e-Prime - Advances in Electrical Engineering, Electronics and Energy 12 (2025) 101003

Contents lists available at ScienceDirect

e-Prime - Advances in Electrical Engineering, Electronics and Energy

journal homepage: www.elsevier.com/locate/prime

Smart forecasting: Enhancing virtual power plant performance with analytical frameworks

Maja Celeska Krstevska *, Petar Krstevski, Mare Srbinovska, Vesna Andova, Aleksandra Krkoleva Mateska

Ss Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies, 1000 Skopje, Republic of North Macedonia

ARTICLE INFO

Keywords:
Electricity production
Forecast
Solar radiation
Virtual power plant

ABSTRACT

This paper presents a comprehensive framework for efficient electricity forecasting within a Virtual Power Plant (VPP) context. The focus is on forecasting both electricity consumption and production from photovoltaic (PV) systems, as essential components for successful operation of a sustainable energy infrastructure/system. In order to investigate the applicability of the framework, a case study approach was employed. The aim was to evaluate the effectiveness of various statistical modeling techniques, including time series analysis and regression analysis. Three datasets were analyzed: (i) electricity consumption at Ss. Cyril and Methodius University in Skopje (UKIM) facilities, (ii) solar radiation data for Skopje city region, and (iii) historical PV power generation data. The study delved into the application of Auto Regressive Integrated Moving Average (ARIMA) models to forecast PV power generation from a 22 kW PV system. A comparative analysis with linear regression highlighted the superior performance of ARIMA models, particularly in capturing seasonal patterns and handling the inherent variability of solar radiation. ARIMA (1,0,5) model was selected as the most efficient, with highest accuracy during summer, confirmed by the Mean Absolute Error (MAE) of 1.219 and $R^2 = 0.899$, likely due to stable and high levels of solar radiation. The highest error is observed in winter with MAE of 1.594 and $R^2 = 0.631$, which is attributed to the lower and more variable solar radiation typical of this season.

The findings of this research contribute to the advancement of VPP management strategies by providing accurate and reliable electricity forecasts. These forecasts can be utilized to optimize energy scheduling, demand response, and improve the grid integration of renewable energy sources, ultimately leading to more efficient and sustainable energy systems.

* Corresponding author.
E-mail addresses: celeska@feit.ukim.edu.mk (M.C. Krstevska), petark@feit.ukim.edu.mk (P. Krstevski), mares@feit.ukim.edu.mk (M. Srbinovska), vesnaa@feit.ukim.edu.mk (V. Andova), krkoleva@feit.ukim.edu.mk (A.K. Mateska).

https://doi.org/10.1016/j.prime.2025.101003
Received 27 December 2024; Received in revised form 21 March 2025; Accepted 23 April 2025
Available online 24 April 2025

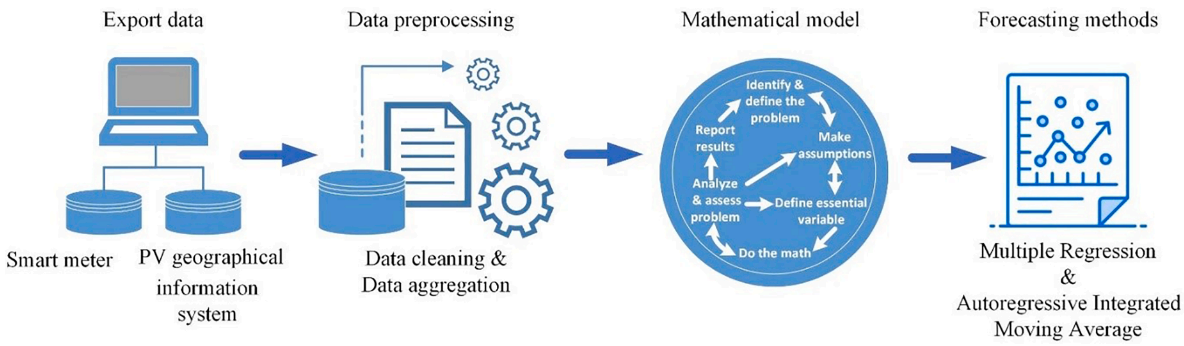


Diagram showing the workflow of the proposed methodology

1. Introduction

Currently, most countries are promoting policies to support the sustainability and growth of electricity generation from renewable energy sources. Their goal is to mitigate the effects of climate change by enabling the transition towards a reliable, sustainable, and cost-effective energy system [1]. A climate-friendly energy system is characterized by decentralization, digitalization, and electrification [2]. The digitalization of power systems can be achieved through gradual installation of advanced metering technologies, enabling bidirectional communication between grid users and system operators. The new technologies are essential for implementing decentralized control strategies and optimizing the use of available resources. The transformation of energy systems is further driven by the emergence of new players, such as demand aggregators, storage and virtual power plants (VPPs), which can help maintain the security and quality of electricity supply as renewable energy integration increases [3]. Flexibility within power systems is essential for integrating high levels of variable renewable electricity, as a key source of supply for the future energy grid. A mix of generation and storage technologies, including flexible generation systems, is vital to offset the intermittency of variable renewable energy sources [4]. Furthermore, the power grids, especially distribution grids, are faced with significant challenges due to large scale integration of distributed energy resources (DERs) [5]. Some of the DERs are connected at consumers' premises thus establishing new entities known as prosumers. The optimization of DERs' and prosumers‘ operation may be supported by establishment of microgrids within the distribution systems. By allowing isolated operation of microgrids during emergency states, the consumers can still be supplied by local DERs which increases the resilience of the system.

In the context of the previously described changes in the energy systems, VPPs are gaining importance because of the possibility to integrate various entities for providing flexibility services to the grid and receiving financial benefits for the VPP participants. The operation of VPPs is not limited to this scenario as their participation in the electricity markets has been the main business model for these entities. These activities require increased security of information exchange between VPP participants. The research presented in [6] introduces a novel framework for Virtual Power Plants (VPPs) to optimize DERs administration using blockchain technology. By integrating blockchain with key VPP components, such as cloud services, grid operators, and DER forecasting, the framework enhances real-time bidding, scheduling, and energy efficiency.

