

Journal Scientific and Applied Research, Vol. 31, 2026
International Journal

ISSN 1314-6289 (Print)
ISSN 2815-4622 (Online)

Original Contribution

APPROACH TO MODELING PHOTOVOLTAIC ENERGY YIELD USING CHEBYSHEV POLYNOMIALS

Iliyana S. Ivanova[1], Donika V. Dimanova[2]

[1]Department of Communication and Computer Engineering and Security Technologies, Faculty of Technical Sciences, Konstantin Preslavsky University of Shumen, Shumen 9712, 115 Universitetska St., e-mail: i.s.ivanova@shu.bg
[2]Department of Communication and Computer Engineering and Security Technologies, Konstantin Preslavsky University of Shumen, Shumen 9712, 115 Universitetska St., e-mail: d.dimanova@shu.bg

Abstract: *Photovoltaic energy yield varies in response to changes in meteorological conditions. The main factors affecting the amount of electricity generated include solar irradiance and air temperature, as well as seasonal and daily variations. This paper examines the relationships between these factors and photovoltaic energy yield and presents an approach to modeling those using orthogonal Chebyshev polynomials. The main input variables, data-preprocessing steps, and structure of the approximation model are defined.*

Key words: *Photovoltaic systems, Photovoltaic energy yield, Chebyshev polynomials, Solar irradiance, Temperature, Mathematical modeling.*

1. Introduction

The electrical energy produced by photovoltaic systems depends directly on changes in meteorological conditions [1]. Solar irradiance varies both throughout the day and across seasons [2]. Air temperature also affects the operating characteristics and efficiency of photovoltaic modules. These factors influence the amount of electricity generated to varying degrees [3]. Therefore, their effects cannot always be described accurately enough by a linear model [4].

A mathematical approach is needed to describe these effects and represent nonlinear, interrelated processes. Mathematical models are used to analyze complex interrelated processes in various application areas [5], [6]. Orthogonal polynomials are used in statistical modeling and in the analysis of relationships across different fields. For this purpose, orthogonal Chebyshev polynomials can be used to construct a multidimensional approximation model describing the relationship between meteorological factors and photovoltaic energy yield.

This paper identifies the variables involved in constructing the model, analyzes their influence, and formulates the sequence of steps required for its development.

2. Main Factors and Temporal Characteristics Used as Input Variables

To construct the model, the factors and temporal characteristics that affect photovoltaic energy yield must first be identified. The main variables considered in developing the model are solar irradiance, air temperature, seasonal variations, and time of day [7], [8].

2.1. Solar Irradiance

Solar irradiance is the primary energy source for photovoltaic modules and is used as a key input variable in constructing the model. In this study, it is denoted by R and measured in W/m². The amount of electrical energy produced depends on the solar energy reaching the surface of the photovoltaic modules. As solar irradiance increases, the electric current generated by the photovoltaic cells also increases, leading to higher output power.

The photovoltaic power output can be expressed as (1):

$$P_{PV} = \eta A R \tag{1}$$

where P_{PV} is the photovoltaic power output, η is the efficiency of the photovoltaic modules, A is their active area, and R is the solar irradiance. The photovoltaic power output P_{PV} describes the instantaneous rate of electricity generation, whereas the generated electrical energy E is obtained by integrating the power output over a specified time interval.

The relationship between solar irradiance and power generated is not the same under all conditions. It is affected by module temperature, the angle of incidence of sunlight, the orientation and tilt of the panels, and the operating characteristics of the inverter. Changes in cloud cover can also cause rapid variations in irradiance and, consequently, in instantaneous PV output power and the energy generated over time [9], [10]. Therefore, in constructing the model, solar irradiance is considered together with the other selected input variables.

2.2. Air Temperature

Air temperature is included as one of the main input variables in constructing the model. Its effect on the electrical power generated is primarily related to the temperature dependence of the electrical characteristics of photovoltaic cells. As cell temperature increases, voltage generally decreases, which can lead to a reduction in maximum output power and conversion efficiency [11].

