Storage management optimization based on electrical consumption and production forecast in a photovoltaic system

Anthony Aouad, Khaled Almaksour, Dhaker Abbes

Storage management optimization based on electrical consumption and production forecast in a photovoltaic system

Anthony Aouad[a], Khaled Almaksour[a,*], Dhaker Abbes[a]

[a] *Univ. Lille, Arts et Metiers Institute of Technology, Centrale Lille, Junia, ULR 2697 - L2EP, F-59000 Lille, France*

Abstract

Decentralized energy production, particularly from photovoltaic (PV) systems, is becoming increasingly prevalent, leading to a rise in the number of energy producers and consumers, or ”prosumers”. These prosumers, equipped with their own energy generation and storage systems, are not just passive consumers but active participants in the energy market. They generate their own electricity, often from renewable sources, and can feed excess power back into the grid, store it for later use, or share it within a local energy community. This evolving energy paradigm presents new opportunities and challenges in terms of energy management and optimization, necessitating innovative approaches to ensure efficient and sustainable use of energy resources. This paper introduces an innovative storage management method for grid-connected photovoltaic (PV) systems. The method is designed to minimize either the economic or ecological cost, or to find an optimal balance between the two, under various tariff scenarios. This is achieved while adhering to a full self-consumption constraint imposed by the distribution system operator. The control strategy is underpinned by forecasts of electrical consumption, production, and CO_2 emissions, which are developed using feedforward neural network models. These models are trained on data from a real-scale smart-grid demonstrator at the Catholic University of Lille, France. The results of the study offer a comparative analysis of the economic and ecological benefits of the three proposed strategies, demonstrating that the best compromise is achieved when considering the off-peak tariff option. Furthermore, a real-time controller was implemented on the Energy Management System (EMS) of the demonstrator and tested over a 24-hour period, yielding satisfactory results. This paper, therefore, presents a significant advancement in the field of storage management for grid-connected PV systems.

Keywords: Storage management, Smart Grids, Photovoltaic power, Optimization, Consumption forecast, Production forecast, Photovoltaic self-consumption

1. Introduction

Declining costs of PV equipment and other renewable energy sources, as well as government incentives, are driving decentralized energy production growth in the context of the smart grid. The number of energy producers and consumers (’prosumers’) is increasing significantly. In the first quarter of 2021, the main distribution system operator (DSO) in France, Enedis, had more than 100,000 customers connected to the distribution grid with individual self-consumption contracts. In addition to these connections, there are one hundred and two collective self-consumption operations that are active at the end of the second quarter of 2022. Whereas in 2015, Enedis had only 3,000 individual self-consumption installations, the rate of connection has accelerated exponentially over the last five years. This strong trend reflects a change in consumption patterns in favour of the development of renewable energies and the ecological transition [1], [2]. Energy prosumers generally remain connected to the central electricity network. However, they are also capable of generating and even storing energy, usually with photovoltaic solar panels and batteries, with the aim to maximize their self-consumption rate and reduce their energy costs. Therefore, an adapted multi-objective

*Corresponding author

Email addresses: `anthonyaouad94@gmail.com` (Anthony Aouad), `khaled.almaksour@junia.com` (Khaled Almaksour), `dhaker.abbes@junia.com` (Dhaker Abbes)

supervision is required to ensure energy balance and to properly exploit the storage. The challenge of the development of these supervision strategies is the random behaviour of these multi-source systems, whose time horizons can be very short (dynamic responses) or very long (seasonality of renewable sources). Authors of article [3] show that depending on the time scale, several types of multi-source system supervision exist, including long-term supervision (daily to monthly, based on the prediction and anticipation of input variables), medium-term supervision (usually from one hour to half an hour, which provides forward-looking instructions based on forecasts with average input data), and real-time supervision, which allows forecast correction according to the measurements. Real time supervision is related to the instantaneous nature of electricity. Regarding the energy supervision of multi-source systems, a state of the art followed by an analysis of the specifics of each of the methods is carried out in [4]. The three families of tools offered in this regard are:

- Causal formalization tools, of which the inversion of the power balance makes it possible to determine reference powers (a detailed mathematical model is necessary as well as the behaviour of the system in real time).
- Explicit optimization tools, which allow making an optimal choice by formulating constraints, optimization variables and an objective function to be minimized or maximized (e.g., minimization of CO_2 emissions). This approach is difficult to implement instantaneously. The optimization problem is often complex with several variables and several objectives and therefore time consuming to resolve. However, medium term energy management is becoming increasingly popular (up to 10min).
- Implicit optimization tools, including fuzzy logic ([5] and [6]).

In this paper, we have opted for an explicit optimization method for the energy management of an on-grid PV system with batteries in the context of individual self-consumption. This choice is justified by the existence of a cost function (economic or ecological) to be minimized and by the desire to consider the constraint of non-injection into the network.

The management method uses multiple criteria (economic, ecological, and best compromise between the two), in different tariff scenarios while respecting a full self-consumption constraint imposed by the DSO. It is based on electrical consumption, production, and CO_2 forecasts developed using feedforward neural network models based on data from a real-scale smart-grid demonstrator at the Catholic university of Lille, France.

The management is predictive and done at two-time levels: day ahead and real time. The day ahead planning is obtained by optimizing three possible main objectives, economic, ecological and the compromise between the two. Whereas real time control applies rules to ensure power balance and to respect the self-consumption constraint imposed on the photovoltaic power production by the DSO to avoid any injection of power to the main distribution grid and between independent buildings. The energy management step time is 10 minutes over 24 hours. The proposed method is validated in the real-scale demonstrator of the Catholic university of Lille in France.

The paper is organised as follows: Section 2 presents a literature review. Section 3 provides a description of the studied system. The storage control methodology is presented in section 4. The first part of this section is dedicated for the forecast methodology and the results for the three forecasts (electrical consumption, PV production and CO_2 emissions). The energy management optimization is presented in the second part of this section. Section 5 is dedicated for simulation results and experimental results derived from the smart-grid demonstrator at the Catholic university of Lille. A conclusion is conducted in section 6.

2. Literature Review

Energy management optimization in micro-grids and energy communities is a topic that has been extensively researched. Microgrids are a scaled-down version of utility grids that have gained attention in recent decades due to their distinct features such as the utilization of renewable energy resources and elimination of power transmission requirements. The intermittent nature of distributed generation resources and the need for improving the economic feasibility of microgrids have made energy management an important research area. In this context, Oprea et al. [7] proposed a novel concept of a Smart Adaptive Switching Module (SASM) that leverages fuzzy logic to efficiently manage electricity generated from renewable energy sources (RES). The SASM is designed to gradually switch electric appliances based on a variety of factors including weather sensors, power forecast, storage system constraints, among others. The effectiveness of the proposed SASM architecture was demonstrated through a case study of a RES system located in Hulubesti, Romania. However, a more in-depth analysis of the results obtained from the case study would have provided deeper insights into the practical implications of the SASM. Nsilulu T. Mbungu et al. [8] delve in their study into various control and estimation techniques applied to smart micrograms. They discuss a plethora of control techniques, including linear, non-linear, robust, predictive, intelligent, and adaptive control techniques. They also underscore the importance of accurate data for a better performance index to ensure the efficiency of the power network. This paper provides a fundamental conceptual framework for selecting an optimal design modeling strategy and policy-making decisions to control, monitor, and protect the innovative electrical network. A review of optimization techniques used in microgrid energy management systems found that mixed integer linear programming is the most used optimization technique [9]. Multi-agent systems are most ideal for solving unit commitment and demand management. State-of-the-art machine learning algorithms are used for forecasting applications. The meta-heuristic algorithms are commonly used in economic dispatch application.The review also found that the multi-agent-based techniques and meta-heuristics algorithms outperformed other conventional techniques in terms of system efficiency due to the decentralized nature of the EMS problem in microgrids and the capability of these techniques to act effectively in such scenarios. However, it was also evident that the use of advanced optimization techniques was limited in the scope of forecasting and demand management. Abdel-Nasser and Mahmoud [10] introduced a novel approach to PV power forecasting using deep Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN). They argued that LSTM-RNNs are particularly suited for this task due to their ability to model temporal changes in PV output power. However, a more detailed explanation of the LSTM-RNN model and its advantages over other types of neural networks would have provided a more comprehensive understanding of this approach. Furthermore, discussing the practical implications of their findings, particularly how the improved accuracy of their model could impact the operation of smart grids in real-world scenarios, would have added value to their study. Hossain et al. [11] presented a novel approach to forecasting the power output of PV systems using the Extreme Learning Machine (ELM) approach. They claimed that the ELM model offers higher accuracy and less computational time in forecasting the daily and hourly PV output power compared to other popular models such as Support Vector Regression (SVR) and Artificial Neural Network (ANN). However, potential issues with the ELM approach, such as overfitting or sensitivity to parameter selection, were not thoroughly discussed in their study. A more detailed discussion of these issues and how they were addressed in their study would have strengthened their methodology. Oprea and Bâra [12] conducted a comprehensive study on the use of big data technologies for the management of photovoltaic power plants. They proposed an ultra-short-term forecast (USTF) algorithm that uses a Feed-Forward Artificial Neural Network (FF-ANN) and a backtracking adjustment of the learning rate for faster convergence. The effectiveness of their approach was demonstrated through two case studies - PV Agigea and PV Giurgiu located in Romania. However, a comparison of their approach with other existing methodologies or algorithms for PV forecasting would have highlighted the advantages of their approach and its potential improvements over existing methods. Following this literature review, this study makes a significant contribution to the field of energy management in grid-connected photovoltaic systems by introducing a novel approach that focuses on individual self-consumption, a topic that has gained increasing relevance in the context of decentralized energy production. The paper presents an explicit optimization method that takes into account multiple criteria, including economic, ecological, and a balance between the two, under different tariff scenarios. This approach provides a comprehensive framework for energy management, demonstrating the potential of advanced machine learning techniques in forecasting electrical consumption, photovoltaic production, and CO_2 emissions. Furthermore, the paper bridges the gap between theoretical research and practical application by validating the proposed method using data from a real-scale smart-grid demonstrator at the Catholic university of Lille, France. Importantly, the study addresses the full self-consumption constraint imposed by the distribution system operator, a critical aspect of energy management in grid-connected photovoltaic systems that has not been extensively explored in previous research, thereby enriching the existing literature in this field.

