

---

Article

Not peer-reviewed version

---

# Using Deep Neural Network methods for forecasting energy productivity based on comparison of simulation and DNN results for central Poland - Świętokrzyskie Voivodeship

---

Michał Pikus \* and Jarosław Wąs

Posted Date: 18 August 2023

doi: 10.20944/preprints202308.1371.v1

Keywords: AI; Forecasting; RES



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# Using Deep Neural Network Methods for Forecasting Energy Productivity Based on Comparison of Simulation and DNN Results for Central Poland—Swietokrzyskie Voivodeship

Michał Pikus <sup>1,\*</sup>, Jarosław Wąs <sup>2</sup>

<sup>1</sup> AGH University of Science and Technology, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Department of Applied Computer Science, al. Mickiewicza 30, 30-059, Krakow, Poland; mpikus@agh.edu.pl

<sup>2</sup> University of Science and Technology, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Department of Applied Computer Science, al. Mickiewicza 30, 30-059, Krakow, Poland; jaroslaw.was@agh.edu

\* Correspondence: mpikus@agh.edu.pl

**Abstract:** Forecasting electricity demand is of utmost importance for ensuring the stability of the entire energy sector. However, predicting the future electricity demand and its value poses a formidable challenge due to the intricate nature of the processes influenced by renewable energy sources. Within this piece, we have meticulously explored the efficacy of fundamental deep-learning models designed for electricity forecasting. Among the deep learning models, we have innovatively crafted recursive neural networks (RNNs) predominantly based on LSTM and combined architectures. The data-set employed was procured from a SolarEdge designer. The data-set encompasses daily records spanning the past year, encompassing an exhaustive collection of parameters extracted from solar farm (based on location in Central Europe (Poland Swietokrzyskie Voivodeship)). The experimental findings unequivocally demonstrated the exceptional superiority of the LSTM models over other counterparts concerning forecasting accuracy. Consequently, we compared multilayer DNN architectures with results provided by the simulator.

**Keywords:** AI; Forecasting; RES

---

## 1. Introduction

Poland's photovoltaic (PV) sector has been witnessing a consistent upswing, with various factors contributing to the advancement of photovoltaic technology in the nation. Similarly to many other European Union (EU) member states, Poland has established ambitious renewable energy targets to bolster the proportion of renewables in its energy mix. Consequently, this impetus has triggered substantial investments in solar energy and PV technology. The government has proactively introduced diverse support mechanisms, such as feed-in tariffs, auctions, and subsidies, to foster the growth of renewable energy initiatives, including solar PV installations. Being an EU member state, Poland enjoys access to funding programs that facilitate the transition to renewable energy sources, which has further accelerated the development of photovoltaic projects in the country. Encouragingly, the cost of PV technology has been consistently declining over the years, rendering solar energy increasingly economically viable and appealing to both investors and consumers alike. Simultaneously, there has been a concerted effort to reduce dependence on imported fossil fuels, leading to a notable emphasis on fostering domestic renewable energy sources, including solar power. Notably, investments in PV technology have enabled Poland to diversify its energy sources, thereby significantly bolstering energy security. A remarkable milestone occurred between 2019 and 2022, as over one million new PV micro-installations were established, primarily focusing on single-family households. These PV micro-installations are thoughtfully tailored to align with specific energy consumption patterns,

spanning distinct periods, such as a year, adhering dutifully to prosumer rules within Poland's regulatory framework.

The principal aim of this study centers around evaluating the feasibility and precision of employing PV energy production calculation methods under Polish conditions. Researchers have observed that presently utilized methods for calculating energy production may exhibit inaccuracies, stemming from inadequate adaptation to terrain, location, wind conditions, and the representative characteristics of PV panels within the installation. In this research, the application of Deep Neural Network (DNN) methods will be explored to verify the accuracy of data provided by SolarEdge, a widely renowned simulator. Emphasis will be placed on hybrid Long Short-Term Memory (LSTM) based architectures to carry out this assessment effectively.

Furthermore, the statistical analysis conducted in this study meticulously evaluates representative characteristics of the one-year analysis period. This comprehensive investigation marks the first-ever assessment of energy production from a micro-inverter installation under simulated conditions within Central Eastern Europe, specifically in the Swietokrzyskie Voivodeship region.