On one hand, the increasing share of variable energies in the power system requires new approaches to maintain its secure operation, but on the other hand renewable generators have the ability to support its

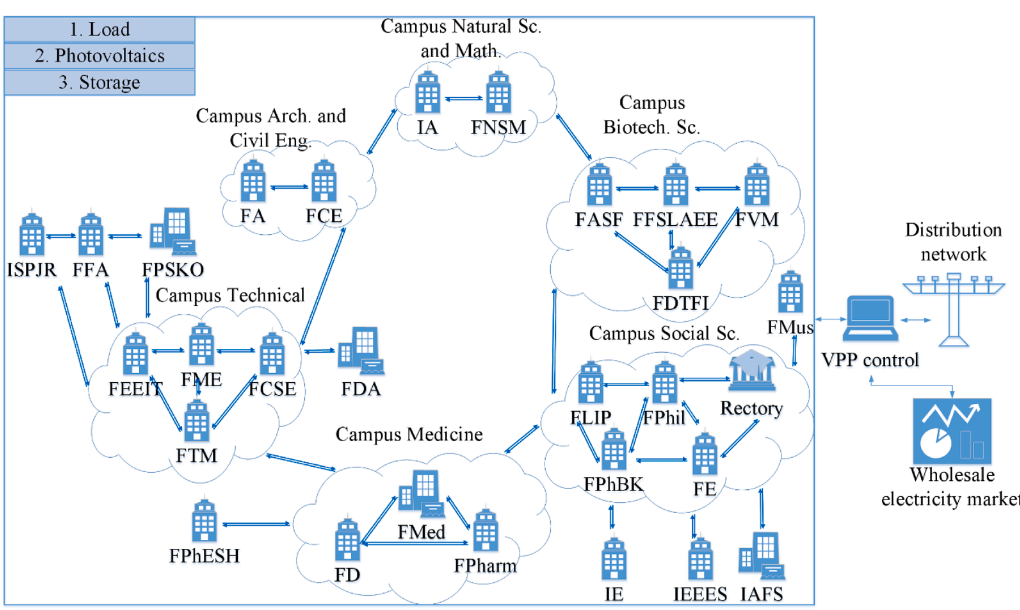


Fig. 1. Diagram of the VPP and its interaction with electricity markets and networks.

Fig. 2. PVPP at FEEIT, phase 1- 10 kWp installed power, positioned on roof area of building 1 and phase 2, 12 kWp installed power, positioned on roof area of building 2.

operation under emergency conditions, as described in [7]. The study presented in [8] reviews the role of renewable energy-based microgrid systems in promoting sustainability and reducing pollution, emphasizing their importance in transitioning away from fossil fuels. It highlights optimal design, advanced control strategies, and energy storage integration as critical for enhancing reliability, resilience, and economic efficiency. The study also explores emerging trends like AI and blockchain, showing a 30 % improvement in power reliability compared to traditional grids. Conversely, the paper [9] explores the electrification of transportation in Pakistan as a solution to severe air pollution and reliance on fossil fuels, analyzing policies, technical aspects, and standards. Drawing on best practices from developed nations and current government incentives, it identifies challenges, opportunities, and strategies for integrating electric vehicles into existing power systems. Through a SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis, the study offers insights to guide Pakistan and similar developing nations in promoting electric vehicle adoption to achieve environmental and energy goals.

The goal of a Virtual Power Plant (VPP) is to enable the joint participation of multiple energy-producing devices in electricity and flexibility markets. Popławski et al. [10] demonstrated this concept through a case study, highlighting the role of VPPs as data acquisition tools for PV energy forecasting. Similarly, Ullah and Mirjat [11] analyzed the interactive operational characteristics of VPPs in distributed energy systems, further emphasizing their role in market integration. Namely, apart from selling electricity VPPs can provide services to the grid operators utilizing the advanced control and communication systems, [12]. Despite numerous benefits in the concept of VPP (whether „technical“ or „commercial“), there is lack of knowledge in several aspects when dimensioning and establishing a VPP.

Forecasting electricity production is essential for PV plants to participate effectively in electricity markets and optimize resource planning. Fara et al. [13] demonstrated that advanced models such as ARIMA and artificial neural networks (ANN) improve forecasting accuracy. Başaran et al. [14] conducted a systematic review of photovoltaic output power forecasting techniques, identifying key trends in the field. Similarly, Sobri et al. [15] provided a comprehensive overview of solar photovoltaic generation forecasting methods, emphasizing their applicability in energy management.

However, forecast uncertainty increases with longer time horizons, presenting challenges for Virtual Power Plant (VPP) energy management, as noted by Popławski et al. [10]. Numerous forecasting approaches have been documented in the literature. Mellit et al. [16] categorized these into four main types: statistical models, machine learning techniques, physical models, and hybrid approaches. Statistical models, such as ARIMA, rely on historical data trends, as demonstrated in case studies by Atique et al. [17] and Kushwaha et al. [18]. Singh and Pozo [19] provide a methodology for developing ARMA models to forecast photovoltaic solar power output. Their work emphasizes the integration of such models into energy optimization frameworks, highlighting the importance of accurate forecasting in operational planning. Machine learning techniques, including ANN-based models, have been extensively explored by Alcañiz et al. [20] and Runge et al. [21]. Kempitiya et al. [22] propose an artificial intelligence-based framework for optimizing bidding strategies in multiple frequency reserve markets. They address the challenges of uncertainty in renewable energy and the need for data-driven decision-making in reserve procurement. Physical models, which utilize numerical weather predictions and satellite imagery, have been analyzed by Qing et al. [23] and Theocharides et al. [24], showcasing their potential for improving forecast accuracy. Shi et al. [25] develop a scenario-oriented approach to joint procurement and pricing of energy and reserves. Their model connects reserve procurement directly with potential uncertainties, aiming to minimize expected system costs while ensuring reliability. Finally, hybrid approaches, which integrate elements of multiple forecasting methods, have been proposed by Abumohsen et al. [26] and Khadke et al. [27], demonstrating enhanced predictive performance through combined methodologies. Fernández-Muñoz and Pérez-Díaz [28] present optimization models for the day-ahead scheduling of hybrid wind-battery virtual power plants. Nguyen et al. [29] propose a two-stage stochastic optimization model for VPPs participating in both day-ahead and balancing markets. They emphasize the role of energy storage systems in providing reserve services and managing the uncertainties of renewable energy outputs.