3. Description of the smart-grid demonstrator

The demonstrator is located at Lille Catholic university campus in the north of France. It contains 4 academic buildings named HEI1, HEI2, HA and RIZOMM (Fig. 1). It consists of the following parts:

- Two photovoltaic rooftop generators. A 189 kWp PV system with a surface of 1200 m² is installed on the roof of RIZOMM building. This system is connected to the grid via 10 FRONIUS PV inverters (2 x 8.2 kVA, 1 x 10 kVA, 1 x 15 kVA, 4 x 17.5 kVA, 2 x 20 kVA). The second PV system is installed on the roof of HEI2 building with an installed power of 28 kWp and a surface of 150 m². It is connected to the grid via two 12.5 kVA FRONIUS PV inverters.

- A Li-ion Eaton storage system with a capacity of 250 kWh is installed in the underground of HEI1 building. It consists of five strings of 50 kWh of capacity for each string. The storage is connected to the grid via two inverters and has a rated power of 80 kW in discharge and 40 kW in charging.
- Six charging points for electric vehicles with a charging power of 22 kW for each point. They are connected to the grid via HA building LVDB (Low Voltage Distribution Board).
- Measurement station (Socomec DIRIS Digiware D-70) is installed in the main distribution board of the campus. Power consumption/production measurements are provided in real-time by the station to the EMS of the demonstrator.
- EMS of the demonstrator is installed on a dedicated server. Modbus TCP/IP protocol is used for the communication of the EMS with all devices. The data acquisition is done using Python codes which allow the filtering and the storage of the relevant data on a NoSQL database. Others Python codes are used to provide different forecasts (detailed in section 4.1) and control the storage.

This campus network is connected to the public distribution grid via 1 MVA 15kV/0.4kV transformer (Fig. 2).

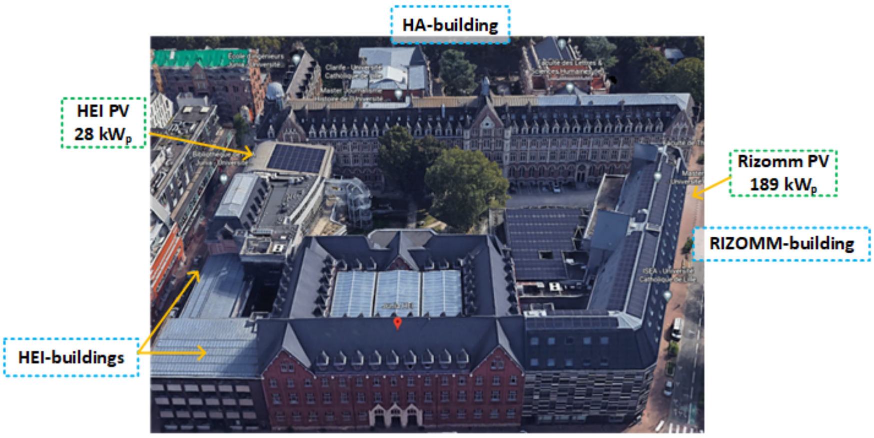

Figure 1: Lille Catholic university demonstrator

The engineering school JUNIA own two buildings HEI1 and HEI2 while the Catholic university own the other buildings HA and RIZOMM. As the buildings belong to two separate legal entities, the law in France does not allow the exchange of energy between the two entities.

The Catholic university buildings with the 189 kWp photovoltaic system are the case study considered for this work and the energy produced is to be consumed by the local loads of the Catholic University. Thus, the storage control method proposed in the study considers the self-consumption constraint that prevents any excess photovoltaic production from being injected from the Catholic university to the HEI buildings or to the public distribution grid.

4. Proposed storage management method

Energy management is done in two stages: day ahead and real time. The day ahead planning is optimized following three possible main objectives, economic, ecological (CO_2 cost) and the compromise between the two. Whereas real time control applies rules to manage power flows and to respect the self-consumption constraint imposed on the photovoltaic power production by the distribution system operator to avoid any injection of power to HEI buildings or

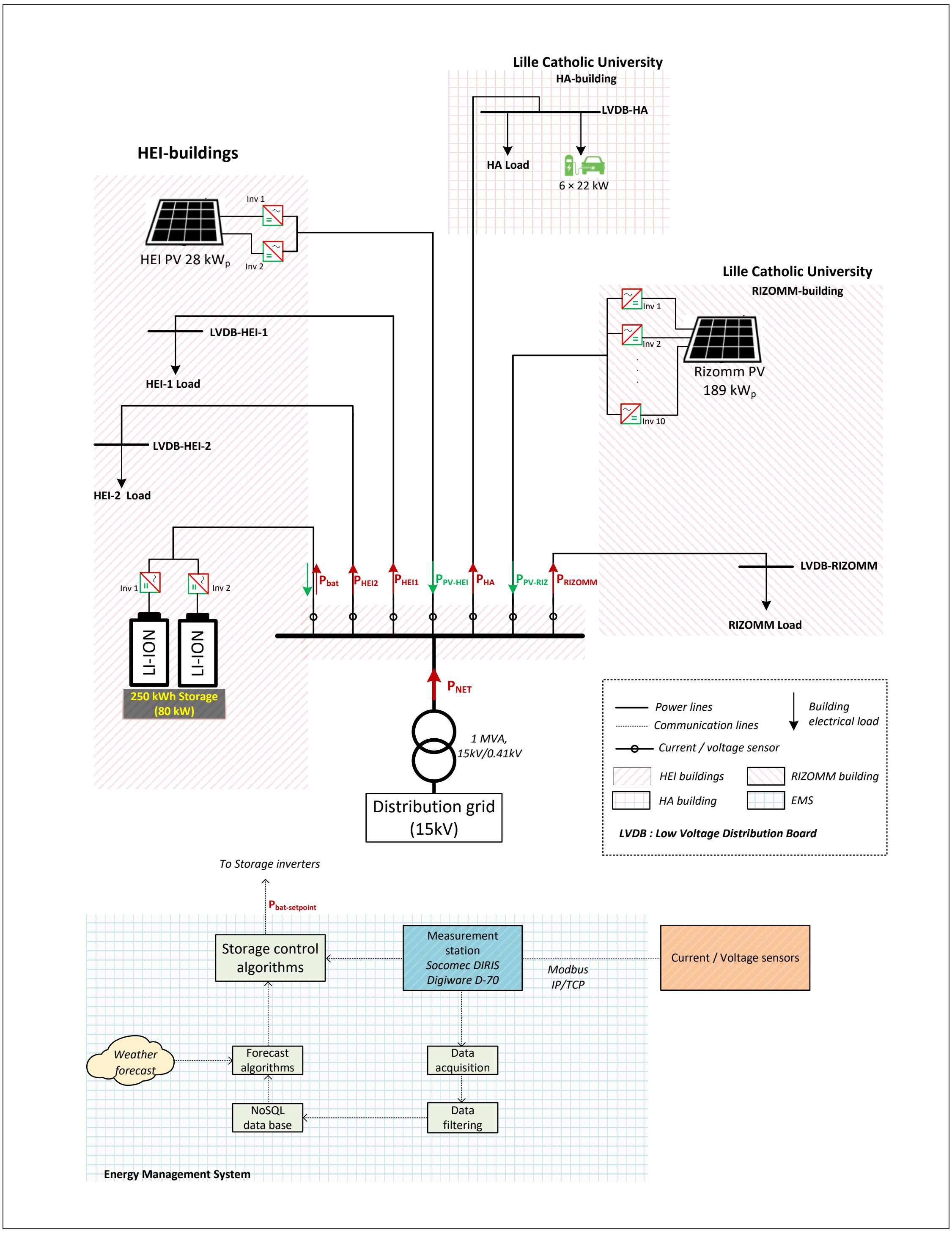

Figure 2: Diagram of the Lille Catholic University demonstrator

to the main distribution grid. To control the storage with the desired objectives, day-ahead forecasts of electrical consumption, local PV production and CO_2 intensity are needed. The forecast methodology is presented in the following sections as well as the energy management optimization.

