
5	+	+	+	+	0.973	0.016	0.393	0.281	0.598	0.211	0.023	0.008	0.291	0.126	0.962	**0.689**	0.830	0.032	0.389
10	X	X	+	+	0.886	0.013	1.644	0.765	1.257	0.285	0.049	0.011	0.528	0.171	0.897	1.827	1.352	0.053	0.579
12	X	X	+	X	0.883	0.034	1.821	1.346	1.288	0.450	0.050	0.018	0.568	0.289	0.899	1.777	1.333	0.052	0.569
13	X	+	+	+	0.896	0.024	1.630	1.081	1.220	0.419	0.048	0.016	0.529	0.260	0.839	2.855	1.690	0.066	0.704

generation were analyzed, in most cases the lowest standard deviations were obtained for the self-attention model, indicating that this model has a high ability to generalize across various data sets. For those reasons, models using self-attention should be seriously considered in real world applications.

Figure 9 presents the influence of the model hyper parameters on the results produced. Figure 9(a) presents the influence of the window shift hyper parameter which indicates how many minutes ahead the prediction was made with respect to produced errors. We chose campus load dataset to perform those experiments because this dataset tend to be more challenging as in most cases it was a little bit more difficult for the ML models to capture trends in forecasting load than in forecasting PV generation. Analysis of this figure shows a strong relationship between the number of minutes ahead and the error produced, which was expected behavior. Figure 9(b) presents the influence of the input width hyper parameter which indicates how many data from the past (in minutes) we took into consideration to generate the prediction. In this case, it can be seen that even a small input width can produce accurate results, and extending the value of this hyper parameter is not needed, and the value of 5 minutes is sufficient.

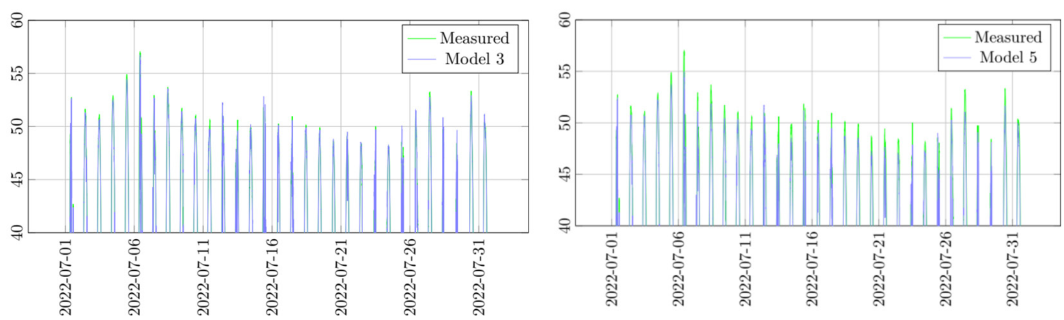
Table 4 presents the number of parameters used by each model that reflects the computational complexity of those models. The trainable parameters indicate the number of parameters that need to be adjusted in each epoch of learning. In addition, the optimizer parameters indicate how many parameters are necessary for performing optimization. The sum of those values are summarized as Total parameters. It is worth to note that the number of the parameters determines the computational complexity of the model and consequently the time of both training and inference by the model. From the analysis of Table 4 it can be seen that the self-attention mechanism has the greatest influence on the number of parameters. Additionally, due to the convolution nature of the CNN layer, it can reduce the number of parameters. Figure 10 presents the heatmap of Self-Attention scores that was generated during



(a): Load prediction for the Model 3 (R^2=0.992)

(b): Load prediction for the Model 5 (R^2=0.987)

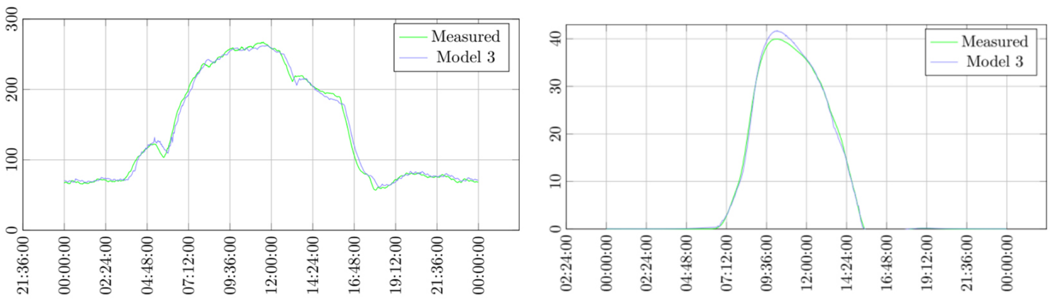
Figure 6. Prediction results of campus load for the example month (2022-12)



