Machine Learning Algorithms and PV Forecast for Off-Grid Prosumers Energy Management

Simona-Vasilica Oprea
Adela Bâra
Bucharest University of Economic Studies, Romania
simona.oprea@csie.ase.ro
bara.adela@ie.ase.ro

Abstract

The actual context characterized by the high prices of the conventional power gives more and more credit to the Renewable Energy Sources (RES) to cover load requirements in large amounts. However, the volatility of RES (especially solar and wind) restricts their smooth integration into the residential consumption energy mix. One of the main challenges is to maximize the consumption of appliances from RES taking into account their availability. To fulfil this objective, first, a performant forecast is necessary to create the day-ahead schedule and optimize the operation of appliances. In this paper, we propose a framework to perform PV forecast with machine learning algorithms and various data sources for the energy management of the off-grid prosumers.

Key words: renewable energy sources (RES); maximizing consumption from RES; day-ahead forecast; machine learning; prosumers
J.E.L. classification: papers P28, Q47, Q29

1. Introduction

Pollution, along with global warming, greenhouse gases and heat waves are the worst systemic threat to human civilization according to the Organization United Nations. In the last two decades, strong support has been given to the large-scale renewables-based power plants, but the European Union (EU) targets for 2030 to reduce the CO2 emissions (by 45%) and integrate RES (32%) are more ambitious, leading the policy makers to empower active consumers (prosumers) and citizen energy communities (COM) beyond example. Yet, the electricity costs are constantly increasing and may range between 10-12% in the income of families without self-generation. Especially in rural areas where disconnections are more frequent, self-generation (photovoltaics-PV, Wind Turbines-WT) and storage devices represent an alternative to reduce the electricity costs and increase the consumers' comfort and secure operation. The necessity of additional research is underlined by the EU Clean Energy Package issued in June 2019 that encourages decentralization of generation and brings indisputable economic and environmental advantages, incentivizing consumers and COM to self-generate from RES. The EU Package has provided the starting point for a new shape of the power systems that are moving to a more decentralized structure. It sets targets for the next decades and empowers prosumers in transition to a sustainable low-carbon environment. Thus, prosumers are expected to reconfigure the path of the power system development. They are now able to consume, generate, store, share and sell electricity to the electricity markets, settle the imbalance, and offer flexibility to the power system as the ICT and other technology progress facilitate these activities.

2. Literature review

The prosumers usually have the possibility to generate and consume (off-grid) or consume or inject into the grid (on-grid). For the on-grid prosumers, the surplus of energy is injected into the grid, whereas for off-grid prosumers, the surplus is either stored or the PV panels stop generating and the energy is lost. Thus, it is more difficult for off-grid prosumers to predict their generation as sometimes the generation is interrupted and unknown not due to the lack of weather favorable conditions, but due to the lack of local load. Another option would be to generate and locally share the energy with the neighbors (Oprea and Bâra, 2021), (Oprea and Bâra, 2020), but in many countries the infrastructure and market platforms are missing.