4.1. Forecast methodology and validation

Electrical consumption and local PV production data of the demonstrator are available for three years. Machine learning (ML) is used to develop models that provide the different needed forecasts. In order to choose the best ML algorithm and using historical data from the smart grid demonstrator, multiple machine learning models were developed and compared including models from the boosting family of machine learning algorithms which try to convert weak learners to strong learners [13] (Catboost, Light Gradient Boosting Machine, Extreme Gradient Boosting, Gradient Boosting Regressor, AdaBoost Regressor) as well as decision tree algorithms which, by learning straightforward decision rules derived from the data features, create a model that forecasts the value of a target variable [14] (Decision Tree Regressor, Extra Trees Regressor, Random Forest Regressor) and finally a feedforward neural network model consisting of four layers of which two are hidden [15]. A detailed comparison is given for the consumption forecast, where the most suitable model is chosen and reused in the following sections.

In order to compare the different ML algorithms, the total electrical consumption data set of 400 days of the demonstrator is used to provide and compare forecast provided by different algorithms. The input variables were chosen using correlation studies as well as trial and error with all available parameters. For example Fig. 3 shows the auto-correlation of the consumption data where the peaks of the graph after lagging point 0 correspond to the same hour and minute of the previous seven days and were therefore chosen as input values.The training results of the comparison are presented in table 1 (Where, nRMSE is the normalized root mean square error, R2 is the coefficient of determination, RMSLE is the root mean squared logarithmic error and MAPE is the mean absolute percentage error). An example of the forecast provided by the four best ML algorithms found for the day of Tuesday September 27th

Model	nRMSE (%)	R2	RMSLE	MAPE (%)	Training Time (Sec)
Feedforward Neural Network	3.51	0.9906	0.0371	2.44	120.034
CatBoost Regressor	3.95	0.9876	0.0476	3.15	25.223
Extra Trees Regressor	4.35	0.9864	0.0476	3.64	200.980
Extreme Gradient Boosting	4.44	0.9862	0.0491	3.69	19.742
Light Gradient Boosting Machine	4.96	0.9790	0.0614	4.64	6.172
Decision Tree Regressor	6.23	0.9706	0.0689	4.15	5.124
Random Forest Regressor	7.98	0.9636	0.0796	4.63	49.540
Gradient Boosting Regressor	9.47	0.9396	0.0985	6.99	19.874
Linear Regression	15.68	0.8350	0.1664	12.25	8.409

Table 1: Results table for consumption prediction using various machine learning models.

2022 is shown in Fig. 4. The feedforward neural network performed the best however, it is closely followed by the CatBoost machine learning model which has a lower training time, noting that the training time includes the model tuning time in PyCaret. Additionally, CatBoost and the neural network are the only models where the error did not significantly increase when testing with unseen data. An ensemble model using the previous two models should yield the best result, however after looking at the feature importance plot, i.e. the contribution of each input variable to the output result, (Fig. 5) of the CatBoost model, the neural network model was chosen as it yielded better results using the year and school holiday inputs which are disregarded by the CatBoost model. Both factors are considered important as the demonstrator decreases its consumption from year to year as it gets more efficient and school holidays have decreased consumption when compared to regular days. Based on this comparative study, feedforward neural network is used to provide all type of forecast presented in the following sections.

4.1.1. Consumption forecast

Day ahead consumption load forecast is calculated for each building of the demonstrator at midnight. The input layer of the model consists of 15 variables: temperature, day of the week, month, hour, weekend (or not), day of the year, public holiday (or not), school holiday (or not), as well as the consumption of the last seven days. Consumption

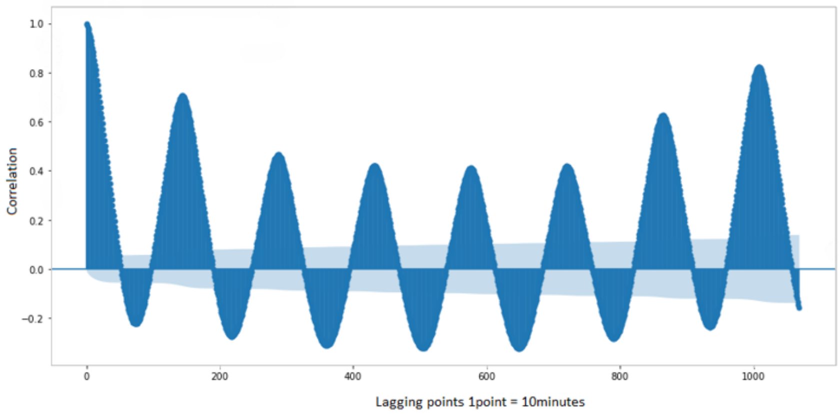

Figure 3: Auto-correlation coefficient for consumption data over a period of a week i.e. over 1008 lagging points

data is sourced from the smart grid demonstrator database, temperature data from Solcast [16], the time period for the historical data is a sliding window of 400 days since the demonstrator keeps evolving and data beyond this period decreases the accuracy of the models.The available data is divided into two parts; 70% of data chosen randomly and used for the training stage and the rest (30%) used for validation.

The models were developed using TensorFlow [17] (for the neural network) and PyCaret [18] libraries in the python programming language.

The feedforward neural network model is composed of 4 layers including 2 hidden layers consisting respectively of 1024 neurons and 512 neurons as shown in Fig. 6. The number of neurons chosen for the network is based on training time and results. Fig. 7 represents the forecast results during a normal week of operation of October 2021 (without holidays or school closures). The same structure is used for all forecast results presented in the following sections.

4.1.2. PV production forecast

Thanks to three years of PV production data set available for both PV systems of the demonstrator, and using historical meteorological data, a feedforward neural network model was developed. Photovoltaic power forecasts are based on meteorological forecasts using 4 main variables from the Solcast [16] database: global horizontal irradiation (GHI), direct normal irradiation (DNI), temperature, and cloud coverage. These variables were chosen using correlation studies as well as trial and error with all available parameters. It should be highlighted that the weather data used are the weather forecast obtained every day at midnight for 24 hours. These data were saved in a database to test the accuracy of the forecast algorithm with realistic data that contains the weather forecast error. The feedforward neural network model is composed of 4 layers including 2 hidden layers consisting of 1024 neurons and 512 neurons respectively. A rectified linear activation function (ReLU) is used for both hidden layers. All data used for training and prediction are normalized using equation (1):

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

where X represents the original value, $X_{\min}$ is the minimum value of input data, and $X_{\max}$ is the maximum value of input data.70% of available data are used for training and 30% of data are used for validation. The mean squared error (MSE) is used to calculate the loss during the NN training and the mean absolute error (MAE) is used with the MSE for the validation of the NN model. The model is trained for 300 epochs, with a batch size of 720. The training is performed using historical meteorological data obtained from Solcast database and PV production data of the demonstrator. The prediction of PV production is calculated for 24 hours using the NN model and the forecast data of the four meteorological variables mentioned above as input data. The weather prediction provided by Solcast is generated using satellite forecasting based on latest cloud image for four hours horizon with an update each 30

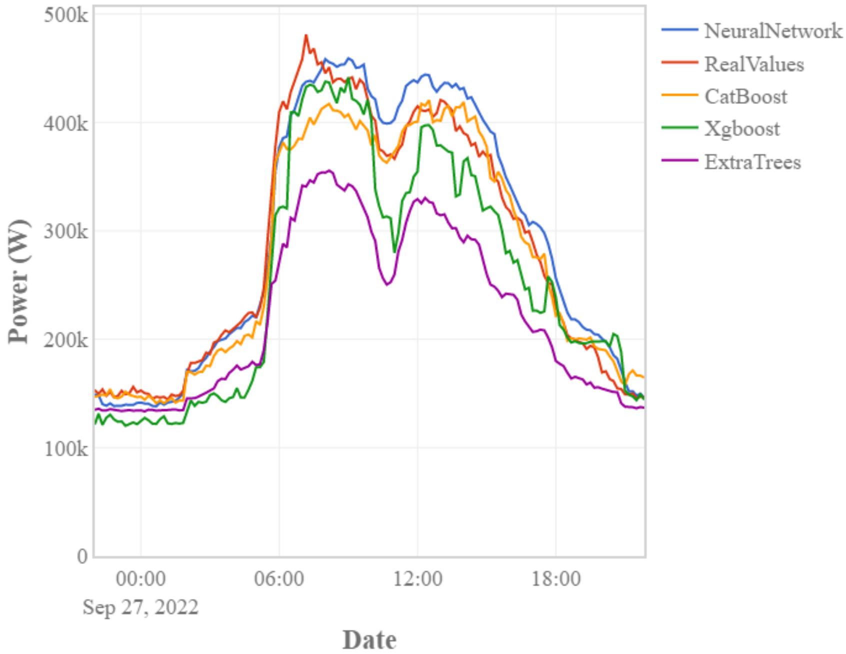

Figure 4: Consumption power prediction for various models with real values in red