The operating temperature of photovoltaic modules is influenced by air temperature, incoming solar irradiance, and heat-transfer conditions. It depends on several factors acting together, including solar irradiance, air temperature, wind speed, installation conditions, and heat exchange with the surrounding environment. Under high irradiance, module temperature can significantly exceed air temperature, intensifying the effect of temperature on the modules' electrical characteristics [12].

Therefore, temperature should not be considered independently of solar irradiance. Both parameters simultaneously affect the operating conditions of a photovoltaic system: higher irradiance increases the energy available for conversion but can also heat the modules and thereby alter their efficiency. This indicates a nonlinear relationship and interaction between solar irradiance and temperature, which should be accounted for when constructing the approximation model.

In the proposed model, air temperature is denoted by Q and is used together with solar irradiance R. Chebyshev polynomials make it possible to account for the effect of temperature as an individual input variable and for its interactions with the other input variables through combined polynomial terms.

2.3. Seasonal Variations

The photovoltaic energy yield varies throughout the year. These variations are determined by the length of daylight, the Sun's elevation and position above the horizon, and seasonal changes in meteorological conditions. As a result, the amount of incoming solar energy - and, consequently, photovoltaic energy yield - exhibits a clear annual pattern.

To represent this periodicity in the proposed model, the day of the year, D, is used. Its value indicates the position of each observation within the annual cycle, thereby providing the model with additional temporal information.

Including this temporal variable allows the model to account for the seasonal pattern of photovoltaic energy yield, which cannot be captured using only the current values of solar irradiance and temperature. Days with similar meteorological conditions may occur at different times of the year and differ in daylight duration and the Sun's position, both of which affect the energy produced.

In the model, the day of the year is denoted by D. After normalization, its values are mapped to the interval $[-1,1]$, allowing them to be used as arguments of the Chebyshev polynomials. In this way, seasonal dependencies are represented by polynomial terms in D, and by their interactions with solar irradiance R, temperature Q, and time of day H.

2.4. Time of Day

When hourly data are used, variations in hourly photovoltaic energy yield throughout the day must be considered [13]. These variations are mainly determined by the daily pattern of solar irradiance, which is directly related to the position of the Sun above the horizon [14].

Photovoltaic power generally increases after sunrise, reaches higher levels during the hours of greatest solar irradiance, and gradually decreases in the afternoon and evening. The specific daily profile may vary depending on the season, cloud cover, module orientation and tilt, and other meteorological and operating factors. Accordingly, the daily model uses solar irradiance, air temperature, and day of the year as input variables, whereas the hourly model additionally includes time of day.

For this reason, time of day is included as an additional input variable when hourly data are used. In the mathematical relationship, it is denoted by H and indicates the temporal position of each observation within the daily cycle. After normalization, the hourly values are mapped to the interval $[-1,1]$, allowing them to be used as arguments of the Chebyshev polynomials. This enables the model to account for the characteristic intraday variations in photovoltaic energy yield and for interactions between time of day and solar irradiance R, temperature Q, and day of the year D.

When hourly data are used, including H is particularly important because identical or similar meteorological values may occur at different times of day and correspond to different operating conditions of the photovoltaic system.

Based on the relationships discussed, the following main input variables are used to construct the mathematical model:

Table 1. Model Variables

Variable	Symbol	Unit of measurement	Role in the model
Solar irradiance	R	W/m²	Input variable
Air temperature	Q	°C	Input variable
Day of the year	D	-	Input variable

Variable	Symbol	Unit of measurement	Role in the model
Time of day	H	h	Input variable
Electrical energy produced	E	kWh	Output variable

The output variable of the model is the electrical energy generated by the photovoltaic system, denoted by E and measured in kWh. The relationship between the input variables and the generated electrical energy is expressed as (2):

$$\mathrm{E} = \mathrm{f}(\mathrm{R}, \mathrm{Q}, \mathrm{D}, \mathrm{H}) \tag{2}$$

3. Data Preparation for Modeling

Before the model is constructed, the data must undergo preprocessing, including:

3.1. Synchronizing photovoltaic energy-yield data with the corresponding meteorological observations.

3.2. Checking for missing, invalid, or anomalous values and, where necessary, applying appropriate methods to handle them.