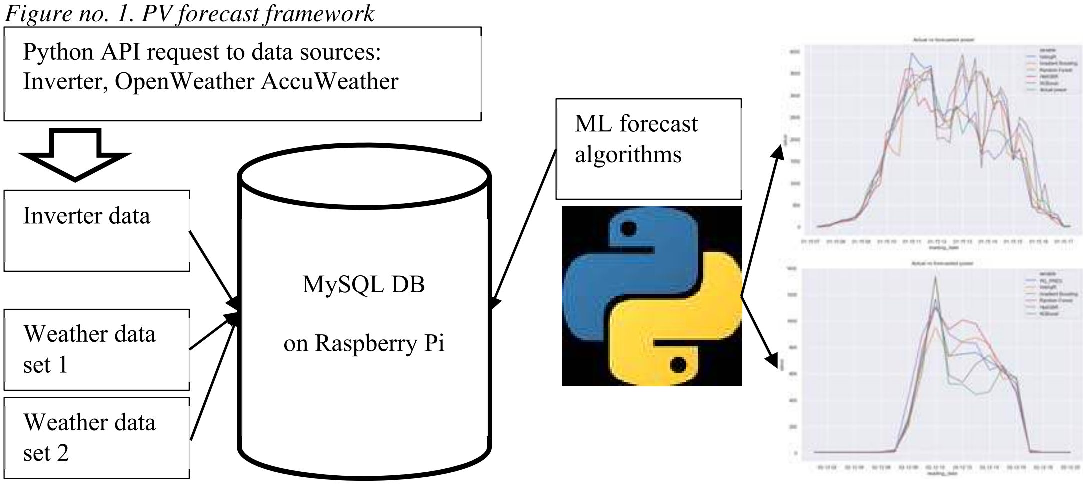
In any case, the PV forecast is necessary to correctly schedule the operation of programmable appliances such as water heater, electric heating system, storage facilities and batteries for the battery-operated appliances. The control of the appliances is performed by means of smart plugs that can be scheduled to be on or off. Thus, especially for off-grid prosumers, it is very important to use the RES when they are available energizing the appliances and storing the energy in storage facilities, battery-based appliances, including EV batteries, water heaters, heating, etc., otherwise the green energy is lost. An accurate forecast provides savings and maximizes the usage of RES leading to CO2, deforestation and standard coal reductions improving the environment. Savings come from the optimal schedule of the appliances. For instance, the water heater can operate at noon when the PV panels usually generate more and store the warm water in the water tank, not in the evening when the warm water is used.

Thus, in this paper, we propose a framework with various data sources and Machine Learning (ML) algorithms to perform PV generation forecast and improve the energy management performance at the prosumers' level. Day-ahead probabilistic PV forecast for buildings in Germany based on historic and weather data was proposed in (El-Baz, Tzscheutschler and Wagner, 2018). A comparative study between parametric using Matlab modelling tool and non-parametric (in R) PV forecast is presented in (Almeida *et al.*, 2017) showing that the parametric forecast displayed more significant bias. However, the forecast can be performed as input data for other objectives such as maximizing the prosumers' profits (Panapakidis, Koltsaklis and Christoforidis, 2021), (Agathokleous, Tuan and Steen, 2019). In (Giordano *et al.*, 2020), an interesting approach for the scheduling batteries to optimally match local RES, loads, including storage is envisioned. PV generation is forecasted using statistic algorithms and clustering techniques.

3. Research methodology

The proposed framework consists in storing the invertor and weather data into tables T_METER_READINGS (for invertor) and T_WEATHER_READINGS (for weather data provided via websites APIs) into an open database such as MySQL that can be installed on Raspberry Pi. In Python, data is read and processed to perform forecast for the next 24 hours. The 30-days ago generated power recorded at 5 minutes is stored into T_METER_READINGS table, whereas the hourly weather data from two websites (OpenWeather https://openweathermap.org/ and AccuWeather https://www.accuweather.com/) is stored into T_WEATHER_READINGS table as .json files as in Figure 1.

Figure no. 1. PV forecast framework



Source: Authors' contribution

The data is read in Python as local variables and processed using join, label encoders, interpolation, scaling the input data with MinMaxScaler() and StandardScaler() and building new features with group by functions and SQL analytic functions.

Interpolation is used to both eliminate null values and mediate the differences in readings. Feature selection aims to identify the most relevant features that are involved in training the model. Then, ten ML algorithms are proposed to train the model: Linear Regression, Ridge, Stochastic Gradient Descent Regressor, Multi-layer Perceptron Regressor, Gradient Boosting Regressor, Random Forest Regressor, XGB Regressor, HistGradientBoosting Regressor, Decision Tree, Voting Regressor. The hyperparameters of the ten algorithms are presented in Table 1. Also, a dataframe acc_df was created to store the accuracy of the models. In practice, the last five algorithms proved to perform better.