This paper introduces an analytical framework for electricity forecasting tailored to VPPs, focusing on a case study at Ss. Cyril and Methodius University in Skopje (UKIM), North Macedonia. The framework addresses the unique demands of VPPs and considers the local context of the university. The motivation for this design is threefold: the successful operation of the PV power plant (PVPP) at the Faculty of Electrical Engineering and Information Technologies (FEEIT); successful applications of VPPs in Western European countries; and UKIM’s commitment to the energy transition and global sustainability goals. The conceptual design of the VPP, depicted in Fig. 1, illustrates potential electrical pathways among the university’s 74 facilities.

A significant research gap remains in integrating forecasting techniques with VPP optimization strategies for joint electricity and reserve market participation. The study addresses the gap by proposing a forecasting framework tailored for VPPs, assessing its performance under seasonal variations, and discussing its applicability in market optimization. The key contributions of this paper include:

- A comparative analysis of ARIMA and regression models for PV power forecasting, demonstrating ARIMA's superior performance in seasonal adaptation.
- A case study analysis of a real-world VPP scenario, incorporating university electricity demand and regional solar radiation data.
- A discussion on integrating forecasting models into joint energy and reserve market strategies, highlighting the potential for enhanced VPP market participation.The paper is organized into four sections. The second section describes the measurement and data collection methods, as well as the statistical tools and evaluation criteria employed. The following section presents statistical analyses, including graphical representations of energy production forecasts. The final section provides the study's conclusions, followed by an acknowledgment section that recognizes the contributions and support received during the research.

2. Methodology

The methodology presented in this paper was tailored to the specific conditions for its application at the University, but it is replicable to similar cases. Therefore, in the next parts of this section the observed system elements and the data acquisition process are first described, followed by the applied statistical tools used in the procedure for forecasting PV generation in a VPP. The last subsection presents the used performances indexes used in the analysis and comparison of the statistical tools.

2.1. System description and data acquisition

This paper analyses data collected from a smart meter installed at FEEIT, along with online solar radiation data, to forecast electricity production from a 22 kW PV plant located on the faculty building roof (Fig. 1). At present FEEIT is the only unit at the University that has installed PV plant. In the future, the University VPP will consist of PV installations located on the rooftops of the various faculty buildings dispersed among six campuses. The VPP will be equipped with adequate communication and energy management systems to optimize the overall generation and consumption of the University.

The proposed VPP solution is considered to be cost effective because North Macedonia is situated in the temperate continental climate zone and the region experiences about 1400 kWh/m^2 of direct solar radiation per year. Solar radiation data was sourced from the Photovoltaic Geographical Information System, [30] and Solar Radiation Data, [31], to make short term forecast for electricity production from PVs in Skopje (Lat. 4159′46“, Lon. 2125′54“). These online sources provide hourly data for: Global radiation on the inclined plane [W/m^2], Sun height [^{o}], 2-m air temperature [°C] and 10-m total wind speed [m/s].

The PV system at FEEIT consists of Canadian Solar CS6K-280P modules, with a total installed capacity of 22 kW, as presented in Fig. 2. This small PV power plant (PVPP) has been operational since 2018. The system also includes an energy storage solution utilizing a Lithium Titanate Oxide (LTO) battery (Yjnlog, model LTO66160H/40 Ah). Energy storage charging is configured to draw power exclusively from the PVPP through the inverter. The inverter, a 3.3 kW three-phase Sunnlogical/Voltronics InfiniSolar 3P 10 kW model, is a versatile hybrid inverter that intelligently integrates solar power, AC utility, and battery sources to ensure a reliable power supply [32].

The smart electricity monitoring system includes both hardware (DTK smart tek, model SOMMY ES9L, Raspberry Pi 3) and software components. The system comprises of a measurement cabinet containing the necessary equipment and an internet-based software application that displays the user's electricity consumption data and automatically generates reports.

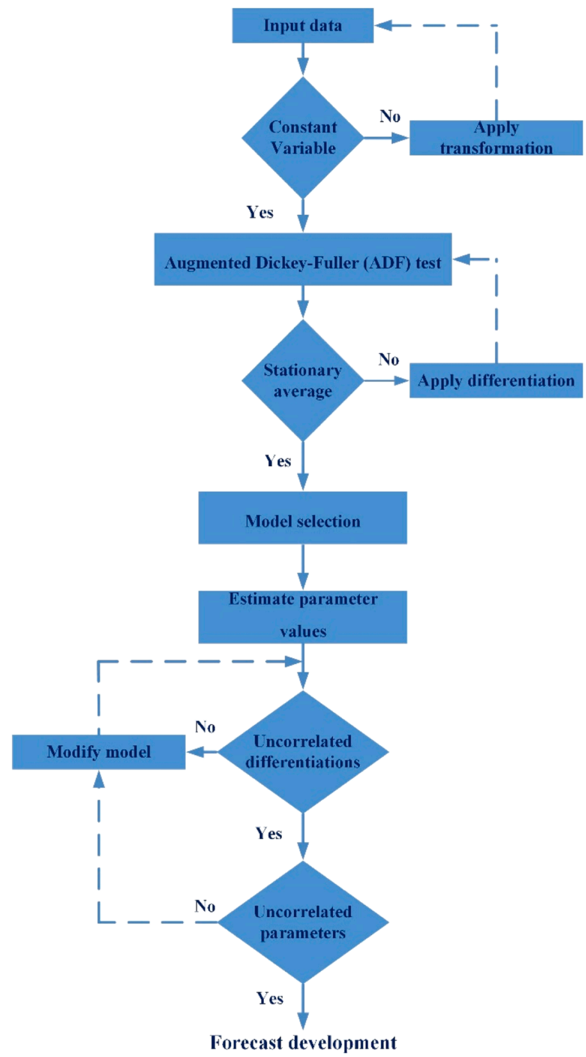


Fig. 3. Operation of the ARIMA model for the energy production forecast of a photovoltaic system.