3.3. Normalizing the input variables. This is necessary because Chebyshev polynomials are defined over the interval $[-1,1]$. Each input variable is normalized using (3):

$$\tilde{x} = 2\frac{x - x_{min}}{x_{max} - x_{min}} - 1 \tag{3}$$

where x is the original value and $\tilde{x}$ is the corresponding normalized value.

In this way, all input variables are mapped to the same interval and can be used to calculate the Chebyshev polynomials. The transformation of the original input ranges to the interval $[-1,1]$ is illustrated schematically in Fig. 1.

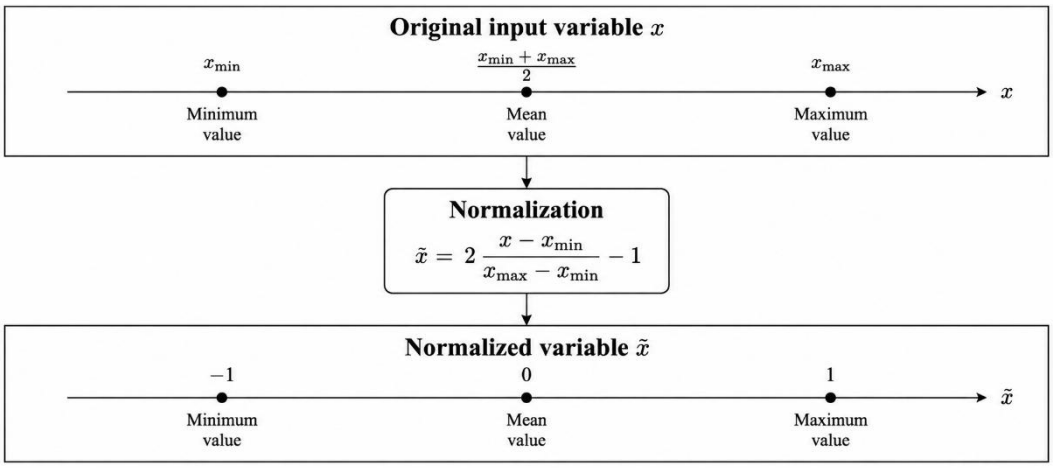


Fig. 1. Schematic representation of the normalization of input variables to the interval $[-1,1]$

4. Chebyshev Polynomials as the Basis of the Model

First-kind Chebyshev polynomials are used to construct the approximation model [15]. Chebyshev polynomials also provide a mathematical basis for developing models that describe various natural processes. The first-kind Chebyshev polynomials are defined by (4):

$$T_n(x) = \cos(n \arccos(x)) \tag{4}$$

where: $x \in [-1,1]$

The first-kind Chebyshev polynomials up to the fifth degree are given in (5):

$$\begin{gathered} T_0(x) = 1 \\ T_1(x) = x \\ T_2(x) = 2x^2 - 1 \\ T_3(x) = 4x^3 - 3x \\ T_4(x) = 8x^4 - 8x^2 + 1 \\ T_5(x) = 16x^5 - 20x^3 + 5x \end{gathered} \tag{5}$$

Chebyshev polynomials are plotted over the interval $[-1,1]$, and their shapes change as their degree increases. The polynomial $T_0(x)$ is a constant function, while $T_1(x)$ varies linearly. Higher-degree polynomials exhibit successive maxima and minima within the range from -1 to 1. As the degree increases, the number of variations within this interval also increases, allowing more complex nonlinear relationships to be represented.

For each normalized input variable in the model, Chebyshev polynomials of different degrees are calculated. For example, the following polynomials are formed for solar irradiance:

$$T_1(\tilde{R}),\ T_2(\tilde{R}),\ T_3(\tilde{R}), \ldots$$

and for temperature:

$$T_1(\tilde{Q}),\ T_2(\tilde{Q}),\ T_3(\tilde{Q}), \ldots$$

These polynomials form the basis functions used to approximate the relationship between the input variables and photovoltaic energy yield [16].