Table no. 1. ML algorithms, hyperparameters and metrics

ML algorithms, hyperparameters and metrics
acc_df = pd.DataFrame(columns = ['Model','Model name','Train Score', 'Test Score', 'R2_Test']) models.append(("Linear Regression:",linear_model.**LinearRegression**())) models.append(("Ridge Regression:",linear_model.**Ridge**(alpha=0.5))) models.append(("SGD Regression:",linear_model.**SGDRegressor**(alpha=1,max_iter=1000,learning_rate='optimal'))) models.append(("MLP:",**MLPRegressor**(solver='sgd', activation='relu', hidden_layer_sizes=(70,70), max_iter=1000,learning_rate='adaptive', alpha=0.01, learning_rate_init=.01, momentum=0.95, nesterovs_momentum=True))) models.append(("Gradient Boosting:",**GradientBoostingRegressor**(n_estimators=200, criterion='squared_error',learning_rate=0.05,max_depth=5, alpha=0.03, loss='squared_error'))) models.append(("Random Forest:", **RandomForestRegressor**(n_estimators = 200, criterion='squared_error',max_depth=10))) models.append(("XGBoost:", **XGBRegressor**(n_estimators=200, n_jobs=8, random_state = 10, max_depth=5,reg_lambda= 0.01, learning_rate=0.05))) models.append(("HistGBR:", **HistGradientBoostingRegressor**(random_state = 10, max_depth=10,learning_rate=0.08, l2_regularization=0.05, max_iter=100))) models.append(("DT:",tree.**DecisionTreeRegressor**())) reg1=**GradientBoostingRegressor**(n_estimators=200, criterion='squared_error',learning_rate=0.05,max_depth=5, alpha=0.03, loss='squared_error') reg2=**RandomForestRegressor**(n_estimators = 200, criterion='squared_error',max_depth=10) reg3=**XGBRegressor**(n_estimators=200, n_jobs=8, random_state = 10, max_depth=5,reg_lambda= 0.01, learning_rate=0.05) reg4= **HistGradientBoostingRegressor**(random_state = 10 ,max_depth=10,learning_rate=0.08, l2_regularization=0.05, max_iter=100) models.append(("VotingR:",**VotingRegressor**(estimators=[('gb', reg1), ('rf', reg2), ('xgb', reg3), ('hist', reg4)])))

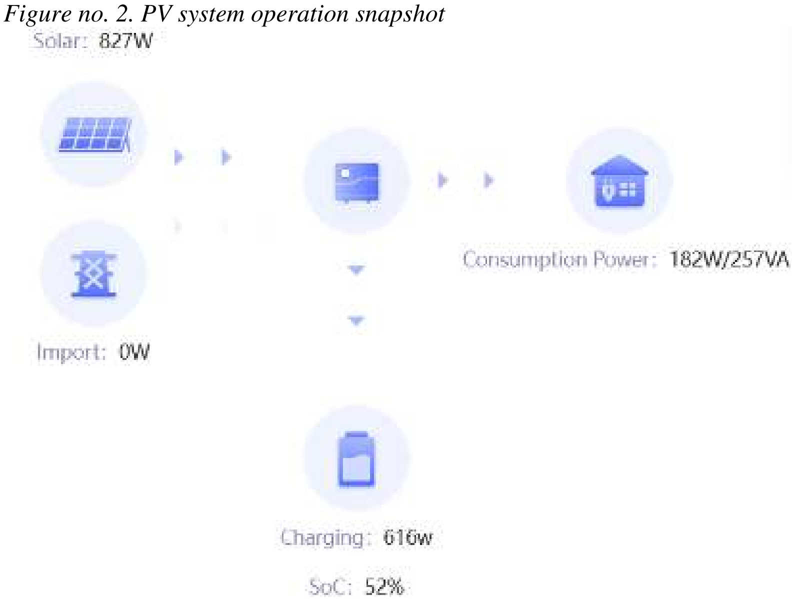
Source: Authors' contribution

The results are graphically displayed, and the metrics are stored and can be easily compared.