2.2. Statistical tools

To analyze the collected data, a statistical time series forecasting model, i.e. Auto Regressive Integrated Moving Average (ARIMA), was employed. Data visualization was carried out using histograms and violin plots. The used dataset originates from two main sources: smart meter data and publicly available weather information. During the study period, data was collected in CSV format, then thoroughly organized and prepared for analysis. Electricity consumption data, initially recorded at 15-second intervals, was aggregated into hourly averages to improve manageability and relevance, calculated by averaging the readings recorded within each hour.

Fig. 3 illustrates the energy production forecast model developed for this study. The analysis starts with a comparison of the energy calculated based on weather station data and the measured energy generated by the PVPP. The objective of this study is to identify the most suitable forecast model for the PVPP installed at FEEIT in Skopje. This includes evaluating simulated and experimental results for short-term (one-day-ahead) electricity production forecasts, as well as long-term forecasts (seasonal analysis and three-day projections). Forecasting electricity production in the context of a PVPP holds significant commercial potential in the future.

The ARIMA model is selected for its simplicity but is applicable only to stationary time series. The "Autoregressive" (AR) component utilizes

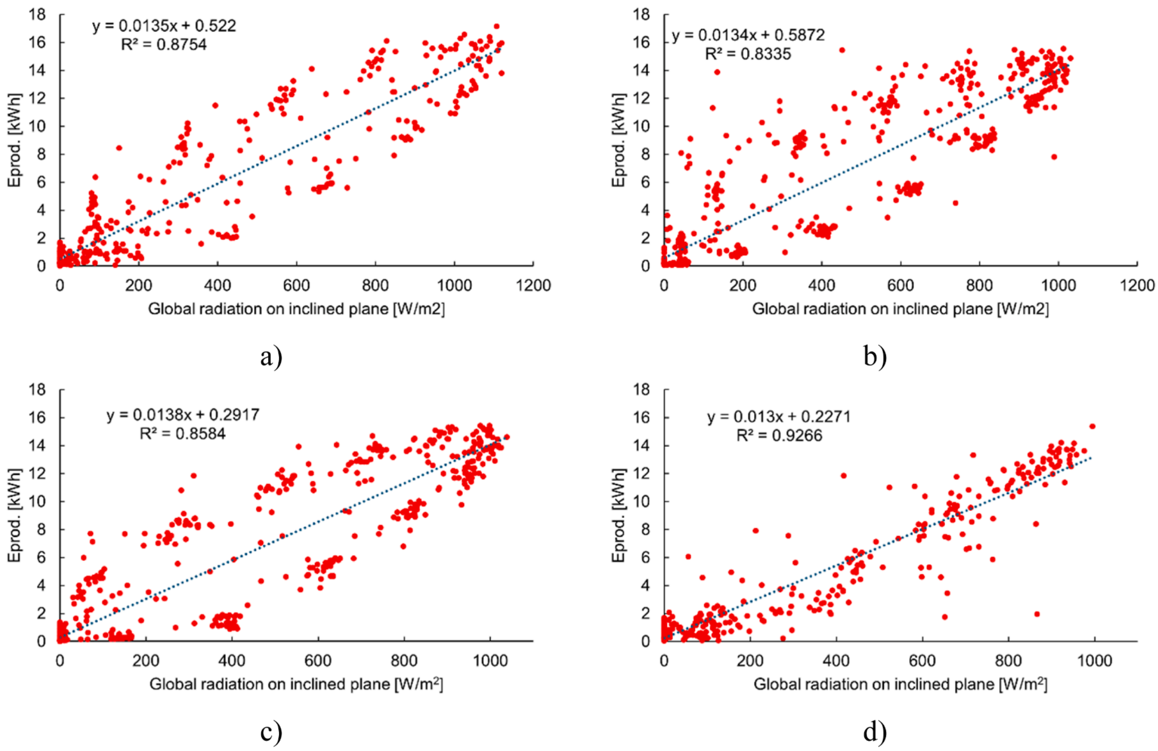


Fig. 4. Linear regression model of correlation between the average hourly solar radiation and the hourly PV plant e nergy production-$E_{prod.}$ for: a) Spring, b) Summer, c) Autumn, d) Winter, 2020.

the relationship between a current data point and its preceding values, using regression to predict (forecast) future values. The "Integrated" (I) component involves differencing the data points to achieve stationarity, a necessary condition for applying ARIMA. Finally, the "Moving Average" (MA) component models the error term as a linear combination of past error terms, allowing the model to adapt to sudden changes in the series [33]. ARIMA is a class of models that interprets a time series based on its own past values, including lags and lagged forecast errors, enabling the forecasting of future values [34]. Consequently, seasonal and non-stationary time series data in this study were transformed into stationary series to utilize the ARIMA model. Any non-seasonal time series that exhibits patterns, rather than random white noise, can be effectively modeled using ARIMA.

We are using the notation ARIMA(p,d,q), where:

p - the number of lag observations in the model, also known as the lag order,

d - the number of times the raw observations are differenced; also known as the degree of differencing, q - the size of the moving average window, also known as the order of the moving average, [35].

The term 'Auto Regressive' in ARIMA means it is a linear regression model that uses its own lags as predictors. Linear regression models work best when the predictors are not correlated and are independent of each other. The model is developed using advanced statistical techniques, with the best approach being chosen and validated through the Augmented Dickey-Fuller (ADF) test and the residual sum of squares.