5. Analysis of Input Variables and Model Construction

The model is constructed in stages, with the selected input variables added step by step to examine their influence on the representation of photovoltaic energy yield. This makes it possible to assess the contribution of each variable and determine an appropriate model structure. The main stages in constructing and evaluating the model are shown in Fig. 2.

5.1. Initial Model with Solar Irradiance

The initial model examines the relationship between solar irradiance and photovoltaic energy yield. Solar irradiance is selected as the first input variable because of its primary influence on the amount of energy that can be converted by photovoltaic modules. Chebyshev polynomials of different degrees are calculated for the normalized solar irradiance values. The resulting model serves as a baseline for comparison with models developed by subsequently adding further input variables.

This stage shows how well solar irradiance alone describes variations in photovoltaic energy yield and helps determine whether additional input variables are needed.

5.2. Inclusion of Air Temperature

Air temperature is then incorporated into the model because it affects the operating characteristics and efficiency of photovoltaic modules. Chebyshev polynomials of different degrees are calculated for temperature and combined with those for solar irradiance, allowing the model to account for variations in both variables. The results are compared with those of the initial model to determine whether including temperature improves the representation of photovoltaic energy yield and to assess its contribution to the model.

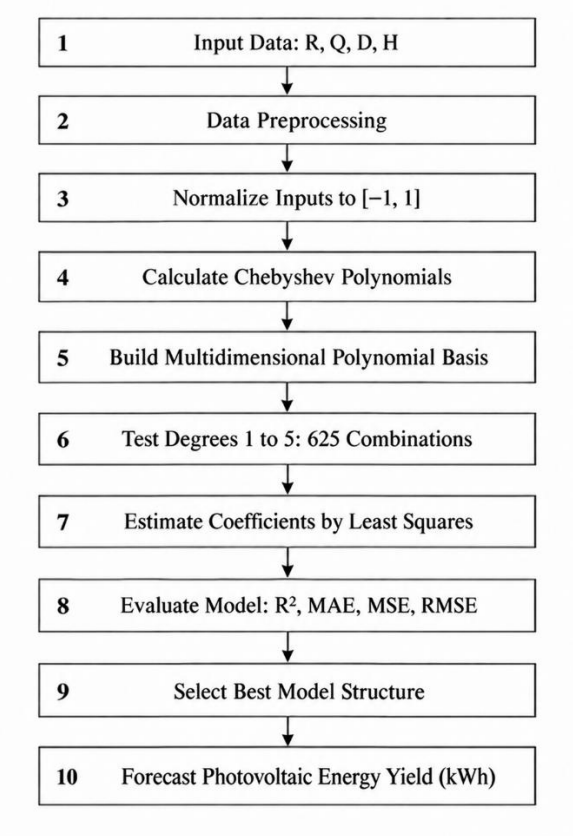


Fig. 2. Main stages in constructing the photovoltaic energy-yield model

5.3. Inclusion of Seasonal Variations

The model is then extended by adding the day of the year to represent seasonal variations in photovoltaic energy yield. Daylight duration, the Sun's position, and typical meteorological conditions vary throughout the year. Including this variable enables the model to account for these seasonal patterns. Chebyshev polynomials of different degrees are calculated for the day of the year and combined with those for solar irradiance and air temperature.

5.4. Inclusion of Time of Day

Time of day is also included among the input variables to account for the characteristic variation in photovoltaic energy yield throughout the day. After sunrise, power output gradually increases, reaches higher levels during periods of stronger solar irradiance, and decreases toward the end of daylight. Therefore, hourly information helps the model represent the hourly energy-yield profile more accurately.

The daily model uses three main input variables - solar irradiance, air temperature, and day of the year - whereas the hourly model additionally includes time of day.

