Transition Into Solar Elevation Angle Domain for Photovoltaic Power Generation Forecasting

Paweł Parczyk[1][0009−0004−3287−9520], **Robert Burduk**[1][0000−0002−3506−6611]

[1] *Wrocław University of Science and Technology*
Department of Systems and Computer Networks
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
pawel.parczyk@pwr.edu.pl, robert.burduk@pwr.edu.pl

Abstract. *Nowadays, we observe an energy transition in which fossil fuel plants are replaced with renewable energy sources. It leads to new challenges. One of them is the inability to control energy production. Therefore, yield forecasting has become a pressing problem. However, previous work related to the issue of forecasting yields from photovoltaic installations was based mainly on the observation of past data. Therefore, the research did not focus on the phenomena that form the production. This work presents an alternative approach to searching for the source of a phenomenon instead of observing the effects. The article presents an analysis of the course of the sun and its impact on yields. The experiments that were performed showed the competitiveness of the proposed method compared to the Long Short-Term Memory.*

Keywords: *Renewable Energy Sources (RES), Photovoltaic (PV), Photovoltaic forecasting, Solar elevation, Solar movement, Yield forecasting*

1. Introduction

In modern society, electricity plays a very important role. It could be said that society is addicted to it. Nowadays, it is hard to imagine life without access to energy. Therefore, the blackout has become a major threat with the potential to cause significant consequences [1]. Unfortunately, fossil fuels on which the power system is currently based lead to environmental pollution. Moreover, their resources are limited. To counteract these effects, we have started an energy transition [2], which is based on Renewable Energy Sources (RES).

However, RES also has its own limitations. Production from RES strongly depends on many factors such as location, temperature, dust occurrence, weather, and cloud cover [3]. In other words, wind turbines need wind, and photovoltaics need the sun to produce energy.

Additionally, photovoltaic production is much higher on long summer days than on short and cloudy winter days [4]. Apart from seasonal, long-term changes, solar power plants are also exposed to sudden changes in sunlight. It happens when weather conditions change rapidly. This may cause short but significant fluctuations in power, which introduce instability to the grid. Therefore, energy forecasting may allow us to overcome this issue.

This work presents a method for predicting yields from photovoltaic installations, which is based on the analysis of the impact of the sun's height above the horizon on power generation.

2. Model assumptions

As the study shows [5], topics related to photovoltaic forecasting have become popular. Scientists are focused on methods related to RNN, such as LSTM [6] or more complex BiLSTM-CNN models [7]. However, the most common approach is to look at the dataset and use measured values to form the models as a black box. In other words, they do not consider phenomena that form the data.

In this section, the sun's movement and its effect on production are described.

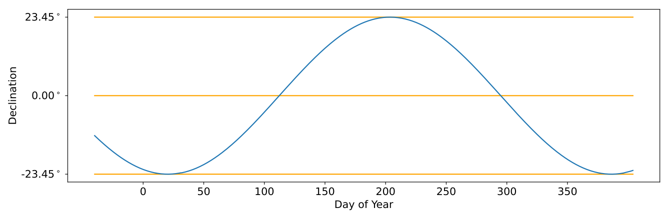

Figure 1. Declination angle during the year.

The sun's height above the horizon α depends on several factors: a constant one, a yearly and daily changing. The first is the latitude L of the location of the point on the Earth's surface. The second is the sun's declination angle δ, i.e., the planet's rotation axis angle from the plane of the normal to the disk along which the Earth orbits the sun. This value ranges from -23.45 to +23.45 degrees, and changes occur annually. Moreover, it is the same for each location on the planet at the time. Figure 1 shows the behavior of the declination value for subsequent days of the year. We can observe these changes as the seasons of the year. The solar hour *HRA*, which changes daily. It is the third factor influencing the sun's altitude. The sun's height above the horizon for any moment of the year can be calculated from the equation 1.

$$sin(\alpha) = sin(L)sin(\delta) + cos(L)cos(\delta)cos(HRA) \tag{1}$$

Figure 2 portrays the annual course, omitting negative angles. There are four specific moments during the year. These are the summer solstice, winter solstice, and two equinoxes. Figure 3 compares the sun's altitude on those specific days.