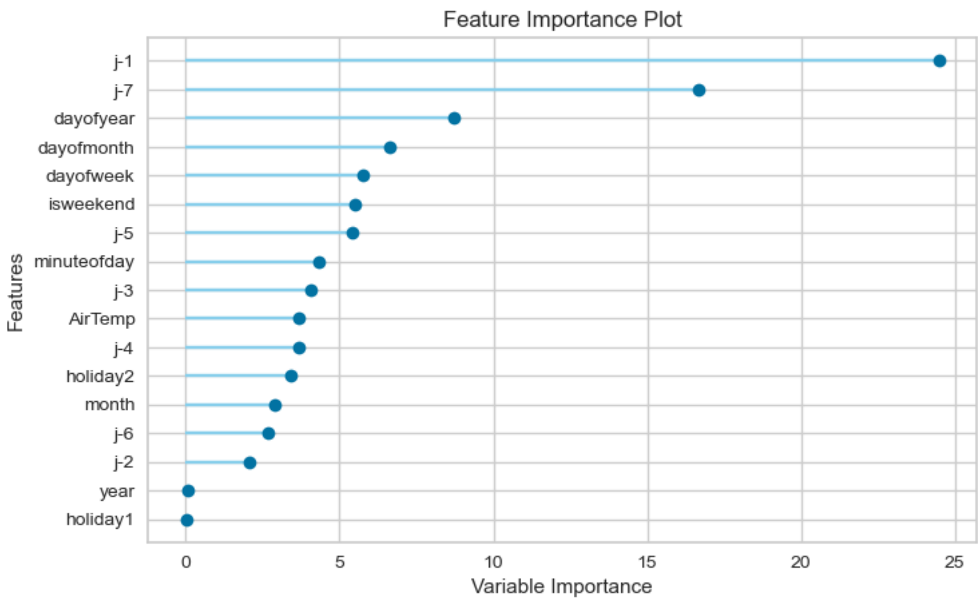

Figure 5: Feature importance plot for all consumption forecast inputs, holiday 1 being a school holiday and holiday 2 a public holiday

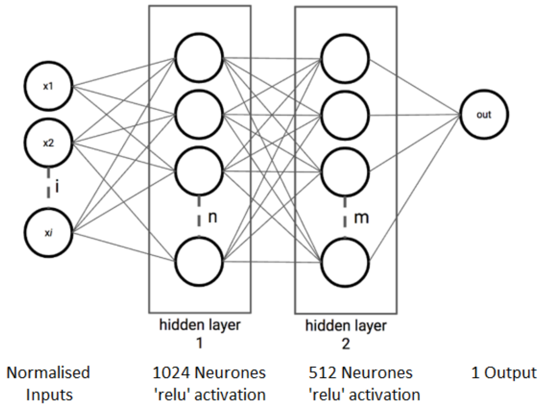

Figure 6: Structure of the Neuron Network

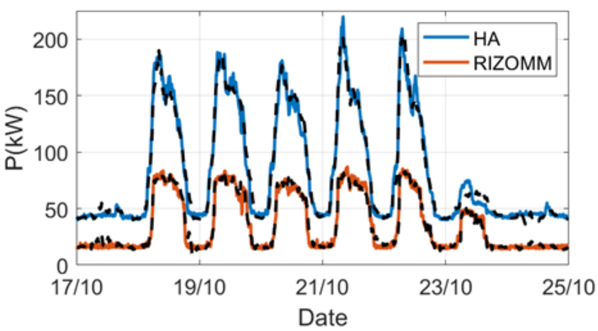

Figure 7: Load forecast of HA building and Rizomm building for a week of October 2021. Dashed lines are forecast results and continues lines are measured values

minutes. Numerical Weather Prediction is used for weather prediction for more than 4 hours up to 14 days with lower accuracy. The prediction of PV production is updated every two hours to increase the prediction accuracy. The model has an average normalized root mean square error of 10.01% for days with different weather conditions. . Fig. 8, Fig. 9 and Fig. 10 represent the forecast results for different types of days, including sunny, cloudy, or partially cloudy. PV forecast has higher forecast errors than the consumption forecast as it is totally dependent on meteorological data forecasts which have their own associated errors. Errors of PV production forecast for the three types of day are presented in table 2. The forecast errors for partially cloudy day are bigger than the errors of sunny and cloudy day. This could be explained by the error in weather forecast which cannot predict the fluctuation of the irradiation because of variable and partially cloud coverage.

Day type	RMSE (W) (%)	nRMSE (%)	MAE (W)	nMAE (%)	R2
Sunny day (14th of August 2021)	6788	4.74	4261	2.97	0.97
Cloudy day (26th of September 2021)	7827	5.46	4280	2.99	0.8
Partially cloudy day (29th of July 2021)	13297	9.26	8140	5.68	0.88

Table 2: Results table for PV production prediction of RIZOMM PV system for different weather conditions

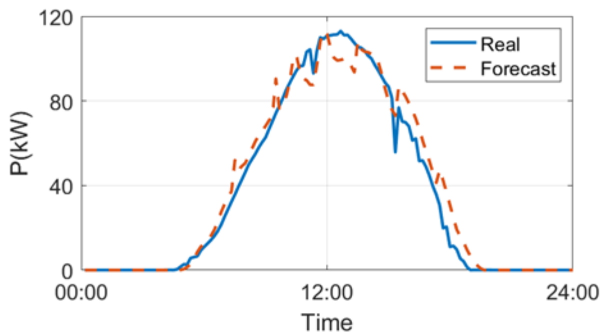

Figure 8: PV production forecast for RIZOMM PV system of a sunny day (14th of August 2021)

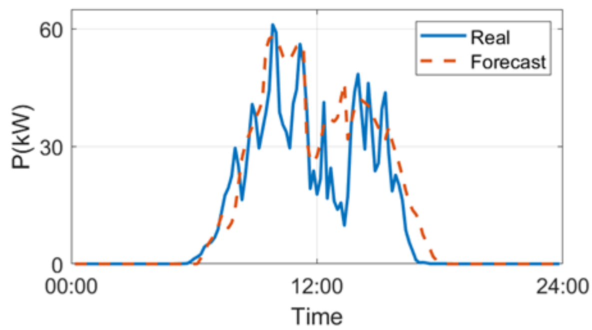

Figure 9: Production forecast for RIZOMM PV system of a cloudy day (26th of September 2021)

The PV production forecast is provided using the same NN algorithm with the same configuration for the second PV system of the demonstrator which belong to HEI building. Fig 11 presents an example of PV production forecast for a sunny day and a cloudy day. The different metrics of the PV production forecast of HEI are provied in table 3.

4.1.3. CO_2 forecast

The CO_2 content of the electric kWh corresponds to the CO_2 emissions generated by the production of this kWh of electricity as well as the global warming potential of each electric energy source by analyzing its life cycle [19].

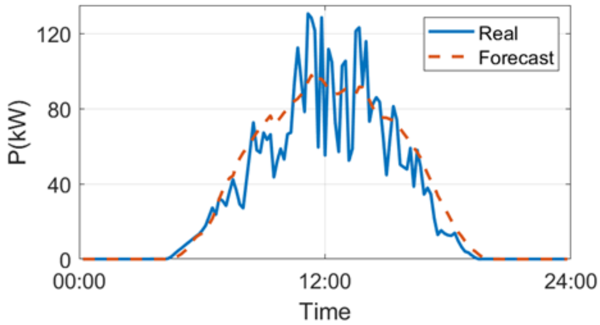

Figure 10: PV production forecast for RIZOMM PV system of a partially cloudy day (29th of July 2021)

Day type	RMSE (W) (%)	nRMSE (%)	MAE (W)	nMAE (%)	R2
Sunny day (14th of August 2021)	1660	6.67	1080	4.34	0.95
Cloudy day (26th of September 2021)	1354	5.44	721	2.9	0.8

Table 3: Results table for PV production prediction of HEI for different weather conditions

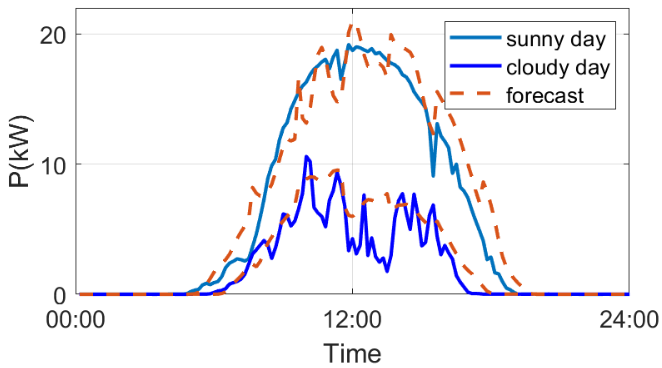

Figure 11: PV production forecast for HEI PV system of a sunny day (14th of August 2021) and cloudy day (26th of September 2021)

Forecasts are based on consumption, wind, and solar day-ahead forecasts, as well as the CO_2 content of the last 2 weeks given by the national transmission system operator (TSO). The method is inspired by the works of the United Kingdom's national grid operator National Grid ESO which developed a carbon intensity forecasting tool [20].The data is sourced from the french national grid operator RTE who has an open data platform [21] that includes historical consumption and production data on the national level since 2012. The model has a normalized root mean square error of 3.80%. Fig. 12 represents the forecast results during a week of August 2021.