4. Findings

The PV system consists in self-generation of 5.94 kWp PV capacity connected through a off-grid invertor and 8.3 kWh LiFePO4 batteries. A snapshot of the flows is extracted from the application of an inverter as in Figure 2. The PV system generates 827 W that are consumed by the appliances and storage facility. The input from the grid is zero. The State of Charge (SoC) of the storage facility is 52%.

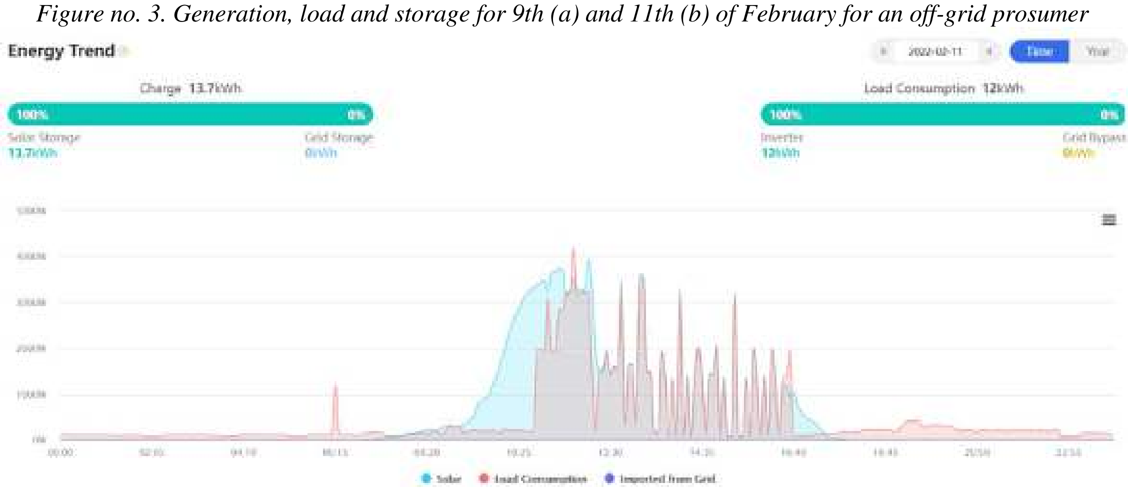
Figure no. 2. PV system operation snapshot