2.3. Performance indexes

Selecting the best ARIMA(p,d,q) model involves a systematic approach to identify the optimal combination of parameters. Different statistics are commonly used to quantify the discrepancy, or error, between forecast and actual observation to evaluate the performance of a method, [36]. The common error definition for the assessment is the hourly error e_h, which is defined as:

$$e_h = P_{m,h} - P_{p,h} \tag{1}$$

where $P_{m,h}$ is the average actual power in the hour and $P_{p,h}$ is the prediction in the hour provided by the forecasting method [37,38]. Starting from the hourly error definition, the other error indexes used for the assessment can be derived:

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|P_{m,h,i} - P_{p,h,i}\right| \tag{2}$$

The MAE values for the different seasons indicate the average absolute difference between the predicted electricity production at time i, $P_{p,h,i}$, and actual electricity production at time i, $P_{m,h,i}$. Lower MAE values suggest better predictive accuracy.

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_{m,h,i} - P_{p,h,i}\right)^2} \tag{3}$$

RMSE provides a measure of the average magnitude of the prediction errors, with larger errors being more penalized due to squaring. Lower RMSE values indicate better model performance.

- R-squared

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(P_{m,h,i} - P_{p,h,i}\right)^2}{\sum_{i=1}^{N}\left(P_{m,h,i} - \overline{P_{m,h}}\right)^2} \tag{4}$$

Table 1
Parameters for assessing accuracy of the ARIMA model for PVPP energy production.

	Linear regression	ARIMA model		
	R^2 [/]	MAE [kWh]	RMSE [kWh]	R^2 [/]
Spring	0.875	1.404	2.034	0.893
Summer	0.834	1.219	1.776	0.899
Autumn	0.858	1.275	1.916	0.862
Winter	0.927	1.594	2.874	0.631

where $\overline{P_{m,h}}$ is the mean of the observed data. R^2 indicates the proportion of variance in the observed data that is explained by the model. Values closer to 1 indicate better model performance.

3. Results and discussion

In this section the results of our study on the electricity forecasting from the PVPP at FEEIT are presented from different perspectives.

To enhance the accuracy of our forecasts for electricity production, firstly linear regression model was applied over characteristic days in each season. This initial analysis was performed to understand the correlation between historical data and forecasted electricity production values for spring, summer, autumn, and winter. The linear regression models served as a preliminary step before applying more sophisticated ARIMA models for a more detailed forecast. The linear regression analysis provided insights into the seasonal variations and trends in electricity production. The R^2 values for each season's regression analysis indicate the proportion of the variance in the dependent variable that is predictable from the independent variable. The results are visualized in the regression graphs in Fig. 4, where each graph represents the linear fit for one of the four seasons. The R^2 values given in Table 1 suggest the strength of the linear relationship between the variables, with higher R^2 values indicating a better fit. Notably, the R^2 value of 0.927 in winter indicates a very strong linear relationship, while the summer value of 0.834, though slightly lower, still represents a robust model fit. The regression graphs presented in Fig. 4 are a clear visualization of these relationships and form the basis for further analysis using ARIMA models to produce the final forecasts for electricity production. Correlation with electricity production, with other meteorological data was also investigated by lasso regression, [39] but did not show any statistical significance.

By comparing the linear regression results with those of the ARIMA model, we can highlight the improvements in forecast accuracy achieved through more sophisticated time-series analysis. The ARIMA model's ability to capture temporal dependencies and seasonal patterns offers a more delicate understanding and prediction of electricity production, especially during seasons with higher variability in solar radiation.

Short-term solar radiation forecasts were developed using the ARIMA. Three days were considered for the forecasting process. The statistics are significant in the case of ARIMA (1,0,5), this being the reason why this variant was chosen for analysis. By comparing the results of the forecast with the measured values it was noticed that the ARIMA (1,0,5) model is more efficient than the ARIMA (1,0,4) and (1,0,3). The ARIMA (1,0,5) model was chosen after conducting several iterations. The statistical test used to determine the most suitable model was the Augmented Dickey-Fuller test. In the context of ARIMA, this test is applied to the residuals of a fitted model rather than the original series, ensuring that these residuals do not exhibit autocorrelation. Fig. 5 illustrates the ARIMA model's fit, showing the comparison between measured data (blue) and forecasted data (red), along with 95 % confidence intervals, for three representative days in each season. These "characteristic" days were selected based on the values of global radiation on an inclined plane, solar elevation angle, 2-m air temperature, and 10-m total wind speed for each season. Table 1 presents the results for three performance metrics—MAE, RMSE, and R^2—evaluating the ARIMA forecasts under four scenarios.

The seasonal analysis of MAE values reveals that the ARIMA model's predictive accuracy varies across the year. The model performs best during summer, with an MAE of 1.219, likely due to stable and high levels of solar radiation. In contrast, the largest error occurs in winter