(a): PV generation predicted with Model 3 (R^2=0.995)

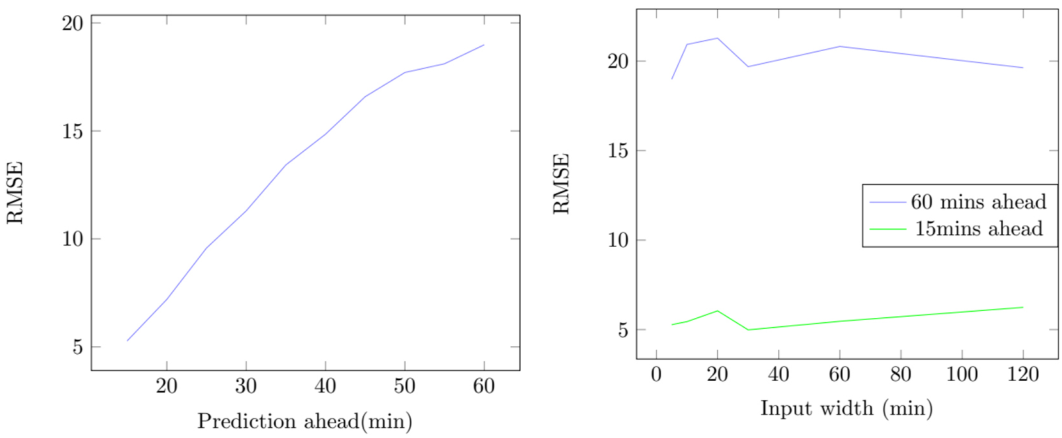
(b): PV generation predicted with Model 5 (R^2=0.993)

Figure 7. Prediction results of PV generation for example month (2022-07)



(a): Campus load prediction for 2022-12-21 (R^2=0.998) *(b): PV generation prediction for 2022-10-29 (R^2=0.999)*

Figure 8. Prediction results of example day with Model 3



(a): Errors in relation to the number of minutes prediction ahead *(b): Errors in relation to the window span width in minutes*

Figure 9. Analysis of the influence of the model parameters on produced errors in predicting campus load

Table 4. The number of the parameters for each models. The time of the learning and prediction was captured on a workstation with the AMD Ryzen 9 3900X 12-Core Processor with 64 GiB RAM and NVIDIA GeForce RTX 3090

No	Kalman	Attention	LSTM	CNN	Total params	Trainable params	Optimizer params	Learning time [min]	Prediction time [s]
1	+	X	X	X	180773	60257	120516	116.34	12.34
2	+	X	+	+	650789	216929	433860	615.39	51.42
3	+	X	+	X	699749	233249	466500	581.15	49.19
4	+	X	X	+	205541	68513	137028	132.25	16.04
5	+	+	+	+	825413	275137	550276	768.91	60.28
6	+	+	X	X	528005	176001	352004	160.19	19.38
7	+	+	+	X	1046981	348993	697988	744.5	58.92
8	+	+	X	+	380165	126721	253444	168.6	19.26

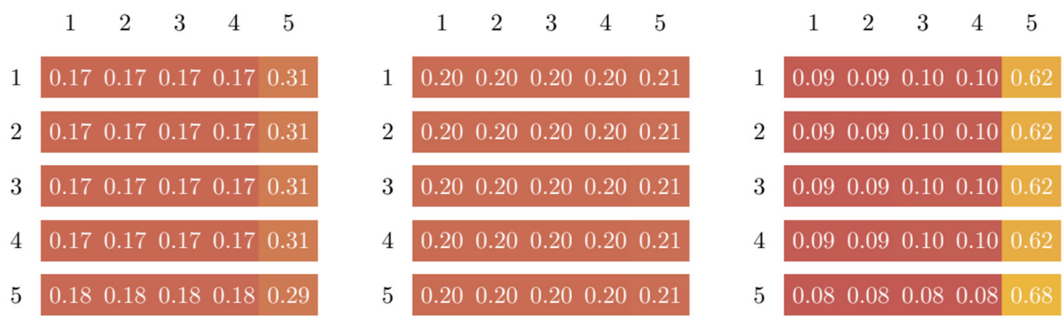
prediction of the testing dataset. Figure 10a presents averaged through all test samples. In the figure it can be seen that in most cases the last sample (the results from minute that is closest to the predicted time) has the most influence on the result of the model. However, it can also be seen that the data from all other minutes also have a very similar influence on the result.