## 2. Materials and Methods

Simulation-based methods and AI predictions represent distinct approaches applied in the estimation of electricity generation from renewable energy sources (RES). Let's delve into each of these methodologies:

### 2.1. *Simulation-Based Techniques*

Simulation-based methods entail the development of computer models that emulate the behavior of renewable energy systems under various conditions. These models consider a myriad of factors, including weather patterns, geographical location, technical specifications of the renewable energy sources (PV farms), and the unique characteristics of the energy grid.

Advantages of Simulation-Based Methods:

- Precision: Simulation models can offer precise predictions when properly calibrated and grounded in real-world data.
- Flexibility: They can be tailored to accommodate diverse scenarios and configurations.
- Understanding: Simulations facilitate comprehension of the intricate interactions between different components within the energy system.

Disadvantages of Simulation-Based Methods:

- Computational Intensity: Developing and executing simulations can be computationally demanding, particularly for large-scale systems.
- Data Requirements: High-quality data pertaining to weather, energy demand, and system parameters are crucial for ensuring accurate simulations.
- Complexity: Constructing a reliable simulation model necessitates expertise and domain knowledge.

### 2.2. *AI Predictions for Electricity Generation from RES*

AI predictions encompass the use of machine learning algorithms, such as neural networks, to scrutinize historical data and make forecasts regarding future electricity generation from renewable sources. These algorithms discern patterns and relationships from data and can generate projections based on these identified patterns.

Advantages of AI Predictions:

- Speed: AI models can rapidly process vast data-sets and provide real-time predictions.
- Adaptability: They possess the ability to adapt to changing conditions and learn from newly acquired data.

- Scalability: AI models demonstrate efficacy in handling large-scale systems.

Disadvantages of AI Predictions:

- Data Quality: The accuracy of AI predictions heavily relies on the quality and sentative character of the training data.
- Black Box Nature: Neural networks can be challenging to interpret, making it difficult to comprehend the underlying reasoning behind predictions.
- Limited Generalization: AI models may encounter difficulties when dealing with unseen situations or data that significantly deviate from the training set.

**Combining Both Approaches:** In practice, a hybrid approach, integrating simulation-based methods and AI predictions, emerges as a viable solution to capitalize on the strengths of each technique. For instance, simulation models can furnish synthetic training data for AI models, augmenting their accuracy. Moreover, AI models can optimize specific parameters within the simulation model, thereby enhancing overall performance.

Ultimately, the selection between simulation-based methods and AI predictions hinges on factors such as the availability of data, the complexity of the RES system, the desired level of accuracy, and the computational resources at hand. In numerous cases, a combined methodology presents the most robust and dependable outcomes.

### 2.3. Analyzed designed PV farm

Data for further research were downloaded from SolarEdge designer:

- The PV farm has a total installed DC power capacity of 27.25 kilowatts peak (KWp). This is achieved through the use of 50 JAM72D30-545/MB solar panels.
- The PV system utilizes a Solar Edge SE25K inverter, which has a capacity of 25,000 watts. The inverter converts the DC power generated by the solar panels into AC power, suitable for use in homes and businesses.
- To maximize energy production and improve efficiency, the PV farm is equipped with 50 Solar Edge S650B optimizers. These optimizers individually track and optimize the performance of each solar panel, ensuring that the entire system operates at its highest potential.
- The maximum AC power output of the PV farm is 25 kilowatts (KW). This represents the highest amount of electricity that the system can produce and supply to the grid or connected loads at any given moment.
- The PV farm's estimated annual energy production is 28.10 megawatt-hours (MWh). This figure indicates the total amount of electricity the system is expected to generate over the course of a year under average sunlight conditions.
- The performance index of the PV farm is 91%. This metric represents the efficiency of the system in converting sunlight into electricity. An index of 91 suggests that the PV farm is performing at a high level of efficiency, considering various factors such as shading, temperature, and soiling losses.