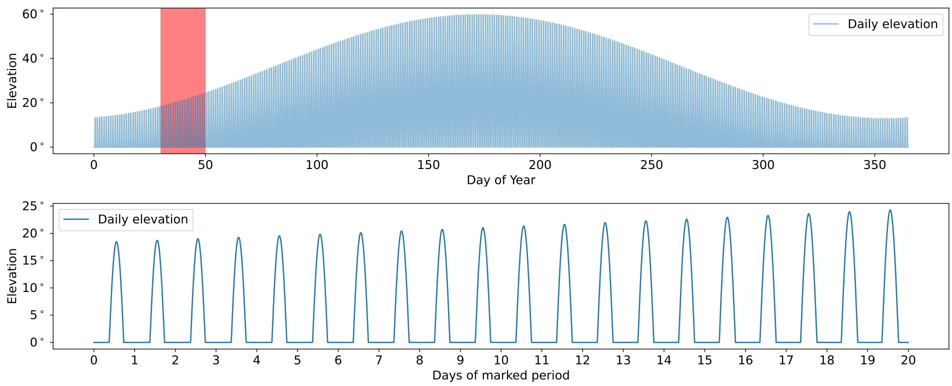

Figure 2. Elevation angle during the year (top) and a closer look for red marked rage (bottom) for a location with a latitude of 53.671^o.

The time-based approach (such as RNN-based models) is based on the relation between future and past data in the time domain. This relation can be described by equation $Y_{t+1} = M_1(time_t, \theta) + \epsilon_t$. Based on the presented phenomena, we can assume that the production is a function of the sun's position rather than time. Therefore, we propose transitioning from the time domain to the solar elevation angle domain. In that case the problem can be reformulated as $Y_{t+1} = M_2(elevation(time_t), \theta) + \epsilon_t \rightarrow Y_{t+1} = M_2(\alpha_t, \theta) + \epsilon_t$.

The principle of operation of the model is based on assigning subsequent production values Y_t to groups depending on the elevation angle at which they occurred.

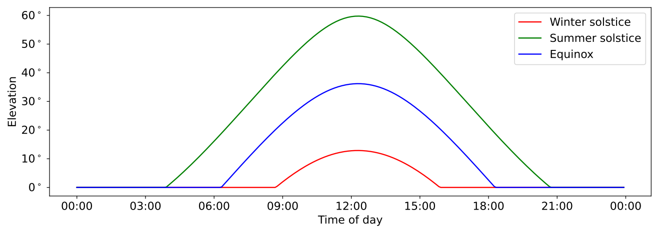

Figure 3. Comparison of Sun's elevation during most characteristic days: summer solstice, equinoxes, and winter solstice for a location with a latitude of 53.671^o.

First, during the model fitting, the sun's elevation angle is calculated for the selected period. The *window_size* parameter defines the length. Then, the angle values are quantized. This process assigns similar elevation angles to the individual corresponding groups. For example, assume that the maximum available sun elevation angle for the installation under analysis is 60 degrees and the minimum is 0 degrees. The number of bins is specified by *elevation_angle_bins*. When divided into six groups, each group will cover a range of 10 degrees (the first group will be from 0 to 10 degrees, the second from 10 to 20, etc.). Then, all values are aggregated into one value. A weighted average of all Y_t values belonging to each group is calculated. Those averages constitute the model θ.

Forecasting involves assigning production values to future points in time. This is done in a similar way to fitting. First, the sun's elevation angles in the processed period are calculated, and then values are assigned to these angles. This forecast becomes a stair-like forecast. Then, to eliminate this effect, the forecast is smoothed using linear interpolation. An example of a 6 hours ahead forecast is presented in Figure 4. The forecast was made for two consecutive days, one of which is sunny and the other is a day with temporary cloud cover. The upper chart presents a stair-like forecast, while the lower presents a smooth one. Moreover, it can be observed that for extremely good weather, the model underestimates the forecast, while for fine weather, it is much more precise.

3. Evaluation

Tests were performed to verify the effectiveness of the method proposed. Tests were carried out using the rolling window scheme with a constant window size.