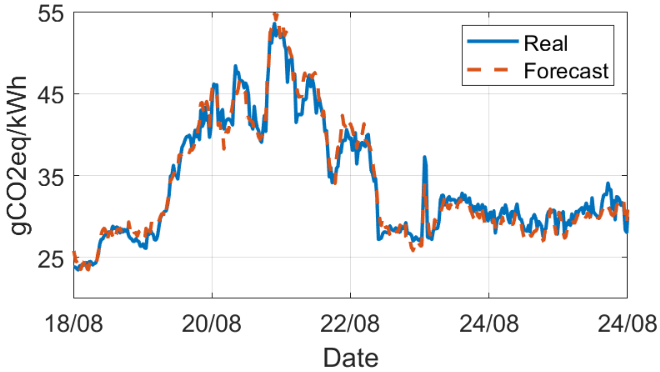

Figure 12: Forecast of grams of CO_2 equivalent per kWh of a week in August 2021

4.2. Energy management optimization

Day-ahead planning is based on PV, load and CO_2 forecasts calculated at midnight. It is obtained by deterministic optimization using Sequential Least SQuares Programming (SLSQP) algorithm using objective function given by equation (2). SLSQP is chosen because of the non linear constraints due to battery and inverter efficiencies' outlined in section 4.2.2 and equation (4).

$$min \quad \sum_{i=1}^{144} (p_{c-i} - p_{pv-i} + p_{bat-i}) \times \frac{e_i}{6} \tag{2}$$

Where:

- p_{c-i}: day ahead load consumption forecast at 10-minute intervals in kW (average value over 10-minute period)
- p_{pv-i}: day ahead photovoltaic power production forecast at 10-minute intervals in kW (average value over 10-minute period)
- p_{bat-i}: battery charging or discharging power in 10-minute steps in kW (average value over 10-minute period)
- e_i: cost coefficient of gCO_2equiv/kWh or €/kWh. The division of e_i by 6 gives the cost coefficient of an average power value over 10 minutes period.

The cost coefficient will depend on the chosen strategy. The energy management step time is 10 minutes over a period of 24 hours, which was found to be satisfactory. Real time power flow management is rule-based. 13 represents the forecast diagram of the different elements of the cost function in equation2.

4.2.1. Strategies and Constraints

For the specific case of the Catholic university SG demonstrator, the locally produced power by the PV system should be fully consumed by the local load thus, no injection to the distribution grid of extra produced power is permitted. To fulfil this condition and other objectives, the control of the storage is done at two-time levels: day-ahead optimization and real-time controller.

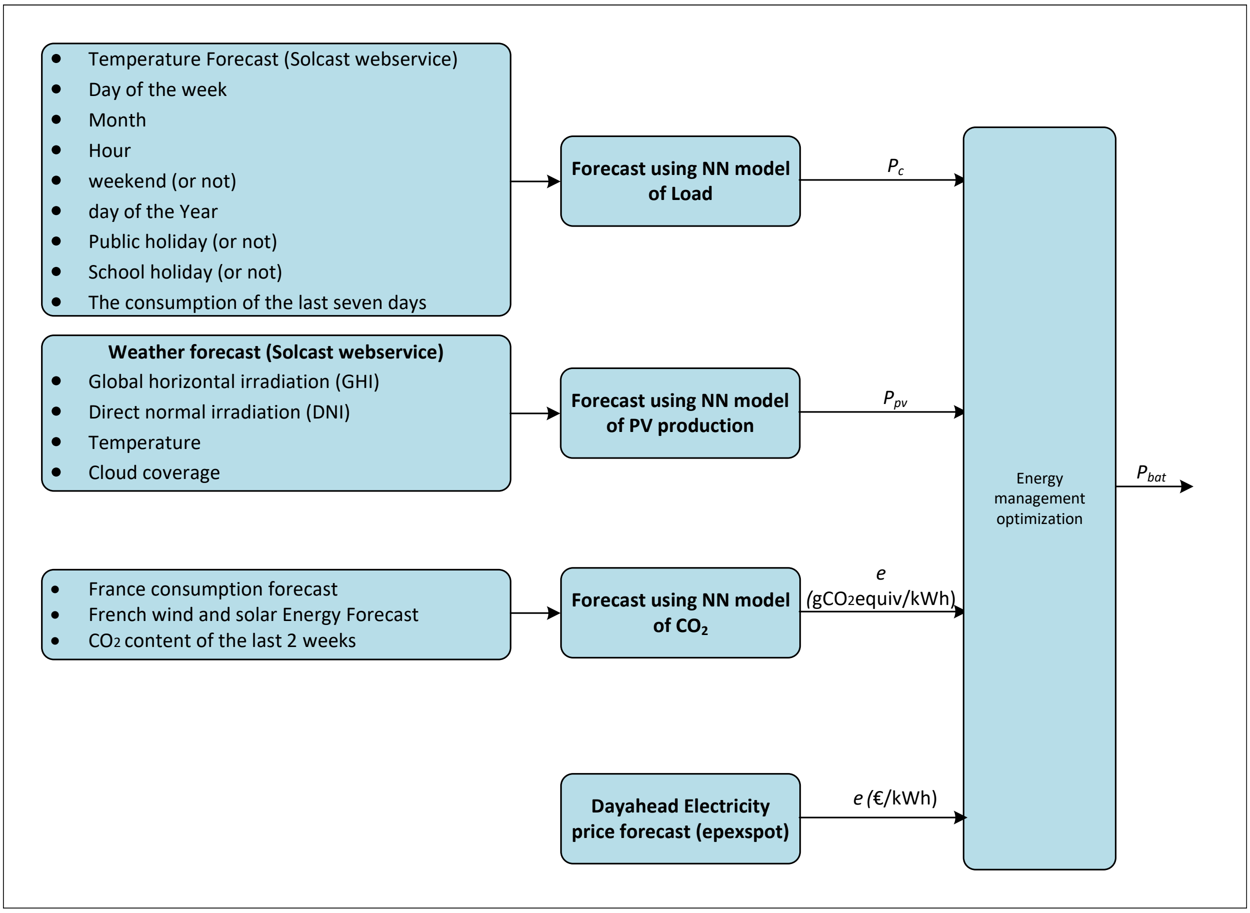

Figure 13: Forecast diagram

Three strategies for the control of the storage are proposed in this study for the day-ahead optimization; economic, ecological and compromise strategies. The smart grid demonstrator is currently subscribed to an off-peak tariff option, which is the basis for economic results in the best compromise and ecological strategies. However, to evaluate the proposed storage control mechanism in a more complex economic tariff, the economic strategy uses the day-ahead SPOT market prices as its cost function. Fig. 14 shows an example of the difference between both tariffs for a typical day in January. Additional constraints could be added if the results of the three strategies cited above do not fulfil the full self-consumption condition. Since the day-ahead optimization is based on PV production, and consumption forecast, a real-time controller is used to verify the full-self consumption condition of the produced energy and take corrective actions if needed. These actions could concern the storage by ignoring the day-ahead optimization and applying a set point to absorb the extra produced power. In some cases the real-time controller could curtail the PV production if the storage has non available capacity. The different strategies of the day-ahead optimization and the associated constraints are presented in the following sections.

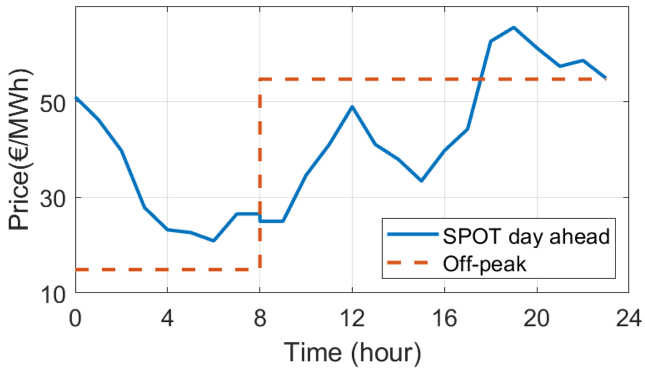

Figure 14: Example of SPOT price variation during a day vs off-peak tariff variation

- **Best compromise strategy**

The best compromise strategy takes advantage of the off-peak tariff option which has a lower electricity price during off-peak hours (mainly at night) and higher costs during the day. It primarily targets economic gains, but also implements an ecological aspect by using the predicted CO_2 per kWh, as the cost coefficient (see equation (2)), to minimize the carbon impact of the facility by charging during the least carbon intensive hours (at night) and discharging during the most carbon intensive hours (during the day) using the following constraints:

- $SOC_{min} \leq SOC \leq SOC_{max}$.
- $SOC = SOC_{max}$ at the end of the charging period (off-peak)
- $0 \leq p_{bat-i} \leq P_{charge,max}$ during the charging period (off-peak hours)
- $0 \geq p_{bat-i} \geq P_{discharge,min}$ during the discharge period (peak hours)
- $SOC = SOC_{min}$ at the end of the discharging period (peak)

Where:

- SOC: State of Charge of the battery in %
- SOC_{min}: Minimum allowable SOC in %
- SOC_{max}: Maximum allowable SOC in %
- $P_{charge,max}$: Maximum charging power in W

– $P_{charge,min}$: Minimum charging power in W

The chosen constraints also apply the physical constraints of the storage as well as limiting the system to a single battery charge/discharge cycle per day (by limiting the total energy) to increase battery lifespan.