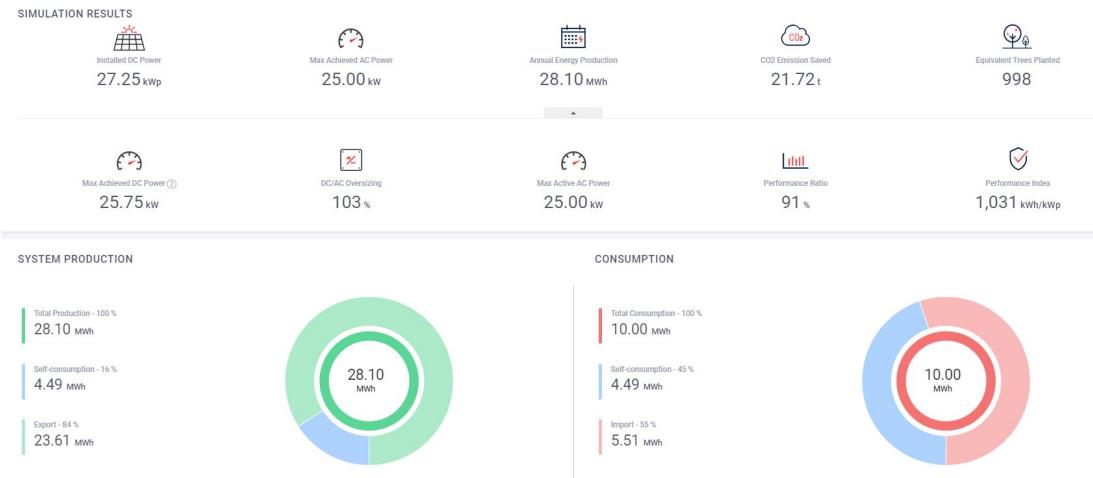


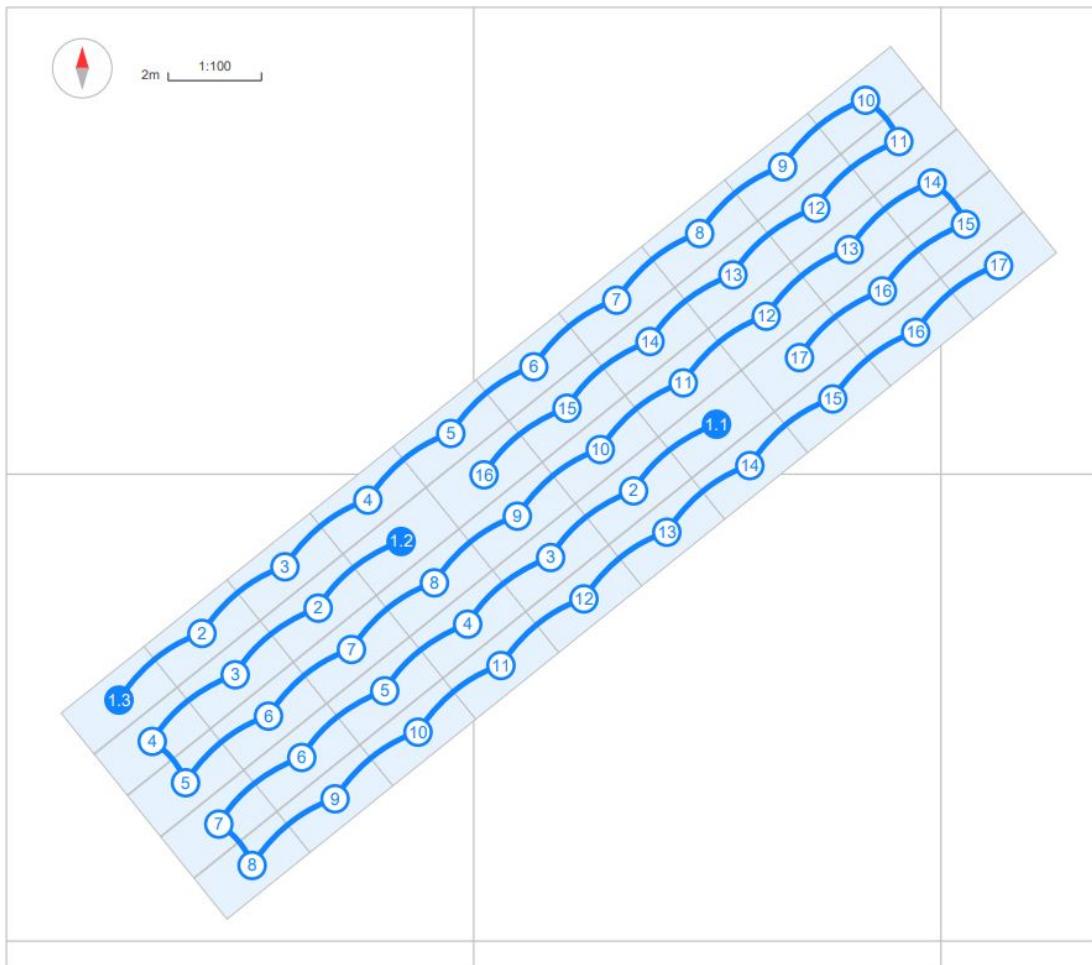
Figure 1. Microinstallation parameters



Figure 2. Estimated monthly energy production



Figure 3. Proposed localization of PV panels on real location, Azimuth 145, slope 10



**Figure 4.** Two mppt strings connection between PV panels

#### 2.4. Parameters used as input for DNN

The list provided contains various parameters related to a solar photovoltaic (PV) system and its performance, as well as some parameters related to energy storage and consumption. These parameters can be used as input features for a Deep Neural Network (DNN) model to analyze and predict the performance and behavior of the solar PV system. Here's a description of each parameter:

- **Timestamp:** The date and time at which the data is recorded, providing a temporal context for the other parameters.
- **Global Horizontal Irradiance:** The total solar irradiance received on a horizontal surface, including direct and diffuse sunlight.
- **Diffuse Horizontal Irradiance:** The solar irradiance from the sky's scattered radiation received on a horizontal surface.
- **Diffused Normal Irradiance:** The solar irradiance received on a surface perpendicular to the sun's rays, after scattering in the atmosphere.
- **Global Tilted Irradiance:** The total solar irradiance received on a tilted surface, which is typically the PV panel's orientation.
- **Temperature:** The ambient temperature of the solar PV panels or system.
- **GTI Loss:** Loss in Global Tilted Irradiance, potentially due to shading or orientation issues.
- **Irradiance Shading Loss:** Loss in irradiance due to shading on the PV panels.
- **Array Incidence Loss:** Loss caused by the angle at which sunlight strikes the PV panels.
- **Soiling and Snow Loss:** Loss due to the accumulation of dirt, dust, or snow on the PV panels, reducing their efficiency.