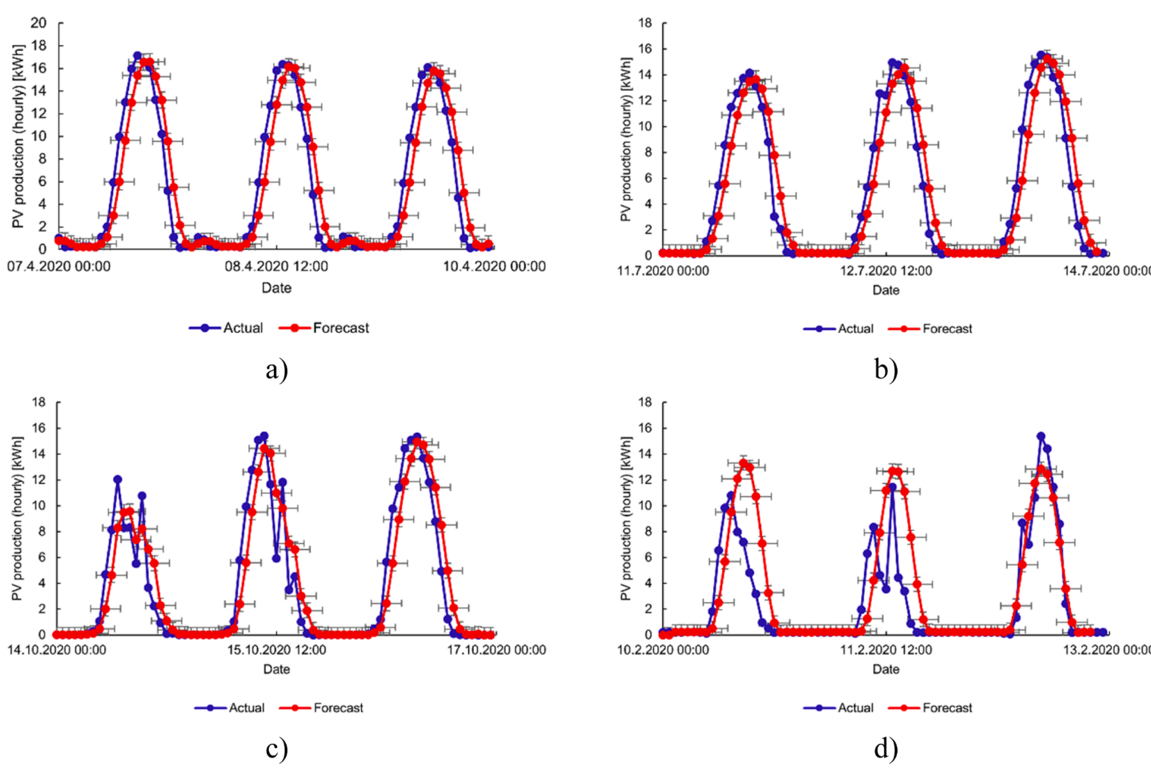


Fig. 5. The PVPP energy production forecast for 3 days in each season: a) Spring, b) Summer, c) Autumn, d) Winter, 2020.

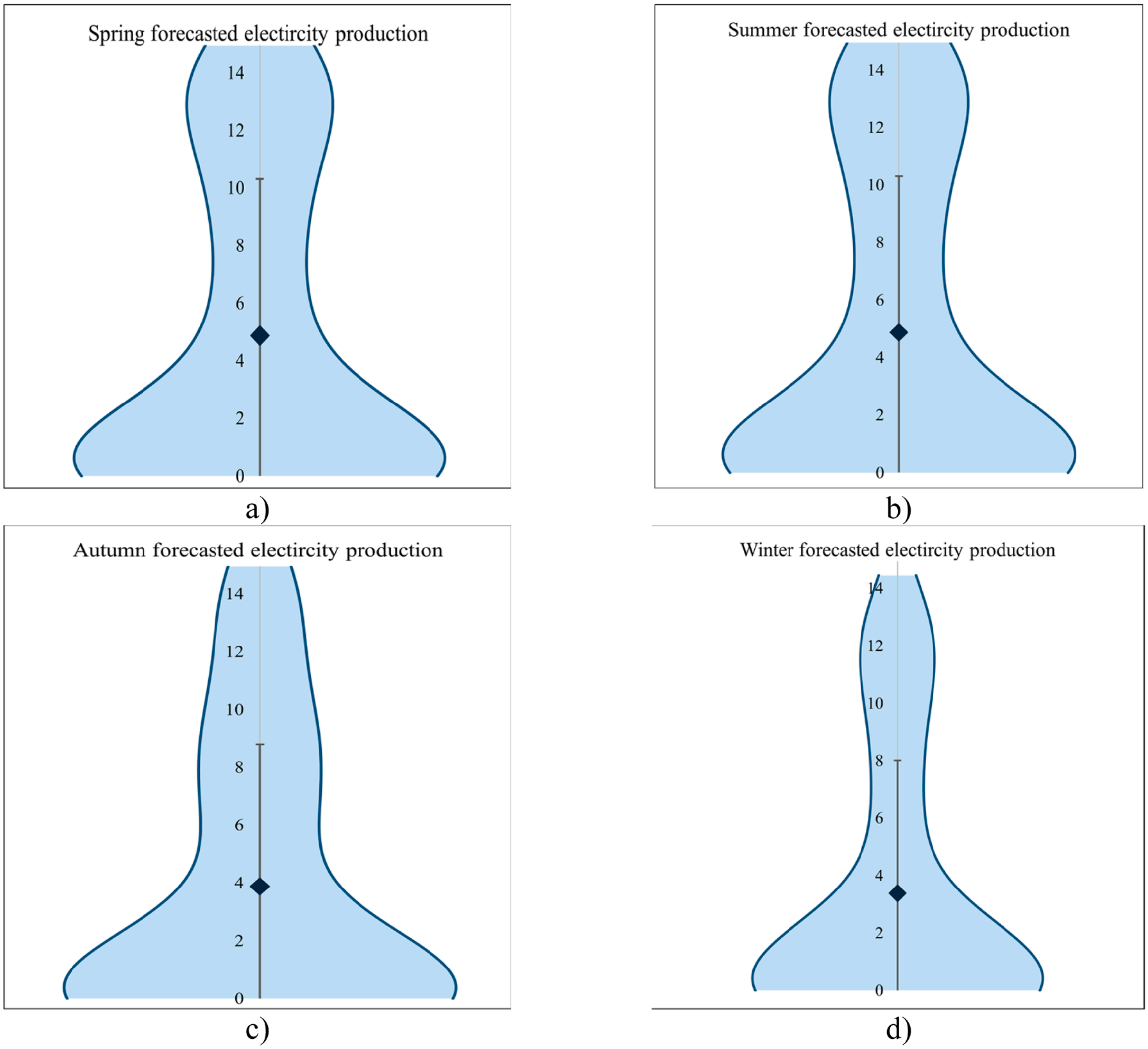


Fig. 6. Violin plots showing the distribution of forecasted electricity production for whole season: a) Spring, b) Summer, c) Autumn, d) Winter, 2020.