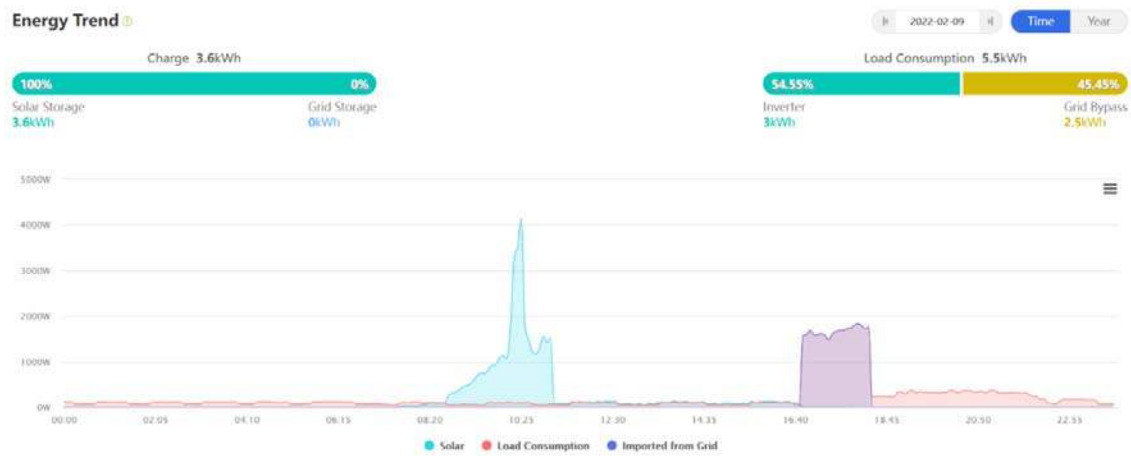
Source: Authors' contribution

To understand the off-grid prosumers, two diagrams are presented in Figure 3 (a) and (b). The first one shows the operation of various appliances as the load is varying from 06:15 in the morning until 16:40 afternoon when the PV generation become lower. The second day shows a much smaller activity up to 16:40 when the appliances were fed from the grid. However, in the second day, the solar energy was lower and the grid input was required to energize an appliance of 1800 W (water heater). This scenario indicates that the consumers were not at home during the day. These charts were extracted in February (9th and 11th) from the inverter application.

Figure no. 3. Generation, load and storage for 9th (a) and 11th (b) of February for an off-grid prosumer



(a)

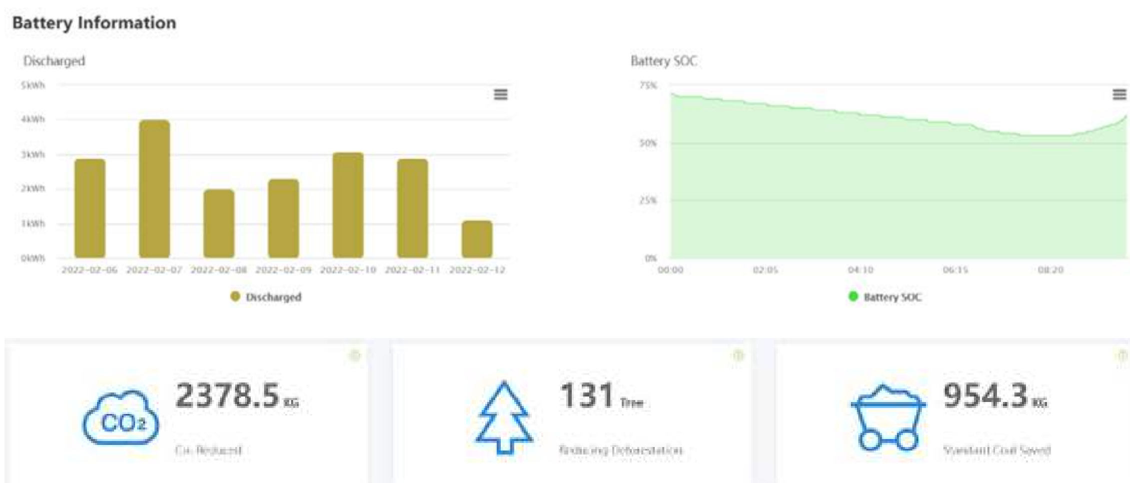


(b)

Source: Authors' contribution

The application of inverter also offers details regarding battery and metrics related to environment as in Figure 4, that can further motivate the prosumer to improve his energy management.