- **Nominal Energy:** The expected energy output from the PV system under ideal conditions.

- Irradiance Level Loss: Loss related to variations in solar irradiance levels.
- Shading Electrical Loss: Electrical losses occurring in the system due to shading.
- Temperature Loss: Electrical losses caused by temperature variations in the PV panels.
- Optimizer Loss: Losses associated with power optimizers used in the PV system.
- Yield Factor Loss: Losses due to factors impacting the system's overall yield.
- Module Quality Loss: Losses caused by the quality of PV modules used in the system.
- LID Loss: Light-Induced Degradation losses in PV modules.
- Ohmic Loss: Electrical losses caused by the resistance of materials in the system.
- String Clipping Loss: Losses due to limiting the output of strings of PV panels.
- Array Energy: The total energy output from the PV array.
- Inverter Efficiency Loss: Losses due to inefficiencies in the inverter.
- Clipping Loss: Losses due to limiting the system's output to protect the inverter.
- System Unavailability Loss: Losses due to system downtime or unavailability.
- Export Limitation Loss: Losses caused by limitations on exporting excess energy to the grid.
- Battery Charge: The amount of energy being charged into the battery storage system.
- Battery Discharge: The amount of energy being discharged from the battery storage system.
- Battery State of Energy: The current energy level or state of charge of the battery.
- Inverter Output: The energy output from the inverter.
- Consumption: The amount of energy consumed by the user or load.
- Self-consumption: The energy consumed from the PV system by the user.
- Imported Energy: The amount of energy imported from the grid.
- Exported Energy: The amount of excess energy exported to the grid.