(MAE of 1.594), which is attributed to the lower and more variable solar radiation typical of this season. These findings indicate that while the ARIMA model is generally effective, its performance is influenced by seasonal factors.

The analysis of RMSE and R^2 values further exemplifies the performance of the ARIMA model across different seasons. In summer, the model exhibits the lowest RMSE (1.776) and the highest R^2 (0.899), indicating excellent predictive accuracy due to stable and high solar radiation. Conversely, winter presents the greatest challenge, with the highest RMSE (2.874) and the lowest R^2 (0.631), reflecting the model's difficulty in accurately predicting electricity production under low and variable solar radiation conditions. Spring and autumn show intermediate results, with RMSE values of 2.034 and 1.916, and R^2 values of 0.893 and 0.862, respectively. These results suggest that while the ARIMA model performs well in these seasons, the transitional weather conditions introduce some variability. These findings underscore the importance of considering seasonal variations in the predictive accuracy of the ARIMA model for the VPP. To enhance forecasting reliability, particularly in winter, incorporating additional weather-related variables or employing more sophisticated modelling approaches may be necessary.

The violin charts show the density of the forecasted electricity production probability. The wider sections of the violin plot indicate a higher probability of electricity production at those values, while narrower sections represent lower probability, [40]. As it can be seen in Fig. 6, charts are relatively symmetrical, indicating a balanced distribution of forecasted values around the mean for each season. If the seasonal variations are analyzed, the forecasted electricity production has a relatively uniform distribution for spring, with a central peak around the median value (~6 kWh). The distribution is slightly flatter at the extremes. For summer, the distribution is more spread out compared to spring, with a peak around the median value (~6 kWh). The wider distribution suggests more variability in forecasted production, which is expected in summer due to longer daylight hours and potentially varying cloud cover. When investigating the violin chart for autumn, the distribution is narrower, with a pronounced peak around 4 kWh. This suggests that the forecasted electricity production in autumn is more consistent and lower compared to other seasons. Also, for winter, the distribution is narrower and more concentrated around lower values, with a peak around 4 kWh. This is typical for winter due to shorter daylight hours and lower solar radiation. Each chart includes a marker (likely representing the median or mean value) along the centre line. In spring and summer, this marker is slightly higher than in autumn and winter, reflecting the expected seasonal variations in solar radiation and resulting electricity production.

Our study aligns with prior research on PV power forecasting, particularly those employing statistical and machine learning models [8, 9]. Studies such as [10] demonstrate that hybrid models combining ARIMA with machine learning techniques achieve higher accuracy. However, our research highlights the importance of seasonal adaptation, as ARIMA performed well in summer but struggled in winter due to solar radiation variability.

In comparison with [11], which focused on forecasting for standalone PV systems, our study extends the scope to VPP applications, considering joint optimization of electricity and reserve markets. Research in [12] discusses the role of forecasting in demand-side management, which complements our findings on optimizing VPP market participation. Furthermore, studies on reserve market optimization [13, 14] emphasize the need for accurate short-term forecasting, a critical aspect addressed in our proposed framework.

The results in this paper suggest that accurate forecasting can significantly impact VPP market strategies. As highlighted in [15], joint energy and reserve markets require precise short-term predictions to optimize bidding strategies. Our study provides insights into how forecasting techniques can be integrated into these market mechanisms. For instance, the accuracy improvements demonstrated by ARIMA models in stable solar conditions can be leveraged to optimize reserve scheduling. Conversely, incorporating external variables such as temperature and cloud cover, as suggested in [16], could enhance winter forecasting accuracy. Future research could explore hybrid models combining ARIMA with deep learning techniques to address these limitations.

4. Conclusion

This paper highlights the successful application of ARIMA models for forecasting electricity production from a 22 kW PVPP installed on the roof of a faculty building. Among the configurations tested, the ARIMA (1,0,5) model proved to be the most effective, demonstrating statistical significance and strong performance.

An initial linear regression analysis served as a baseline, revealing significant seasonal patterns in the data. The R^2 values indicated strong fits across all seasons, with the highest in winter (0.927) and the lowest in summer (0.834). This showed that linear regression effectively captured much of the variance in electricity production. However, the ARIMA model provided a more advanced approach, leveraging its capacity to account for temporal dependencies and the dynamic nature of solar radiation. It outperformed linear regression in spring ($R^2 = 0.893$) and summer ($R^2 = 0.899$), underscoring the importance of selecting models that can handle the complexities of time series data.

Interestingly, the ARIMA model's performance in winter ($R^2 = 0.631$) was lower than that of the linear regression model ($R^2 = 0.927$). This is likely due to the increased variability and complexity of solar radiation during winter, which posed challenges for the ARIMA model. These findings suggest that while ARIMA generally offers a more robust forecasting method, it may require further refinement or supplemental techniques to better address periods of high variability and low solar radiation.

In the context of statistical methods like ARIMA, appropriate data preprocessing can be fundamental to achieving good performance and reducing computational costs. While this study did not remove nighttime data samples, such an approach could be considered in future work, especially since these periods typically show no power production. Focusing on periods of actual energy generation could potentially enhance model accuracy and efficiency. Additionally, handling any missing data points through interpolation or imputation ensures the continuity and reliability of the time series data.

Violin charts provided a visual confirmation of the ARIMA analysis, highlighting the expected seasonal variability in electricity production. Specifically, summer exhibited the greatest potential variability, while winter showed the least. These observations are consistent with the ARIMA model's findings, further supporting its suitability for this type of forecasting.

In summary, while linear regression models can offer useful preliminary insights, more advanced approaches like ARIMA are crucial for achieving accurate and dependable forecasts of solar electricity production. The findings of this study provide valuable guidance for energy management decisions within the Virtual Power Plant (VPP) framework, particularly in planning for seasonal variations. Future research could focus on incorporating additional variables or exploring other advanced time series models to improve forecast accuracy, especially during challenging seasons such as winter.