5.5. Selection of Polynomial Degree

Selecting the degree of the Chebyshev polynomials is an important step in constructing the model. The degree determines the complexity of the polynomial representation and its ability to describe variations in the relationship between the input variables and photovoltaic energy yield. Lower-degree polynomials represent the general pattern of the relationship, while higher degrees can capture more complex nonlinear variations.

However, an excessively high degree may unnecessarily complicate the model and cause it to overfit the data used for its construction. Therefore, the degree is not fixed in advance. Several polynomial degrees are evaluated in sequence, and the resulting models are compared.

When selecting the polynomial degrees, degrees from 1 to 5 are considered. For each input variable, the analysis examines whether increasing the polynomial degree improves the representation of photovoltaic energy yield. For the daily model, five possible degrees are evaluated for each of the three input variables, resulting in 5^3=125-degree combinations. For the hourly model, five possible degrees are evaluated for each of the four input variables, resulting in 5^4=625-degree combinations.

The input variables do not need to have the same polynomial degree. For example, the relationship with solar irradiance may be represented best by one degree, while the effects of temperature, seasonal variations, or time of day may require another. Therefore, the maximum degree is selected separately for each input variable, and all 625 combinations are compared using accuracy and complexity measures. The principle for generating and evaluating degree combinations for the four input variables is shown in Fig. 3.

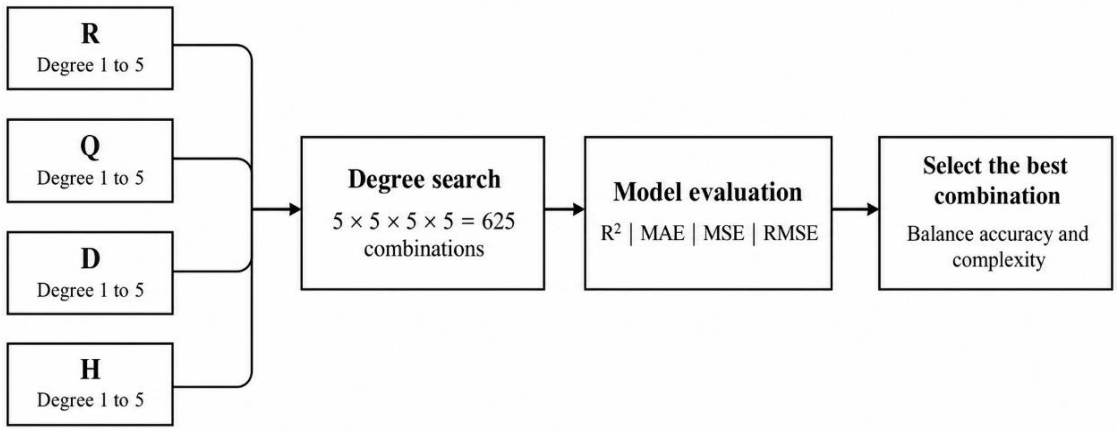


Fig. 3. Formation and Evaluation of Chebyshev Polynomial Degree Combinations

Different combinations of polynomial degrees are investigated, and the resulting models are compared in terms of accuracy. The aim is to select a model that describes the observed relationships adequately without increasing their complexity unnecessarily.

The final polynomial degrees are selected after evaluating the results using appropriate accuracy metrics. Thus, the degrees are not predetermined but are selected based on the model's performance on the available data.

5.6. Formation of the Overall Model Structure

After the input variables have been identified and suitable polynomial degrees have been investigated, the overall model structure is formed. For each normalized input variable, the corresponding Chebyshev polynomials are used to construct the polynomial basis of mathematical representation. Combining these polynomials allows the model to represent both the individual effects of the input variables and their interactions. The daily model uses solar irradiance, air temperature, and day of the year, whereas the hourly model also includes time of day. The model coefficients are estimated from the corresponding photovoltaic energy-yield data using the least-squares method.

5.7. Model Comparison and Selection

The model variants are compared to assess how the inclusion of additional input variables and the selection of different polynomial degrees affect the results. The analysis considers the contribution of solar irradiance, air temperature, day of the year, and time of day, as well as different combinations of Chebyshev polynomial degrees.