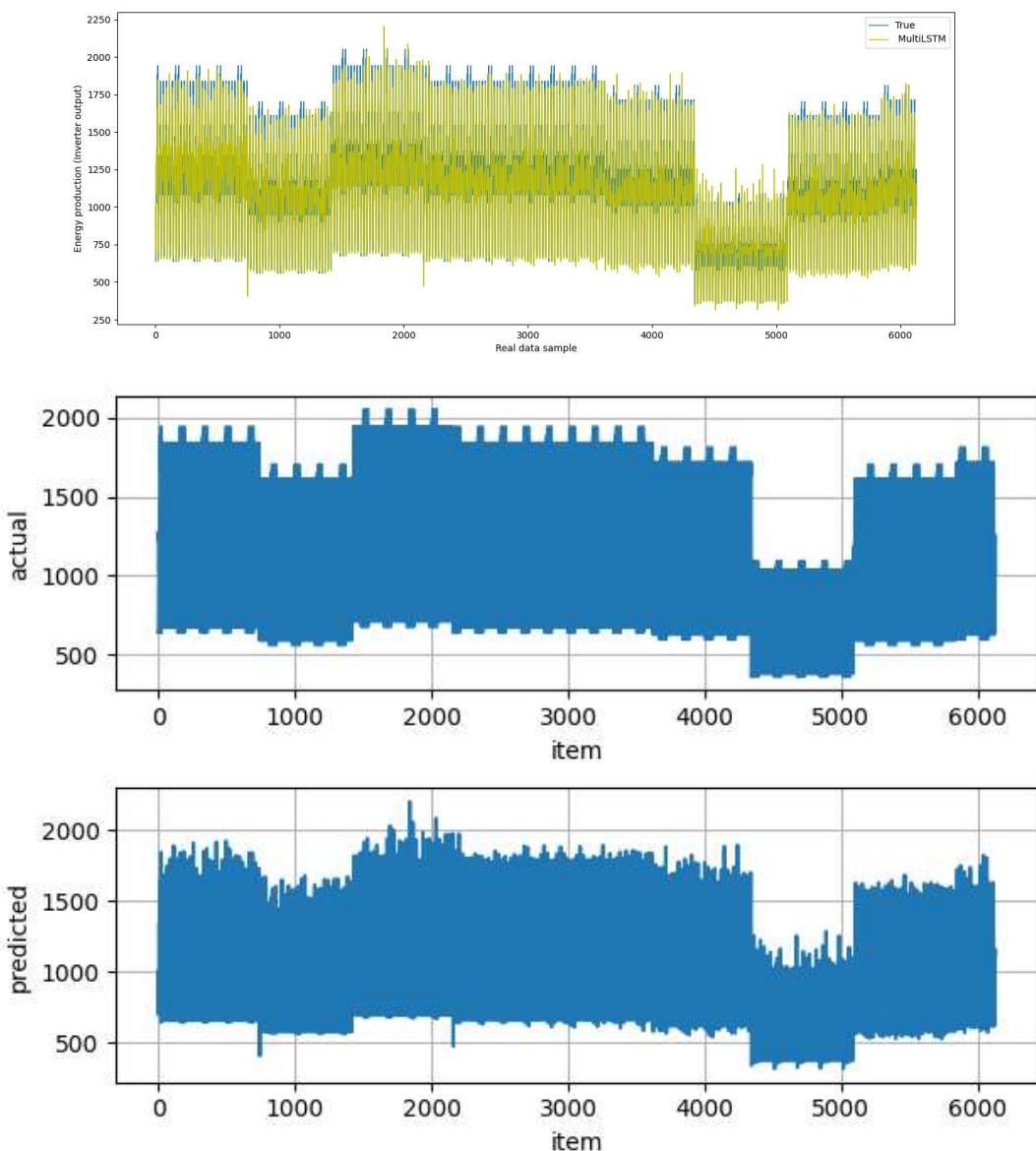
Using these parameters as input to a DNN model, it could be possible to analyze the performance of the solar PV system, predict energy output, optimize operations, and explore various scenarios for efficient energy utilization and storage.