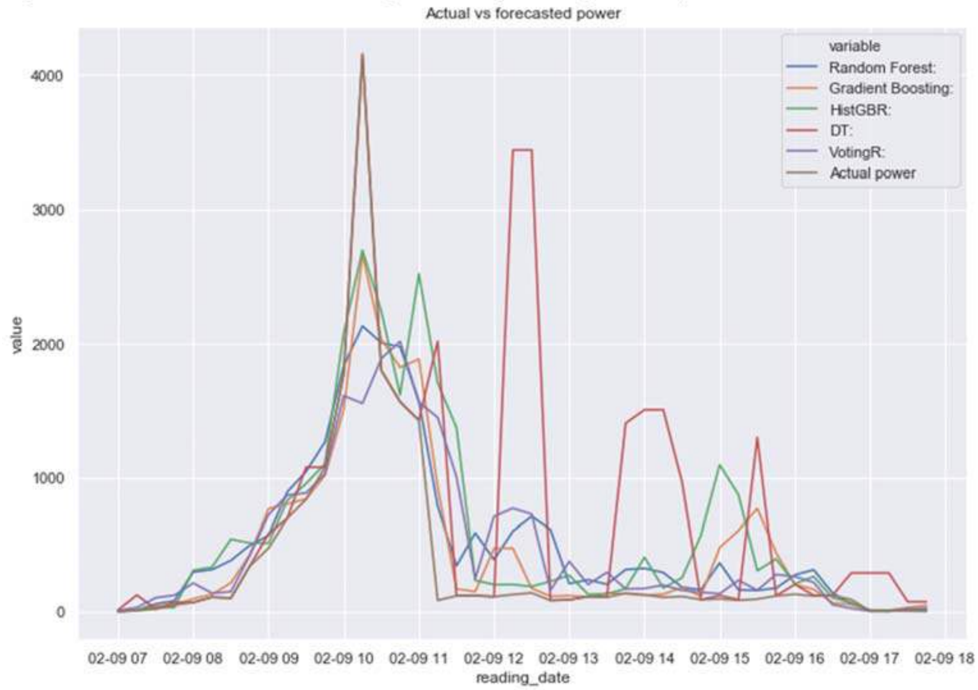
Figure no. 4. Battery info and pro-environmental metrics



Source: Authors' contribution

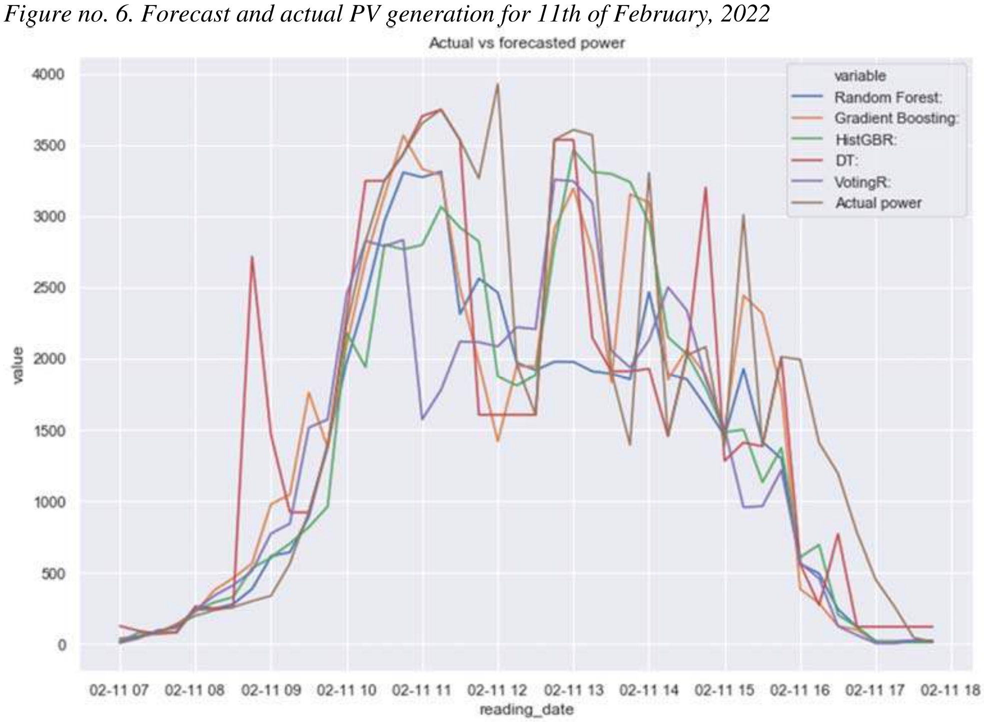
The PV forecast and actual generation for the two days are displayed in Figures 5 and 6.

Figure no. 5. Forecast and actual PV generation for 9th of February, 2022



Source: Authors’ contribution

Figure no. 6. Forecast and actual PV generation for 11th of February, 2022



Source: Authors’ contribution

The graphical results show a good approximation of the actual PV generation. On average, during one month (February) the RMSE is 0.92 and R square is 0.97.