Figure 10b and 10c presents the heatmap for the correct and incorrect prediction of the model accordingly. In such a case, we can see that the most incorrect results was achieved if model focused only on the last part of the data and the minute that was more in the past has only slight influence. The best results were achieved if the model took into consideration the data from all timespan.

To assess the impact of the Kalman filter parameters we performed an additional test with different set of those parameters: Q = matrix for the Discrete Constant White Noise Model, dimensions = 2, time step = 0.0005, variance of the noise = 0.01, P = I*500, R = 2. In this case, we obtained results that were strongly filtered and more different from the original values. After training and prediction processes, we were able to achieve only slightly better results while analyzing mean values from the cross validation. However, standard deviations were much higher, which indicated that a given simpler data model was able to learn it, but the generalization capabilities of the model are not very good for real world data.

In the case of the analyzed models, it is vital to find the trade-off between their accuracy, complexity, and stability. As can be seen in Table 3, the best results were obtained for model 3 and model 5. In the predicting load task model 3 had the highest R2 of 0.986, while the model 5 obtained also very high R2 of 0.982. It suggests a very good predictive power of those models. When analyzing errors in the load prediction task, it can be seen that model 3 has the lowest mean RMSE of 6.235, while model 5 has also very low mean RMSE of 6.385. Model 3 has the lowest mean MAE of 4.335, while model 5 has also very low mean MAE of 5.006. Model 3 has the lowest mean nRSME of 0.014, while model 5 has also very low nRMSE of 0.015. Similar observation was made in the predicting PV1 generation task.

Looking at the standard deviations, which are important because they have a significant impact on the generalization ability of the models when applied to new data in the future, it can be seen that model 5 has the lowest std MSE of 17.371 and std RMSE of 1.257). It reflects the impact of adding self-attention layers to model 5, which supported stabilization of the models by lowering the standard deviation of errors during cross-validation. In addition, model 3 and model 5 obtained very low std R2 (std R2 for model 3=0.003, while std R2 for model 5=0.011. In the predicting PV generation task model 3 outperformed the others with R2=0.993 for PV1 installation and 0.988 for PV3 installation. Model 3 obtained the lowest mean MSE=1.808 for PV1. Model 5 had not only a very good predictive power with R2=0.992 for PV 1 installation and R2=0.973 for PV3 installation, but also had the best generalization ability with the lowest standard deviations std R2=0.0003, std MSE=0.462, std RMSE=0.163, and nRMSE=0.002 for PV1. It can be concluded that Kalman filtering and LSTM layers contributed to a substantial enhancement of the achieved results. The effect of using the Kalman filter can be seen, for example, by analyzing models 5 and 13 in the load prediction task. Model 5 outperformed model 13 with an improvement in R2 of 0.116. The effect of using the LSTM layer can be noticed, for example, by analyzing models 5 and 8 in the load prediction task. Model 5 outperformed model 8, achieving a decrease in mean MSE of 132.677 and an improvement in R2 of 0.036. Besides this, adding CNN layers contributed to reducing the overall complexity of the models, which is visible in Table 4 in the reduced number of parameters.



(a):Averaged attention heatmap for whole training dataset

(b): Attention heatmap for best prediction

(c):Attention heatmap for the worst prediction

Figure 10. Analysis of the self-attention layer heatmap

CONCLUSIONS

Reliable and accurate forecasts of load and renewable energy generation are of paramount importance for energy management in the university campus microgrid. In this work, sixteen deep learning forecasting models, including various combinations of applied layers, were developed and compared to choose the best architecture. A deep neural network, Kalman Filtering and Self-Attention model was found to be a robust and reliable tool for energy management in a university campus microgrid, allowing accurate prediction of photovoltaic power generation and load forecasting. It is reflected in the very high value of the determination coefficient (R2=0.987 for load forecasting, and R2=0.989 for PV power generation forecasting) and the low MAE, RMSE and their low standard deviations (4.598 and 5.749 respectively for load forecasting and 0.953, 1.833 for PV generation). In addition, it can effectively cope with conditions different from those during training, which indicates its very good generalization ability (R2=0.982 MAE=5.006, RMSE=1.257 for load forecasting and R2=0.992 MAE=0.634, RMSE=1.395 for PV generation). In addition to this, this model obtained the lowest standard deviations. It is worth to note that all analyzed components of the proposed models had a positive impact on the results obtained:

- Kalman filtering and LSTM layers – were responsible for significant improvement of the obtained results;
- Self-attention layers – were responsible for increasing stability of the models by reducing standard deviation of the errors during cross–validation;
- CNN layers – were responsible for decreasing the complexity of the models.

Practical application of the proposed model was demonstrated by testing and validating against real-world university campus microgrid energy performance data, in real-world conditions covering all seasons of the year, and not on the simulation results or research on a small scale, which allowed to reduce the error coming from the imperfections of the training set. The proposed deep learning model is dedicated to university campuses, taking into account their specificity. In the proposed approach, applying high resolution using 1-minute interval data allowed for enlargement of the data set compared to models based on 1-hour or 15-minute intervals. The developed deep learning model allows for reliable and accurate forecasts of campus energy demand and PV energy generation, which is necessary to optimally schedule dispatchable sources (microturbines and boilers) and the electrical storage system. It helps minimize operating costs (cost of natural gas to feed micro turbines and boilers and cost of electrical energy). In the future, the proposed deep learning system can be further developed and connected to cost-related data (natural gas price, electricity price, fees for the local electricity production, etc.). The limitation of this work could be that the data from the same location and similar climatic conditions were used to check the generalization ability of the model and to train and test the model. Therefore, additional research can be performed to check the model's ability to adapt to different geographical locations and climatic conditions. Another limitation of the model is that its accuracy may decline over time due to climate change. Therefore, it may be necessary to incorporate a retraining module containing new weather data with new climate patterns.

REFERENCES

1. Aslam S, Herodotou H, Mohsin SM, Javaid N, Ashraf N, Aslam S. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. Renewable and Sustainable Energy Reviews. 2021;144: 110992. https://doi.org/10.1016/j.rser.2021.110992
2. Amjady N, Keynia F, Zareipour H. Short-term load forecast of microgrids by a new bilevel prediction strategy. IEEE Transactions on smart grid. 2010;1: 286–294. https://doi.org/10.1109/TSG.2010.2078842
3. Krechowicz A, Krechowicz M, Pawelec A. Machine learning models for electricity generation forecasting from a PV farm. IFIP international conference on artificial intelligence applications and innovations. Springer; 2024; 252–264. https://doi.org/10.1007/978-3-031-63215-0_19
4. Bojarczak P, Osuch–Słomka E, Stanisławek S, Lucedarski B. Detection of selected gear tooth defects using deep neural networks. Advances in Science and Technology Research Journal. 2025.
5. Krechowicz A, Deniziak S, Kaczmarski D. Machine learning approach to gait deviation prediction based on isokinetic data acquired from biometric sensors. Gait & Posture. 2023;101: 55–59. https://doi.org/10.1016/j.gaitpost.2023.01.015