Model performance is evaluated by comparing the actual and predicted electrical energy generated by the photovoltaic system using the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) [17–19]. The daily and hourly model configurations are evaluated separately because they use different numbers of input variables and therefore involve different sets of polynomial-degree combinations. These metrics are used to identify a model structure that describes the observed relationships accurately while maintaining an appropriate level of complexity [20]. The daily model uses three input variables, whereas the hourly model uses four. The maximum polynomial degree is determined separately for each input variable. The resulting model configurations provide a basis for forecasting photovoltaic energy yield.

6. Conclusion

Photovoltaic energy yield is shaped by the combined effects of various meteorological and temporal factors. Solar irradiance is the primary energy input, while temperature, seasonal variations, and time of day can also affect the amount of electrical energy produced.

This paper identifies the main input variables used to develop a mathematical model based on orthogonal Chebyshev polynomials. It outlines the main stages of data preparation, polynomial basis construction, selection of polynomial degrees, and model evaluation. The contribution of this paper is the formulation of a structured methodology for representing the relationship between photovoltaic energy yield and meteorological variables using a multidimensional basis of first-kind Chebyshev polynomials. The methodology specifies the input variables, data-preprocessing steps, polynomial-degree search, and model-evaluation procedure.

The next stage of the research will focus on validating the proposed model using real data to assess its applicability and accuracy.

Acknowledgment

The authors gratefully acknowledge the support provided by the project CoE UNITe BG16RFPR002-1.014-0004 funded by PRIDST, co-funded by the European Union.

References

[1] Iheanetu K. J. Solar Photovoltaic Power Forecasting: A Review. Sustainability, Vol. 14, No. 24, 2022, Article 17005, ISSN 2071-1050. DOI: https://doi.org/10.3390/su142417005.

[2] Diagne M., David M., Lauret P., Boland J., Schmutz N. Review of Solar Irradiance Forecasting Methods and a Proposition for Small-Scale Insular Grids. Renewable and Sustainable Energy Reviews, Vol. 27, 2013, pp. 65–76, ISSN 1364-0321. DOI: https://doi.org/10.1016/j.rser.2013.06.042.

[3] Antonanzas J., Osorio N., Escobar R., Urraca R., Martinez-de-Pison F. J., Antonanzas-Torres F. Review of Photovoltaic Power Forecasting. Solar Energy, Vol. 136, 2016, pp. 78–111, ISSN 0038-092X. DOI: https://doi.org/10.1016/j.solener.2016.06.069.

[4] Mellit A., Massi Pavan A., Ogliari E., Leva S., Lughi V. Advanced Methods for Photovoltaic Output Power Forecasting: A Review. Applied Sciences, Vol. 10, No. 2, 2020, Article 487, ISSN 2076-3417. DOI: https://doi.org/10.3390/app10020487.

[5] Simeonova, I. Improving Administrative Security in Municipal Administrations: A Comprehensive Approach Using the Monte Carlo Model. Yearbook of Konstantin Preslavski University of Shumen, Vol. XIII E, Faculty of Technical Sciences, 2023, pp. 219–223, ISSN 1311-834.

[6] Trifonov T. Modeling and Calculation of Passive Audio Crossovers. Annual of Konstantin Preslavski University of Shumen, Vol. IX E, Technical Sciences, 2019, pp. 182–189, ISSN 1311-834X. Available at: http://annuals.ftn.shu.bg/files/godishnik_ftn_2019.pdf.

[7] Al-Dahidi S., Madhiarasan M., Al-Ghussain L., Abubaker A. M., Ahmad A. D., Alrbai M., Aghaei M., Alahmer H., Alahmer A., Baraldi P., Zio E. Forecasting Solar Photovoltaic Power Production: A Comprehensive Review and Innovative Data-Driven Modeling Framework. Energies, Vol. 17, No. 16, 2024, Article 4145, ISSN 1996-1073. DOI: https://doi.org/10.3390/en17164145.