5. Conclusions

Prosumers are not a new concept, they appeared long before the energy transition era where the access to the grid was not possible or too costly. They are heavily associated to the difficulties related to the grid access, environmental concerns driven by oil prices and nuclear dangerous potential that have started in the '60s. It was also an attitude of reducing dependencies and centralized control of the power systems. The expansion of self-generation after the '90s reveals a strong correlation to policies. More and more people envision to be less dependent on a centralized system that sometimes fails to fulfil their necessities. After the '90s, prosumers have become more numerous driven by cheaper and reliable technology, reduced payback period, and increased awareness regarding solar potential. Prosumers will speed up the transition to an economy based on clean energy from RES, especially solar as PV systems cost and efficiency have improved significantly over time. They are becoming cheaper, increasing conversion efficiency, hence producing more energy. Their lifetime also increased to 25 years. Moreover, solar energy can become a solution not only for the environment but also for vulnerable consumers or those who do not have access to the grid. Nowadays, PV systems are able to energize a large range of devices from mobile phone batteries to cars, boats, trains to airplanes. Storage-based modern appliances such as mowers, chainsaw, vacuum cleaners, etc. can be energized when solar energy is available and be used anytime. However, prosumers are still far from being mature and showing their full potential. Multiple barriers, including uncertainties regarding benefits, that impede the development and growth of prosumers have been identified. More research is necessary to reveal their potential, provide solutions and quantify their impact from the economic, environmental and social point of view.

The importance of RES is increasing in the actual context of conventional sources scarcity and high prices that may lead to higher bills. Therefore, in this paper we proposed a framework of various sources integration and ML algorithm to forecast the output of a PV system offering encouraging results for off-grid prosumers. We expect that this framework provides better results for on-grid prosumers as their predictability is higher as it does not depend on the load. As future study, we will implement the proposed framework for on-grid prosumers and further test the algorithms. Additionally, this forecast will be integrated into a more complex application that will include optimal schedule of the appliances and real-time operation.

6. Acknowledgement

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS- UEFISCDI, project number PN-III-P4-PCE-2021-0334, within PNCDI III.

7. References

- Agathokleous, C., Tuan, L. A. and Steen, D., 2019. Stochastic operation scheduling model for a Swedish prosumer with PV and BESS in Nordic day-ahead electricity market. *2019 IEEE Milan PowerTech, PowerTech 2019*. doi: 10.1109/PTC.2019.8810651.
- Almeida, M. P. et al., 2017. Comparative study of PV power forecast using parametric and nonparametric PV models. *Solar Energy*. doi: 10.1016/j.solener.2017.07.032.
- El-Baz, W., Tzscheutschler, P. and Wagner, U., 2018. Day-ahead probabilistic PV generation forecast for buildings energy management systems. *Solar Energy*. doi: 10.1016/j.solener.2018.06.100.
- Giordano, F. *et al.*, 2020. Vehicle-to-Home Usage Scenarios for Self-Consumption Improvement of a Residential Prosumer with Photovoltaic Roof. *IEEE Transactions on Industry Applications*. doi: 10.1109/TIA.2020.2978047.
- Oprea, S. V. and Bâra, A., 2020. Local market mechanisms survey for peer-to-peer electricity trading on blockchain platform. *Scientific Bulletin of Naval Academy*. doi: 10.21279/1454-864X-20-I1-024.
- Oprea, S. V. and Bâra, A., 2021. Devising a trading mechanism with a joint price adjustment for local electricity markets using blockchain. Insights for policy makers. *Energy Policy*. doi: 10.1016/j.enpol.2021.112237.
- Panapakidis, I. P., Koltsaklis, N. E. and Christoforidis, G. C., 2021. Forecasting Methods to Support the Decision Framework of Prosumers in Deregulated Markets. *Proceedings of 2021 9th International Conference on Modern Power Systems, MPS 2021*. doi: 10.1109/MPS52805.2021.9492725.