### 3. Results

Below we present our comparison between values reported by simulations and our DNN models. Based on the DNN (Deep Neural Network) metrics provided, here is a description of the MultiLSTM results:

1. MAE (Mean Absolute Error): 98.531 The MAE represents the average absolute difference between the predicted values and the actual values. In this case, the average absolute error is 98.531, which means, on average, the predictions made by the DNN are off by approximately 98.531 Kwh.
2. NMAE (Normalized Mean Absolute Error): 0.048 NMAE is the MAE normalized by the mean of the actual values. It is a relative measure of the MAE and is useful for comparing performance across different data-sets. In this case, the NMAE is 0.048, indicating that the DNN's predictions have an average absolute error of 4.8% relative to the mean of the actual values.
3. MPL (Mean Percentage Loss): 49.265 The MPL represents the average percentage difference between the predicted values and the actual values. In this case, the average percentage loss is 49.265, which means, on average, the predictions made by the DNN deviate from the actual values by approximately 49.265%.
4. MAPE (Mean Absolute Percentage Error): 0.085 MAPE is a common metric for measuring the accuracy of predictions as a percentage. It represents the average absolute percentage difference between the predicted values and the actual values. In this case, the average absolute percentage error is 0.085, meaning that the DNN's predictions have an average error of 8.5% relative to the actual values.
5. MSE (Mean Squared Error): 18724.956 The MSE represents the average of the squared differences between the predicted values and the actual values. It gives higher weights to larger errors. In this case, the MSE is 18724.956, providing an insight into the average squared error of the DNN's predictions.

6. EVS (Explained Variance Score): 0.889 EVS measures the proportion of variance in the target variable that is explained by the model. It ranges from 0 to 1, where 1 indicates a perfect fit. An EVS of 0.889 means that the DNN explains about 88.9% of the variance in the data, which is considered a good result.
7. R2 (R-squared): 0.867 R-squared is another measure of how well the model fits the data. It represents the proportion of variance in the target variable that can be predicted from the input features. R-squared also ranges from 0 to 1, with 1 indicating a perfect fit. An R2 value of 0.867 indicates that the DNN explains about 86.7% of the variance in the data.
8. RMSE (Root Mean Squared Error): 136.839 RMSE is the square root of the MSE and is a common metric for evaluating the accuracy of predictions. It represents the average magnitude of the errors made by the DNN. In this case, the RMSE is 136.839, providing an insight into the average magnitude of the errors made by the DNN's predictions.