- **Ecological strategy**

The ecological strategy implements a reduced number of constraints but is still limited to one full charge/discharge cycle (two times the capacity) per day to preserve battery life, it uses the predicted CO_2 per kWh as the cost coefficient to optimize the gCO_2equiv for the day (see equation (2)) and is subject to the following constraints:

– $SOC_{min} \leq SOC \leq SOC_{max}$

– $P_{discharge,min} \leq p_{bat-i} \leq P_{charge,max}$

– $\sum \frac{|p_{bat-i}|}{6} \leq 2 \times (E_{SOC_{MAX}} - E_{SOC_{MIN}})$

- **Economic strategy**

The economic strategy uses the same constraints as the ecological strategy, with the main difference being the use of the day ahead spot market price (fig. 14) as the cost coefficient (see equation (2)).

- **Self-consumption optimization**

The self-consumption optimization ensures that the smart grid doesn't re-inject power into the distribution grid. It ensures that PV power never exceeds consumption power by either injecting the excess power into the storage system or by curtailing the PV station. To minimize the curtailed (i.e., wasted) power, whenever the forecasts show that at a certain time or for a given period the production will surpass the consumption, additional constraints are added to the optimization for all strategies, to ensure that the storage will have as much capacity as possible to store the excess energy by adding the following constraints:

– $SOC \leq SOC_{E_{SOC_{MAX}} - E_{Curtail}}$ at the beginning of the overproduction period

– $p_{bat-i} = E_{Curtail}/(L_{overprod})$ during the overproduction period

Where:

– $E_{Curtail}$: The excess energy during the overproduction period up to a limit of $E_{max} - E_{min}$ in Wh

– $L_{overprod}$: Length of the overproduction period (unitless)

- **Real-time controller**

The real-time controller checks every 10 seconds if the production become higher than the consumption. Since the injection of PV production surplus to the public grid is not permitted, the storage should store the extra production to avoid a curtailment of the PV production. The battery set-point calculated by the day-ahead optimization will be ignored in this case and the storage set-point will take a the value of the PV production surplus. If the storage is fully charged, the PV inverters will be controlled to curtail the PV production. The total PV production will be set to be equal the electrical consumption. Fig. 15 represents an overview of the energy management algorithm.

4.2.2. Implementation

The implementation of the optimization algorithm was done using the SciPy library in the python programming language and more specifically the scipy.optimize.minimize function. As per equation (2) the goal is to minimize the total cost (ecological or economic) on a day by day basis. To achieve this objective, p_{bat-i} should be optimized according to the forecasts of power consumption, PV production and CO_2 per kWh or energy cost. There are 144 variables to be optimized simultaneously which form the vector P_{bat} (see eq (3)) of length 144 representing the battery power setpoints that correspond to a 10-minute interval during 24 hours. The initial value of the variables is set to

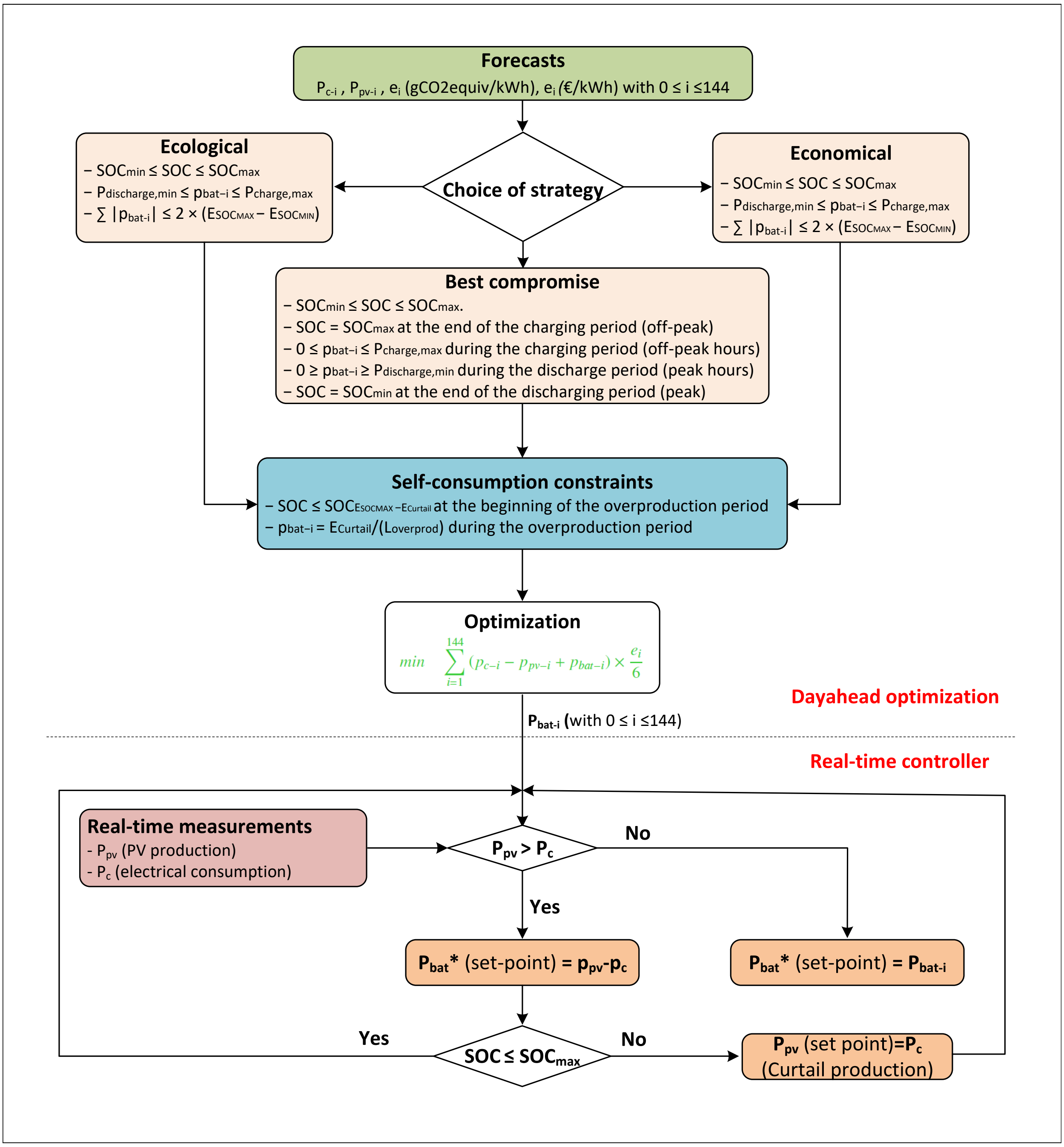

Figure 15: Energy management algorithm diagram

1 kW, SOC_{min} and SOC_{max} are set to 15% and 95% respectively as a compromise between battery lifespan and total usable battery capacity. E_{min} and E_{max} are the energy stored that corresponds to minimum and maximum states of charge, and E_0 is the initial stored energy which is measured before the start of the optimization. As the setpoints p_{bat-i} are in kW per 10-minute interval, when divided by 6 they are expressed in kWh and therefore the energy stored in the battery at a given timestep E_j can be calculated by summing the initial stored energy and the some of the energy of the preceding timesteps multiplied by the total efficiency of the battery (which includes the inverter and depends on the SOC and power setpoint) as per equation (4). The efficiency is found from the characterisation of the battery. Constraints are setup using the equation (4) where 288 inequality constraints are set for each timestep to ensure that the capacity stays within E_{min} and E_{max} as previously described, in addition to other constraints presented in section 4.2.1.

$$P_{bat} = \begin{bmatrix} p_{bat-10} \\ p_{bat-20} \\ \vdots \\ p_{bat-1430} \\ p_{bat-1440} \end{bmatrix} \tag{3}$$

Where: p_{bat-10}, p_{bat-20}, $p_{bat-1430}$, $p_{bat-1440}$ are battery power setpoints at respectively t=10, 20, 1430, 1440 minutes (24 hours).

$$E_j = E_0 + \sum_{i=1}^{j} \frac{p_{bat-i}}{6} \times \eta_{i_{battery}}(SOC_i, p_{bat-i}) \tag{4}$$

5. Results and discussion

The optimisation algorithms is validated in two different ways; first by simulation using one year data set of electrical consumption and PV production, then by an experimental validation on the Catholic university SG demonstrator for 24 hours to validate the dynamic behaviour of the system.

5.1. Simulation results

The proposed storage management method is validated using historical data for the period of 1/3/2019 to 1/3/2020. Two key performance indicators (KPI) are used to compare the results of different strategies; economic indicator and ecological indicator. The economic gain and the ecological gain are directly related to ratio between the battery capacity size and the total daily consumption. For the case of the Catholic university SG demonstrator, the installed battery capacity represents only 6% of the total daily energy consumption. Therefore, relative comparisons provide a more accurate view of the differences between each strategy regardless the battery capacity size. The economic indicator is defined by the decrease in the total cost of electricity. For a full economic strategy, this indicator takes the value of 100% and for the case with no storage used this value become 0. By the same, the ecological indicator is defined by the decrease in the total amount of CO_2 produced from electricity. It takes the value of 100% for a full ecological strategy and 0 for the case with no storage used. A “simple” energy management system is implemented with predefined daily setpoints for charging during off-peak hours and discharging during peak hours to better test the proposed management system’s performance by providing a more appropriate comparative perspective. The storage management method uses forecasts of electricity consumption, photovoltaic production and CO_2 rate or day ahead SPOT market prices to generate power set points for the storage unit daily. The real values are used to calculate the simulation results to compare the different strategies.