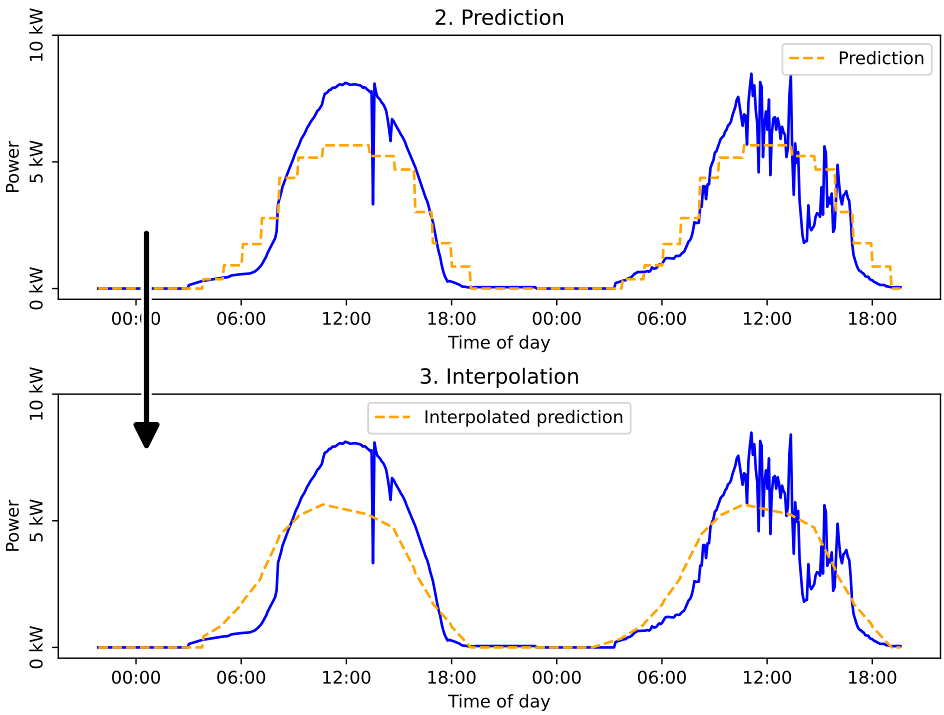

Figure 4. Sample 6 hours ahead forecast (the upper graph: generated from calculated weighted averages for each group of elevation angles) and forecast after applying linear interpolation (the lower graph: forecast after interpolation) for an exemplary period of two consecutive days containing a sunny day (from the left) and a day with transient cloudiness (from right).

The proposed method was compared to the state-of-the-art one-layer LSTM model. The comparison was performed on a high-resolution (5-minute sampling) dataset. Obtained from two different photovoltaic installations installed in Poland. The first installation is 9.9kWp, and the second is 6.12kWp. Both installations face south. The Dataset spans between 21^{st} June 2020 (summer solstice) and 21^{st} December 2023 (winter solstice) therefore, it contains 368350 1-dimensional samples (a power generated in kilowatts) from each installation.

The models were used to generate forecasts in three different time horizons. The first one is a short-term forecast for 6 hours ahead. The second horizon is

Table 1. List of considered proposed hyperparameters. In bold, selected values are marked.

Proposed model	
window_size	7 days, 14 days, 30 days, 45 days, 60 days, **75 days**, 90 days
elevation_angle_bins	10, 26, **40**, 50, 60

Table 2. List of considered reference model hyperparameters. In bold, selected are marked.

LSTM	
window_size	7 days, **14 days**, 30 days, 45 days, 60 days
inputs	2h, 4h, **6h**
max_epochs	25, 70, **100**, 130, 150
layer	**LSTM(32)**

longer and is 12 hours ahead. The third time horizon, which is 24 hours, can be classified as a medium-term forecast. The selection of horizons was based on their usefulness. Short-term and medium-term forecasts are useful for scheduling power plant operations.

A grid search method was used to find the optimal configuration of each model. During selection, models were tested on data covering one year. The considered combinations of hyperparameters are presented in Table 1 and Table 2. In each row in each table, the optimal parameter was marked with bold.

Table 3 presents the results of evaluating individual models on the database. The results were created by averaging the metrics values obtained by individual models during subsequent years (i.e., 2021, 2022, 2023). The evaluation was performed using the commonly used MAE, RMSE, NMAE, and R2 metrics.

Better values are bold in the table. It can be concluded that the proposed method copes better with short-term forecasting for 6 hours than the reference model. However, it is characterized by high variability of the obtained results, as indicated by the high standard deviation value. However, this relationship is not identical for both databases. Therefore, it can be said that the dispersion of results depends strongly on the entered data (with a larger installation, there is a greater dispersion of results) in the case of MEA and RMSE. The NMAE shows that the proposed method is better for 6 hours horizon and 12 hours horizon than LSTM.

Table 3. Yearly models' metrics that were achieved on processing various time horizons.