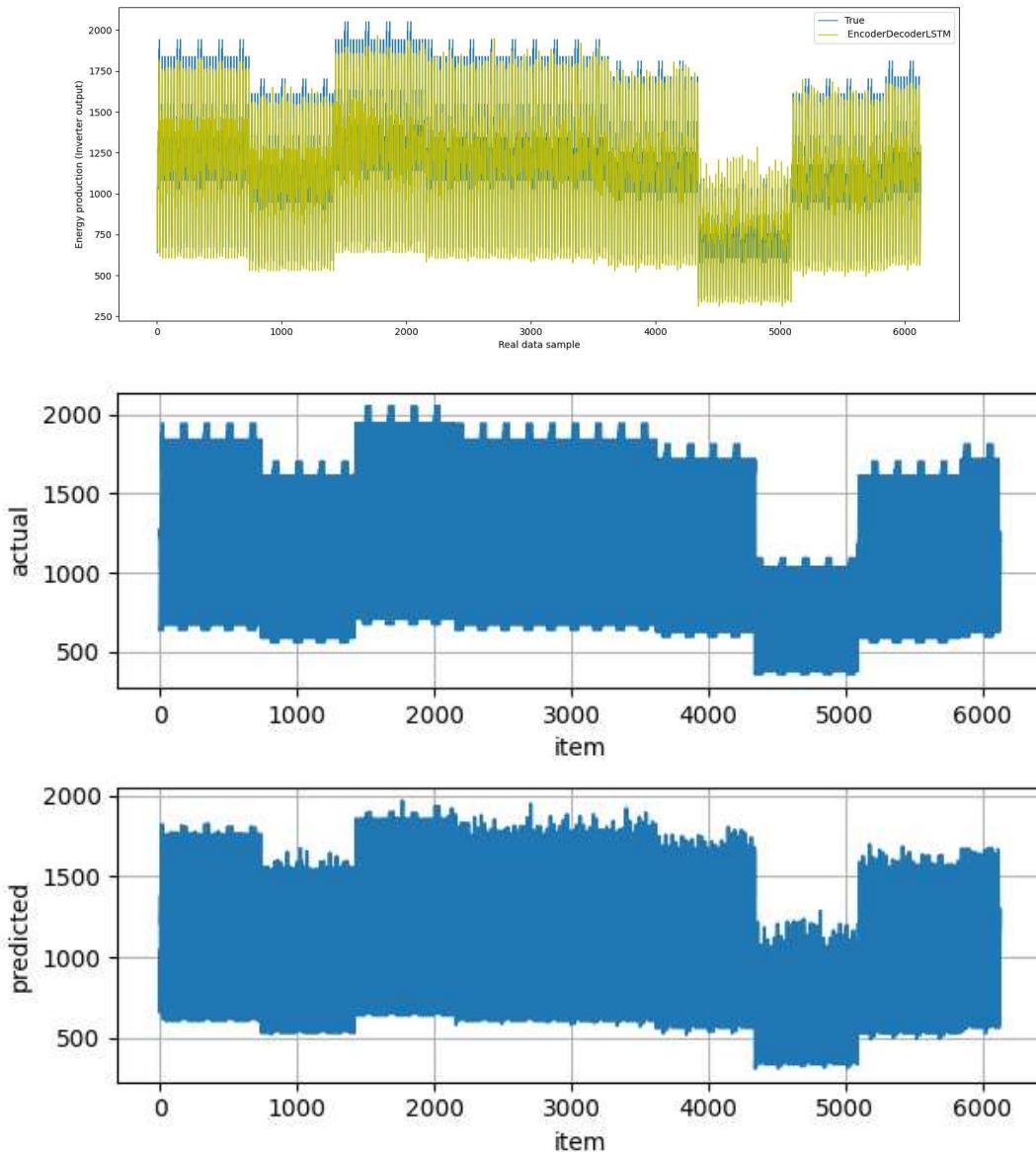


**Figure 5.** MultiLSTM comparison results of inverted output power simulation vs DNN model

**Table 1.** Metrics MultiLSTM

MAE: 98.531	NMAE: 0.048	MPL: 49.265
MAPE: 0.085	MSE: 18724.956	
EVS: 0.889	R2: 0.867	RMSE: 136.839

Overall, the DNN seems to be performing reasonably well, with relatively low errors and high explained variance and R-squared values.



**Figure 6.** EncoderDecoderLSTM comparison results of inverted output power simulation vs DNN model

**Table 2.** Metrics EncoderDecoderLSTM

MAE: 130.454	NMAE: 0.063	MPL: 65.227
MAPE: 0.124	MSE: 29203.207	
EVS: 0.826	R2: 0.805	RMSE: 170.889

In this scenario, the analysis reveals the following performance metrics for the model:

1. Mean Absolute Error (MAE): The MAE is measured at approximately 130.454 KWh, indicating the average discrepancy between predicted and actual values.
2. Normalized Mean Absolute Error (NMAE): The NMAE stands at around 0.063, representing the relative average absolute error.
3. Mean Absolute Percentage Error (MAPE): The MAPE is about 0.124%, signifying the average error as a percentage of the actual values.
4. Mean Squared Error (MSE): The MSE is approximately 29203.207, reflecting the average squared difference between predicted and actual values.
5. Explained Variance Score (EVS): With a value of approximately 0.826, the EVS suggests a reasonably good fit of the model to the data. An EVS close to 1 indicates a stronger fit.
6. Coefficient of Determination (R<sup>2</sup>): The R<sup>2</sup> value is estimated at around 0.805, which serves as a measure of the model's accuracy in explaining the variance in the dependent variable.
7. Root Mean Squared Error (RMSE): The RMSE is evaluated at approximately 170.889 KWh, signifying the square root of the average squared error.

EncoderDecoderLSTM model demonstrates reasonable performance within the context and specific domain of the problem.