5.1.1. Simulation using off-peak tariff

Fig. 16 shows the average relative economic and ecological performance for the simple, ecological, and best compromise strategies using off-peak tariff, The economic gain is limited by the off-peak tariff, it cannot be further increased without increasing the capacity of the storage, which is why the best compromise and simple strategies offer the same economic gain. The ecological gain is where the differences between the strategies emerge. As the best compromise strategy demonstrates its advantages over the simple strategy, the ecological strategy’s gains are a trade-off with the economic gains when compared to the other strategies. Fig. 17 shows a sample of some of the results from the best compromise strategy over a 2-day period; from the 7th of March 2019 to midnight 9th of March 2019.

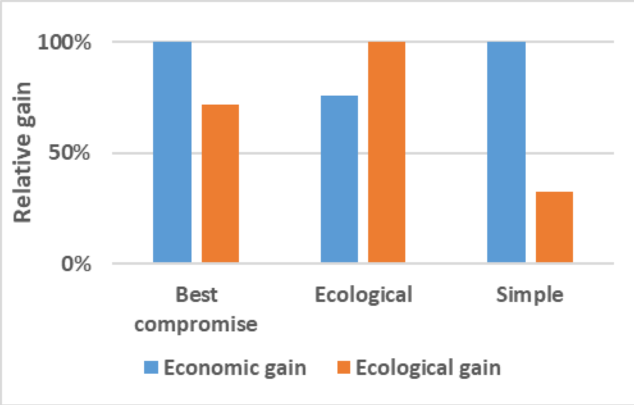

Figure 16: Relative economic and ecological gains per strategy using off-peak tariff

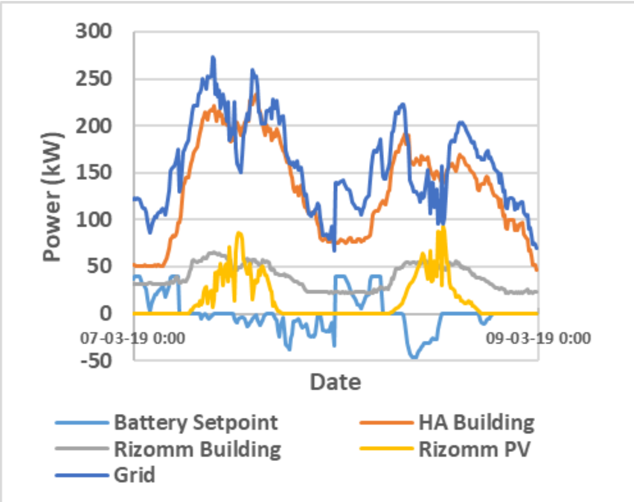

Figure 17: Battery setpoint power profile, HA building power profile, Rizomm building power profile, Rizomm PV power profile and grid consumption power profile for best compromise scenario over a 2-day period

5.1.2. Simulation using spot market prices

The results obtained with the spot market price show the best economic gain in comparing to other strategies. A comparison of the economic gain between the economic and simple strategies is presented in Fig. 18. They show the utility of the proposed management method in more complex tariff scenarios compared to the simple scenario, the latter being adequate with off-peak tariffs but offering limited gains when the price varies hourly. Finally, with regards to the self-consumption constraints, without the presence of the storage, 4481 kWh must be curtailed which represents 3.19% of the total yearly production. The proposed control method manages to reduce the curtailed energy to 20 kWh. This curtailed energy is because of high production during a sunny day in June with low consumption which saturated the storage and lead to curtailing the excess energy.

5.2. Experimental validation of real time controller on the Catholic university demonstrator

In order to test the real time controller outlined in section 4.2.1, the algorithm was implemented on the EMS of the smart grid demonstrator (Fig. 2). For security reasons, the following parameters were adopted in the storage

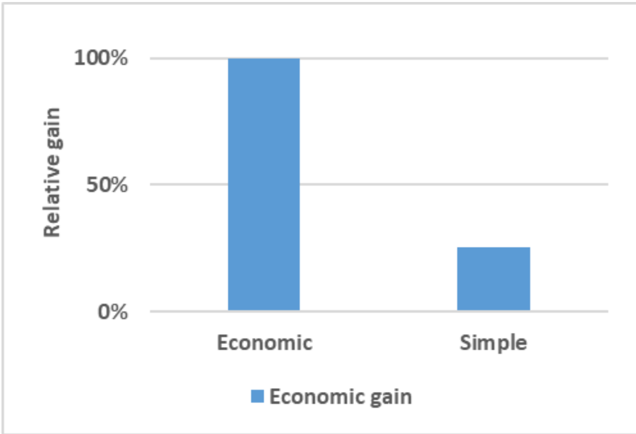

Figure 18: Relative economic gain per strategy using day ahead spot market prices

controller:

- One string of the storage system with a capacity of 50kWh is used.
- The maximum power is limited to 20kW in discharging and 25kW in charging.
- The consumption of only Rizomm buidling is considered in order to reach the worst case scenario (electrical production will surpass consumption for a long period of time) for the real time controller and test the limits of the algorithm.
- SOC_{max} is set equal to 80% ($E_{max} = 40kWh$) and SOC_{min} is set to 30% ($E_{min} = 15kWh$).
- The chosen strategy is ”Best compromise” with full self consumption constraints checking.

The experiment was conducted for 24 hours and started at midnight of Saturday 15th of October 2022. This date was chosen based on day-ahead weather forecast to have a sunny day with good PV production and to have low power consumption and less number of students in the campus on the other hand. Figure 19 shows the results of real time controller test, where the photovoltaic and consumption forecasts predicted a period during which production will surpass consumption, the management method accounts for this by discharging the battery to a specific SOC at the beginning of the overproduction period (as described in 4.2.1).The figure is split into four sections:

- The first section where the battery can only be charged since it is during the off-peak hours and overproduction cannot occur since it is at night, noting that since the chosen strategy is ”Best compromise” the objective here is to charge during the least carbon intensive hours (Fig. 20). Charging stops at 78% due to the small inaccuracies in the theoretical SOC (using characterized efficiency), and falls slightly to 77% due to the re-balancing of the battery cells by the internal battery management system (BMS).
- The second section is where the battery can only be discharged since the net power is still positive (i.e. consumption is higher than production) as per the predictions, and the SOC goes from 77% to the minimum of 30% since the excess energy during the overproduction period was predicted to be higher than the current battery useful capacity (25kWh). The real time controller takes over and charges the battery, in order to ensure the self consumption constraint, near the end of this section as the net power becomes negative for some brief periods, which is due to the forecast errors mentioned previously in section 4.1. It should be noted that real time control can utilize the maximum charging power of the battery and is not limited by the upper and lower bounds shown on the graph as they only limit the day ahead planning.
- The third section is the predicted overproduction period where real time controller takes full control of the battery in order to ensure the full self consumption constraint. The lower and upper bounds are equal to the

curtailment energy over the length of the period as previously mentioned for real time controller constraints in section 4.2.1, these bounds are necessary in order to calculate an accurate SOC after the overproduction period. The battery reaches SOC_{max} near the middle of the period and the real time controller must now curtail the power production by directly communicating with the PV inverters in order to limit their power. This is not shown on the figure for clarity purposes.

- The fourth and final section is the final discharge period, where the battery which is now fully charged must discharge at the most carbon intensive hours (Fig. 20) in preparation for the next day's off-peak charging. Battery cell re-balancing causes the small increase of SOC near the end of the section.

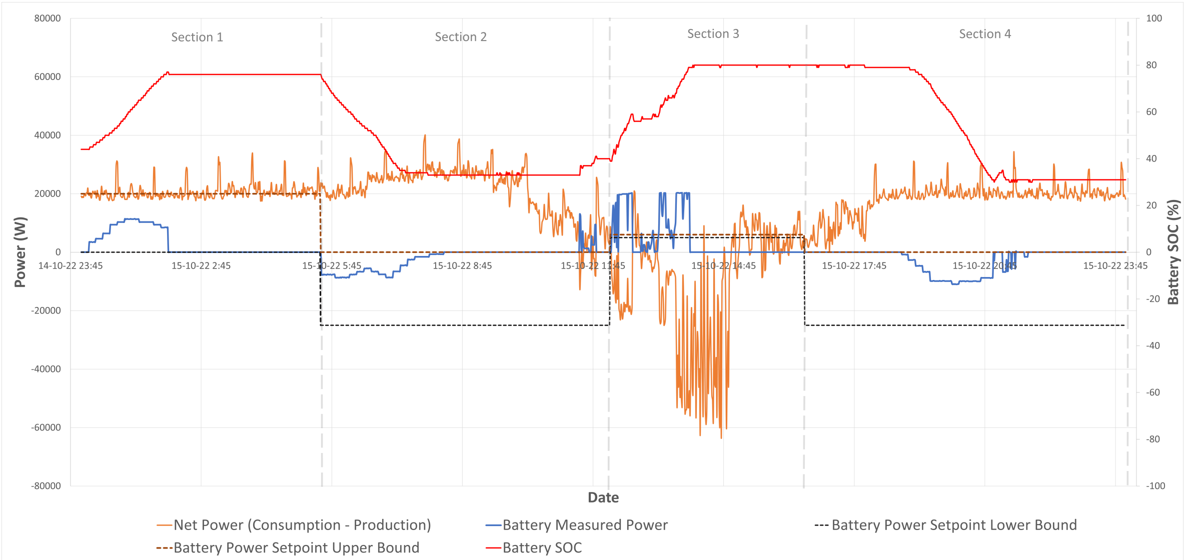

Figure 19: Battery setpoint power profile, net power, battery setpoint lower and upper bounds and the battery SOC for the 15th of October 2022, four sections outline the different stages based on net power, time and bounds