[8] Das U. K., Tey K. S., Seyedmahmoudian M., Mekhilef S., Idris M. Y. I., Van Deventer W., Horan B., Stojcevski A. Solar Photovoltaic Generation Forecasting Methods: A Review. Energy Conversion and Management, Vol. 156, 2018, pp. 459–497, ISSN 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2017.11.019.

[9] Chodakowska E., Nazarko J., Nazarko Ł., Rabayah H. S. Solar Radiation Forecasting: A Systematic Meta-Review of Current Methods and Emerging Trends. Energies, Vol. 17, No. 13, 2024, Article 3156, ISSN 1996-1073. DOI: https://doi.org/10.3390/en17133156.

[10] Voyant C., Notton G., Kalogirou S. A., Nivet M. L., Paoli C., Motte F., Fouilloy A. Machine Learning Methods for Solar Radiation Forecasting: A Review. Renewable Energy, Vol. 105, 2017, pp. 569–582, ISSN 0960-1481. DOI: https://doi.org/10.1016/j.renene.2016.12.095.

[11] Dubey S., Sarvaiya J. N., Seshadri B. Temperature Dependent Photovoltaic (PV) Efficiency and Its Effect on PV Production in the World – A Review. Energy Procedia, Vol. 33, 2013, pp. 311–321, ISSN 1876-6102. DOI: https://doi.org/10.1016/j.egypro.2013.05.072.

[12] Skoplaki E., Palyvos J. A. On the Temperature Dependence of Photovoltaic Module Electrical Performance: A Review of Efficiency/Power Correlations. Solar Energy, Vol. 83, No. 5, 2009, pp. 614–624, ISSN 0038-092X. DOI: https://doi.org/10.1016/j.solener.2008.10.008.

[13] Mohamad Radzi P. N. L., Akhter M. N., Mekhilef S., Mohamed Shah N. Review on the Application of Photovoltaic Forecasting Using Machine Learning for Very Short- to Long-Term Forecasting. Sustainability, Vol. 15, No. 4, 2023, Article 2942, ISSN 2071-1050. https://doi.org/10.3390/su15042942.

[14] Inman R. H., Pedro H. T. C., Coimbra C. F. M. Solar Forecasting Methods for Renewable Energy Integration. Progress in Energy and Combustion Science, Vol. 39, No. 6, 2013, pp. 535–576, ISSN 0360-1285. https://doi.org/10.1016/j.pecs.2013.06.002.

[15] Xu K. The Chebyshev Points of the First Kind. Applied Numerical Mathematics, Vol. 102, 2016, pp. 17–30, ISSN 0168-9274. DOI: https://doi.org/10.1016/j.apnum.2015.12.002.
[16] Niu C., Liao H., Ma H., Wu H. Approximation Properties of Chebyshev Polynomials in the Legendre Norm. Mathematics, Vol. 9, No. 24, 2021, Article 3271, ISSN 2227-7390. DOI: https://doi.org/10.3390/math9243271.
[17] Hyndman R. J., Koehler A. B. Another Look at Measures of Forecast Accuracy. International Journal of Forecasting, Vol. 22, No. 4, 2006, pp. 679–688, ISSN 0169-2070. DOI: https://doi.org/10.1016/j.ijforecast.2006.03.001.
[18] Chai T., Draxler R. R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments Against Avoiding RMSE in the Literature. Geoscientific Model Development, Vol. 7, No. 3, 2014, pp. 1247–1250, ISSN 1991-9603. DOI: https://doi.org/10.5194/gmd-7-1247-2014.
[19] Hodson T. O. Root-Mean-Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not. Geoscientific Model Development, Vol. 15, No. 14, 2022, pp. 5481–5487, ISSN 1991-9603. DOI: https://doi.org/10.5194/gmd-15-5481-2022.
[20] Di Leo P., Ciocia A., Malgaroli G., Spertino F. Advancements and Challenges in Photovoltaic Power Forecasting: A Comprehensive Review. Energies, Vol. 18, No. 8, 2025, Article 2108, ISSN 1996-1073. DOI: https://doi.org/10.3390/en18082108.