#### 4. Discussion

Emphasizing the significance of data abundance for model improvement, it becomes evident that increasing the data-set enhances output quality, particularly after applying feature selection. The accuracy of models has been assessed using R<sup>2</sup> and NMAE metrics, with the data split into 80% for training and 20% for testing.

Choosing the best model manually can be challenging, relying solely on metrics and visual representations. To address this, an automated approach has been adopted. Leveraging benchmarks and a framework, proper hyperparameters and optimized models have been selected, considering factors like calculation costs, weight, and performance.

The automated environment presents a distinct advantage in enabling easy comparison of results. Metrics are measured for each method and learning stage, unveiling the advantages and drawbacks of different architectures. Focus on optimization aspects becomes achievable, especially when evaluating multiple horizons and methodologies. It is worth noting that the evaluation process took approximately one and a half months.

To delve into the intricacies of energy forecasting, the models' behavior has been extensively tested concerning energy production from individual photovoltaic microinstallation based on example from Świętokrzyskie Voivodeship Poland. This examination is based on the total energy output from micro-installation.

In conclusion, this pioneering study seeks to address the challenges posed by existing PV energy production calculation methods in Polish conditions. By identifying potential inaccuracies due to factors such as terrain, location, wind conditions, and PV panel representative characteristics, the research highlights the importance of refining these methods for more precise results.

To achieve this, the study explores the application of Deep Neural Network (DNN) methods, with a particular focus on hybrid Long Short-Term Memory (LSTM) based architectures. Through the evaluation of data provided by the esteemed SolarEdge simulator, the accuracy of energy production calculations can be thoroughly verified.

Moreover, this research stands as a significant milestone as it pioneers the assessment of energy production from a microinverter installation under simulated conditions in the Central Eastern European region, specifically in the Świętokrzyskie Voivodeship area. The statistical analysis conducted in this study meticulously examines the representative characteristics of the one-year analysis period, providing valuable insights into the potential of PV energy generation in this specific context.

Overall, this study's findings and methodologies pave the way for future advancements in PV energy production calculations, fostering the development of more accurate and efficient methods tailored to the unique conditions of Poland and similar regions. With the integration of cutting-edge DNN techniques, the renewable energy industry can take a step closer to realizing its full potential in achieving a sustainable and eco-friendly future.

## References

1. Demirkiran, Muhammet and Karakaya, Abdulhakim. "Efficiency analysis of photovoltaic systems installed in different geographical locations" *Open Chemistry* 20, no. 1 (2022): 748-758. <https://doi.org/10.1515/chem-2022-0190>
2. Roberts, J.; Cassula, A.; Freire, J.; Prado, P. *Simulation and Validation Of Photovoltaic System Performance Models*. In *Proceedings of the XI Latin-American Congress on Electricity Generation and Transmission-CLAGTEE 2015*, Sao Paulo, Brazil, 8–11 November 2015.
3. Milosavljević, D.D.; Kevkić, T.S.; Jovanović, S.J. *Review and validation of photovoltaic solar simulation tools/software based on case study*. *Open Phys.* 2022, 20, 431–451.
4. Jaroslaw Bilski, Leszek Rutkowski, Jacek Smolag, Dacheng Tao: A novel method for speed training acceleration of recurrent neural networks. *Inf. Sci.* 553: 266-279 (2021)
5. Bochenek, B.; Jurasz, J.; Jaczewski, A.; Stachura, G.; Sekuła, P.; Strzyżewski, T.; Wdowikowski, M.; Figurski, M. *Day-Ahead Wind Power Forecasting in Poland Based on Numerical Weather Prediction*. *Energies* 2021, 14, 2164. <https://doi.org/10.3390/en14082164>
6. Shao, X.; Kim, C.-S.; Sontakke, P. *Accurate Deep Model for Electricity Consumption Forecasting Using Multi-Channel and Multi-Scale Feature Fusion CNN-LSTM*. *Energies* 2020, 13, 1881. <https://doi.org/10.3390/en13081881>
7. Jabreel, Mohammed & Moreno, Antonio. (2019). *A Deep Learning-Based Approach for Multi-Label Emotion Classification in Tweets*. *Applied Sciences*. 9. 1123. 10.3390/app9061123.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.