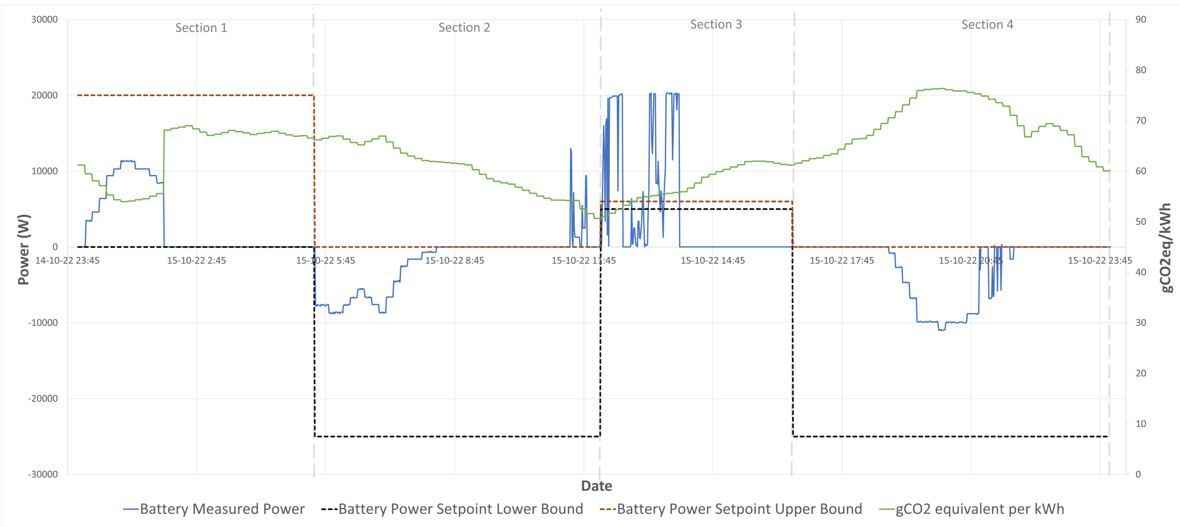

Figure 20: Battery setpoint power profile, grams of CO2 equivalent per kWh, battery setpoint lower and upper bounds for the 15th of October 2022, four sections outline the different stages based on time and bounds

The results show that the real time controller test was a success, with the management method accounting for the overproduction period as well as it possibly could, while using the forecasts with slight inaccuracies due to the efficiency calculations of the battery. Possible improvements could be accounting for forecast errors by varying the length of the overproduction period and more accurate efficiency curve for the battery and inverter system. Noting that the efficiency characterization must be done periodically as the battery ages.

6. Conclusion

In conclusion, a novel approach to the energy management of an on-grid photovoltaic (PV) system with batteries in the context of individual self-consumption has been presented in this study. The proposed method, which uses multiple criteria (economic, ecological, and a compromise between the two) in different tariff scenarios, respects a full self-consumption constraint imposed by the distribution system operator. The originality of this work is found in the application of machine learning algorithms for forecasting electrical consumption, PV production, and CO2 emissions, and in the use of an optimization model for energy management. The forecast models were developed using feedforward neural network models based on data from a real-scale smart-grid demonstrator at the Catholic University of Lille, France. Advantages of this approach include the ability to handle non-linear data and learn complex patterns, which are inherent in electrical consumption, production, and CO2 emissions data. The use of machine learning algorithms for forecasting allows for more accurate predictions, which in turn leads to more efficient energy management. The optimization model enables the system to minimize the economic or ecological cost while respecting the full self-consumption constraint. However, some limitations are also associated with this approach. The performance of the forecast models is dependent on the quality and quantity of the available data. In situations where data is limited or noisy, the accuracy of the forecasts may be compromised. Furthermore, the optimization model assumes perfect forecasts, which is not always the case in real-world scenarios. Future work could focus on incorporating forecast uncertainties into the optimization model. Despite these limitations, it is believed that this approach provides a robust and effective solution for the energy management of on-grid PV systems with batteries. The results from the real-scale demonstrator at the Catholic University of Lille, France, validate the effectiveness of this approach and demonstrate its potential for practical application. This work is hoped to contribute to the ongoing efforts to optimize the integration and management of renewable energy sources into the power grid, and further advancements in this field are eagerly anticipated.

Acknowledgements

Authors want to thank the European city of Lille (MEL) for the financial support under So Mel so connected project as well as the Lille catholic university, JUNIA graduate school of science and engineering and the Hauts-de-France french region.

References

[1] Enedis, Plus de 100 000 opérations d'autoconsommation en france métropolitaine, 2021. URL: https://www.enedis.fr/presse/plus-de-100-000-operations-dautoconsommation-en-france-metropolitaine.

[2] Enedis, Répartition des opérations d'autoconsommation collective actives – maille enedis, 2022. URL: https://data.enedis.fr/explore/dataset/autoconsommation-collective-maille-enedis/analyze/.

[3] B. Robyns, A. Davigny, C. Saudemont, Methodologies for supervision of hybrid energy sources based on storage systems – a survey, Mathematics and Computers in Simulation 91 (2013) 52–71. doi:https://doi.org/10.1016/j.matcom.2012.06.014.

[4] B. Robyns, B. François, G. Delille, C. Saudemont, Energy storage in electric power grids, ISTE Editions, 2015.

[5] V. Courtecuisse, J. Sprooten, B. Robyns, M. Petit, B. Francois, J. Deuse, A methodology to design a fuzzy logic based supervision of hybrid renewable energy systems, Mathematics and Computers in Simulation 81 (2010) 208–224. doi:https://doi.org/10.1016/j.matcom.2010.03.003.

[6] H. Zhang, A. Davigny, F. Colas, Y. Poste, B. Robyns, Fuzzy logic based energy management strategy for commercial buildings integrating photovoltaic and storage systems, Energy and Buildings 54 (2012) 196–206. doi:https://doi.org/10.1016/j.enbuild.2012.07.022.

[7] S.-V. Oprea, A. Bâra, Ştefan Preda, O. Tor, A smart adaptive switching module architecture using fuzzy logic for an efficient integration of renewable energy sources. a case study of a res system located in hulubeşti, romania, Sustainability 12 (2020) 6084. doi:10.3390/su12156084.

[8] N. Mbungu, A. Ismail, M. AlShabi, R. Bansal, A. Elnady, A. Hamid, Control and estimation techniques applied to smart microgrids: A review, Renewable and Sustainable Energy Reviews 179 (2023-06-01) 113251. doi:10.1016/j.rser.2023.113251.

[9] G. Thirunavukkarasu, M. Seyedmahmoudian, E. Jamei, B. Horan, S. Mekhilef, Role of optimization techniques in microgrid energy management systems—a review, Energy Strategy Reviews 43 (2022-09-01) 100899. doi:10.1016/j.esr.2022.100899.

[10] M. Abdel-Nasser, K. Mahmoud, Accurate photovoltaic power forecasting models using deep lstm-rnn, Neural Computing and Applications 31 (2019-07-01) 2727–2740. doi:10.1007/s00521-017-3225-z.
[11] M. Hossain, S. Mekhilef, M. Danesh, L. Olatomiwa, S. Shamshirband, Application of extreme learning machine for short term output power forecasting of three grid-connected pv systems, Journal of Cleaner Production 167 (2017-11-20) 395–405. doi:10.1016/j.jclepro.2017.08.081.
[12] V. Oprea, A. Bâra, Ultra-short-term forecasting for photovoltaic power plants and real-time key performance indicators analysis with big data solutions. two case studies - pv agigea and pv giurgiu located in romania, Computers in Industry 120 (2020-09-01) 103230. doi:10.1016/j.compind.2020.103230.
[13] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artificial Intelligence Review 54 (2020) 1937–1967. doi:10.1007/s10462-020-09896-5.
[14] G. Biau, E. Scornet, A random forest guided tour, Test 25 (2016) 197–227.
[15] N. Kriegeskorte, T. Golan, Neural network models and deep learning, Current Biology 29 (2019) R231–R236.
[16] Solcast, Global solar irradiance data and pv system power output data, 2019. URL: `https://solcast.com/`.
[17] M. Abadi, A. Agarwal, et al., Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2015. Software available from tensorflow.org.
[18] M. Ali, PyCaret: An open source, low-code machine learning library in Python, 2020. PyCaret version 1.0.0.
[19] S. Schlömer, T. Bruckner, et al., Annex iii: Technology-specific cost and performance parameters, Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (2014) 1329–1356.
[20] D. Alasdair, R. W. Bruce, L. Ruffa, J. Kellowaya, F. MacMillana, P. A. Rogers, Carbon intensity - forecast methodology, 2021.
[21] RTE, Données éco2mix nationales consolidées et définitives (janvier 2012 à mai 2022), 2022. URL: `https://opendata.reseaux-energies.fr/explore/dataset/eco2mix-national-cons-def/information/`.