Model	Averaged annual prediction results					
	6h		12h		24h	
	MAE	*RMSE*	*MAE*	*RMSE*	*MAE*	*RMSE*
D1						
LSTM	0.81 ± 0.032	1.47 ± 0.076	0.80 ± 0.065	1.46 ± 0.124	0.64 ± 0.038	1.33 ± 0.082
Proposed	**0.59 ± 0.139**	**1.30 ± 0.220**	**0.60 ± 0.141**	**1.31 ± 0.220**	**0.59 ± 0.141**	**1.31 ± 0.218**
D2						
LSTM	0.44 ± 0.041	0.77 ± 0.066	0.45 ± 0.035	0.79 ± 0.051	0.35 ± 0.032	0.72 ± 0.064
Proposed	**0.33 ± 0.057**	**0.69 ± 0.080**	**0.34 ± 0.059**	**0.70 ± 0.086**	**0.34 ± 0.059**	**0.70 ± 0.087**

	6h		12h		24h	
	NMAE	R2	NMAE	R2	NMAE	R2
D1						
LSTM	0.81 ± 0.100	**0.43 ± 0.034**	0.79 ± 0.058	**0.44 ± 0.037**	**0.65 ± 0.012**	0.48 ± 0.011
Proposed	**0.66 ± 0.054**	0.47 ± 0.078	**0.66 ± 0.057**	0.46 ± 0.084	0.66 ± 0.057	**0.46 ± 0.086**
D2						
LSTM	0.82 ± 0.159	**0.42 ± 0.102**	0.76 ± 0.061	**0.47 ± 0.042**	**0.62 ± 0.030**	0.54 ± 0.030
Proposed	**0.64 ± 0.041**	0.53 ± 0.050	**0.64 ± 0.036**	0.52 ± 0.046	0.64 ± 0.036	**0.52 ± 0.047**

Additionally, it shows that the complexity of data in both datasets is almost the same. The obtained results for both datasets are close to each other regarding R2 and NMAE. In the case of 12-hour forecasts, the results of both models are closer to each other than for the 6-hour forecast, similar to the case of forecasts for the longest horizon. However, the proposed model remains the most effective for all forecast ranges discussed. The interesting fact, which needs deeper research is that the proposed method is better in terms of NMAE and worst in terms of R2 for the 6 and 12-hour horizon, while for the 24-hour horizon, the relation is reversed.

4. Conclusions

In this paper, an alternative approach to the issue of forecasting yields from photovoltaic installations is presented. The proposed method involves transitioning from the time domain to the sun elevation angle domain. Of course, the sun's elevation angle also depends on time and changes periodically. The proposed model was tested on a few years of data and, despite its simplicity, turned out to be competitive with the advanced reference LSTM model.

Moreover, the model is a good starting point for further research. The model is susceptible to development, and further steps will involve more extensive testing and the proposal of alternatives for the weighted average aggregation function.

References

[1] Matthewman, S. and B., R. H. Blackouts: a sociology of electrical power failure. *Socialspace*, pages 1–25, 2014.

[2] Akpan, J. and Olanrewaju, O. Sustainable energy development: History and recent advances. *Energies*, 16(20), 2023. ISSN 1996-1073. doi:10.3390/en16207049. URL https://www.mdpi.com/1996-1073/16/20/7049.

[3] Bonkaney, A. L., Madougou, S., Adamou, R., et al. Impact of climatic parameters on the performance of solar photovoltaic (pv) module in niamey. *Smart Grid and Renewable Energy*, 8(12):379, 2017.

[4] Alhamer, E., Grigsby, A., and Mulford, R. The influence of seasonal cloud cover, ambient temperature and seasonal variations in daylight hours on the optimal pv panel tilt angle in the united states. *Energies*, 15(20), 2022. ISSN 1996-1073. doi:10.3390/en15207516. URL https://www.mdpi.com/1996-1073/15/20/7516.

[5] Iheanetu, K. J. Solar photovoltaic power forecasting: A review. *Sustainability*, 14(24), 2022. ISSN 2071-1050. doi:10.3390/su142417005. URL https://www.mdpi.com/2071-1050/14/24/17005.

[6] Gao, M., Li, J., Hong, F., and Long, D. Short-term forecasting of power production in a large-scale photovoltaic plant based on lstm. *Applied Sciences*, 9(15), 2019. ISSN 2076-3417. doi:10.3390/app9153192. URL https://www.mdpi.com/2076-3417/9/15/3192.

[7] Li, G., Guo, S., Li, X., and Cheng, C. Short-term forecasting approach based on bidirectional long short-term memory and convolutional neural network for regional photovoltaic power plants. *Sustainable Energy, Grids and Networks*, 34:101019, 2023. ISSN 2352-4677. doi:https://doi.org/10.1016/j.segan.2023.101019. URL https://www.sciencedirect.com/science/article/pii/S2352467723000279.