6. Chwaleba K, Wach W. A convolutional neural network – driven model with adaptive feature fusion for polish national dance music recognition. Advances in Science and Technology Research Journal. 2026;20. https://doi.org/10.12913/22998624/210059
7. Krechowicz M, Krechowicz A, Lichołai L, Pawelec A, Piotrowski JZ, Stępień A. Reduction of the risk of inaccurate prediction of electricity generation from PV farms using machine learning. Energies. 2022;15: 4006. https://doi.org/10.3390/en15114006
8. Krechowicz A, Krechowicz M, Poczeta K. Machine learning approaches to predict electricity production from renewable energy sources. Energies. 2022;15: 9146. https://doi.org/10.3390/en15239146
9. R. Singh A, Kumar RS, Bajaj M, Khadse CB, Zaitsev I. Machine learning-based energy management and power forecasting in grid-connected microgrids with multiple distributed energy sources. Scientific Reports. 2024;14: 19207. https://doi.org/10.1038/s41598-024-70336-3
10. Yuan G, Wang H, Khazaei E, Khan B. Collaborative advanced machine learning techniques in optimal energy management of hybrid AC/DC IoT-based microgrids. Ad Hoc Networks. 2021;122: 102657. https://doi.org/10.1016/j.adhoc.2021.102657
11. Ma Y-J, Zhai M-Y. Day-ahead prediction of microgrid electricity demand using a hybrid artificial intelligence model. Processes. 2019;7: 320. https://doi.org/10.3390/pr7060320
12. Indira G, Bhavani M, Brinda R, Zahira R. Electricity load demand prediction for microgrid energy management system using hybrid adaptive barnacle-mating optimizer with artificial neural network algorithm. Energy Technology. 2024;12: 2301091. https://doi.org/10.1002/ente.202301091
13. Da TN, Cho M-Y, Thanh PN. Hourly load prediction based feature selection scheme and hybrid CNN-LSTM method for building's smart solar microgrid. Expert Systems. 2024;41: e13539. https://doi.org/10.1111/exsy.13539
14. Thanh PN, Cho M-Y, Chang C-L, Chen M-J. Short-term three-phase load prediction with advanced metering infrastructure data in smart solar microgrid based convolution neural network bidirectional gated recurrent unit. IEEE Access. 2022;10: 68686–68699. https://doi.org/10.1109/ACCESS.2022.3185747
15. Lan T, Jermsittiparsert K, T. Alrashood S, Rezaei M, Al-Ghussain L, A. Mohamed M. An advanced machine learning based energy management of renewable microgrids considering hybrid electric vehicles' charging demand. Energies. 2021;14: 569. https://doi.org/10.3390/en14030569
16. Wen L, Zhou K, Yang S, Lu X. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy. 2019;171: 1053–1065. https://doi.org/10.1016/j.energy.2019.01.075
17. Trivedi R, Patra S, Khadem S. A data-driven short-term PV generation and load forecasting approach for microgrid applications. IEEE Journal of Emerging and Selected Topics in Industrial Electronics. 2022;3: 911–919. https://doi.org/10.1109/JESTIE.2022.3179961
18. Elabbassi I, Khala M, Eloutassi O, et al. Evaluating and comparing machine learning approaches for effective decision making in renewable microgrid systems. Results in Engineering. 2024;21: 101888. https://doi.org/10.1016/j.rineng.2024.101888
19. Ali W, Ulasyar A, Mehmood MU, Khattak A, Imran K, Zad HS, et al. Hierarchical control of microgrid using IoT and machine learning based islanding detection. IEEE Access. 2021;9: 103019–103031. https://doi.org/10.1109/ACCESS.2021.3098163
20. Ji Y, Wang J, Xu J, Fang X, Zhang H. Real-time energy management of a microgrid using deep reinforcement learning. Energies. 2019;12: 2291. https://doi.org/10.3390/en12122291
21. Ashok Babu P, Mazher Iqbal JL, Siva Priyanka S, Jithender Reddy M, Sunil Kumar G, Ayyasamy R. Power control and optimization for power loss reduction using deep learning in microgrid systems. Electric Power Components and Systems. 2024;52: 219–232. https://doi.org/10.1080/15325008.2023.2217175
22. Madabathula CT, Agrawal K, Mehta V, Kasarabada S, Kommamuri SS, Liu G, et al. Smart green energy management for campus: An integrated machine learning and reinforcement learning model. Smart Cities. 2025;8: 30. https://doi.org/10.3390/smartcities8010030
23. Muqeet HA, Javed H, Akhter MN, Shahzad M, Munir HM, Nadeem MU, et al. Sustainable solutions for advanced energy management system of campus microgrids: Model opportunities and future challenges. Sensors. 2022;22: 2345. https://doi.org/10.3390/s22062345
24. Kim J, Oh H, Choi JK. Learning based cost optimal energy management model for campus microgrid systems. Applied Energy. 2022;311: 118630. https://doi.org/10.1016/j.apenergy.2022.118630
25. Bekhit R, Bianco G, Delfino F, Ferro G, Noce C, Orrù L, et al. A platform for demand response and intentional islanding in distribution grids: The LIVING GRID demonstration project. Results in Control and Optimization. 2023;12: 100294. https://doi.org/10.1016/j.rico.2023.100294
26. Bracco S, Brignone M, Delfino F, Procopio R. An energy management system for the Savona campus smart polygeneration microgrid. IEEE Systems Journal. 2015;11: 1799–1809. https://doi.org/10.1109/JSYST.2015.2419273