CRediT authorship contribution statement

Maja Celeska Krstevska: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Petar Krstevski:** Writing – review & editing, Supervision. **Mare Srbinovska:** Writing – review & editing, Supervision, Methodology. **Vesna Andova:** Writing – review & editing, Supervision, Formal analysis, Data curation. **Aleksandra Krkoleva Mateska:** Writing – review & editing, Validation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the Ss. Cyril and Methodius University in Skopje for the support in facilitating the project „Development and Simulation of a Virtual Power Plant for Sustainable Energy Management at UKIM" without whom this research work wouldn't have been carried out.

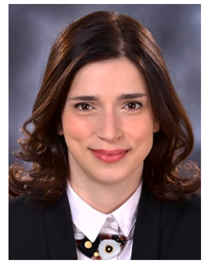
Data availability

The authors do not have permission to share data.

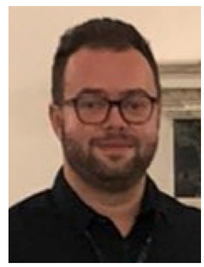
References

[1] P. Forster, t. Storelvmo, K. Armour, W. Collins, J.-L. Dufresne, D. Frame, D. Lunt, T. Mauritsen, M. Palmer, The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.
[2] I. R. E. A. IRENA, World Energy Transitions - Outlook 2023, 2023. IRENA, Abu Dhabi.
[3] N. Naval, J. Yusta, Virtual power plant models and electricity markets - A review, Renew. Sustain. Energy Rev. (20 June 2021).
[4] P. Romanos, A. Kindi, A. Pantaleo, C. Markides, Flexible nucelaer plants with termal energy storage and secondary power cycles: virtual power plant integration in a UK energy storage system case study, e-Prime - Adv. Electr. Eng. Electr. Energy 2 (2022) 100027.
[5] R. Trivedi, S. Patra, Y. Sidqi, B. Bowler, F. Zimmermann, G. Deconinck, A. Papaemmanouil and S. Khadem, "Community-based microgrids: literature review and pathways to decarbonise the local electricity network," 10.3390/en15030918, vol. 15 (3), no. 918, 2022.
[6] D. Kaif, k. Alam, K. Das, Blockchain based sustainable energy transition of a virtual power plant: conceptual framework design & experimental implementation, Energy Rep. 11 (2024) 261–275.
[7] M. Jasiński, T. Sikorski, D. Kaczorowska, J. Rezmer, V. Suresh, Z. Leonowicz, P. Kostyla, J. Szymańda, P. Janik, A case study on power quality in a virtual power plant: long term assessment and global index application, Energies 13 (6578) (2020).
[8] V. Khare, P. Chaturvedi, Design, control, reliability, economic and energy management of microgrid: a review, e-Prime - Adv. Electr. Eng. Electr. Energy 5 (100239) (2023).
[9] A. Masood, S. Hassan, T. Kamal, S. Khan, S. Rizvi, S. Salman, Electrification of transportation: policy framework, technical aspects and challenges in Pakistan - A case study, e-Prime - Adv. Electr. Eng. Electr. Energy 10 (100803) (2024).
[10] T. Popławski, S. Dudzik, P. Szelag, J. Baran, A case study of a virtual power plant (VPP) as a data acquisition tool for PV energy forecasting, Energies 14 (6200) (2021).
[11] Z. Ullah, N.H. Mirjat, Modelling and analysis of virtual power plants interactive operational characteristics in distributed systems, Energy Convers. Econ. (2021).
[12] P. Pal, P. Krishnamoorthy, D. Rukmani, S. Antony, S. Ocheme, U. Subramanian, R. Elavarasan, N. Das, H. Hasanien, Optimal dispatch strategy of virtual power plant for day-ahead market framework, Appl. Sci. 11 (3814) (2021).
[13] L. Fara, A. Diaconu, D. Craciunescu, S. Fara, Forecasting of energy production for photovoltaic systems based on ARIMA and ANN advanced models, Int. J. Photoenergy (2021).

[14] K. Başaran, F. Bozyiğit, P. Siano, P.Y. Taşer, D. Kılınç, Systematic literature review of photovoltaic output power forecasting, IET Renew. Power Gener. (26 January 2021) 3961–3973.
[15] S. Sobri, S. Koohi-Kamali, N.A. Rahim, Solar photovoltaic generation forecasting methods: a review, Energy Convers. Manag. 156 (15 January 2018) 459–497.
[16] A. Mellit, P.A.E. Ogliari, S. Leva, V. Lughi, Advanced methods for photovoltaic output power forecasting: a review, Appl. Sci. 487 (10) (9 January 2020).
[17] S. Atique, S. Noureen, V. Roy, V. Subburaj, S. Bayne, J. Macfie, Forecasting of total daily solar energy generation using ARIMA: a case study, in: IEEE 9th Annual Computing and Communication Workshop and Conference, 2019. Las Vegas, NV, USA.
[18] V. &. P. N. M. Kushwaha, Very short-term solar PV generation forecast using SARIMA model: a case study, in: 7th International Conference on Power Systems (ICPS), 2018. Pune, India.
[19] B. Singh, D. Pozo, A guide to solar power forecasting using ARMA models. IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 2019. Bucharest, Romania.
[20] A. Alcañiz, D. Grzebyk, H. Ziar, O. Isabella, Trends and gaps in photovoltaic power forecasting with machine learning, Energy Rep. 9 (December 2023) 447–471.
[21] J. &. Z. R. Runge, Forecasting energy use in buildings using artificial neural networks: a review, Energies 12 (17) (23 August 2019).
[22] T. Kempitiya, S. Sierla, D. De Silva, M. Yli-Ojanperä, A. D, V. Vyatkin, An Artificial intelligence framework for bidding optimization with uncertainty in multiple frequency reserve markets, Appl. Energy 115918 (2020).
[23] X. &. N. Y. Qing, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy 148 (April 2018) 461–468.
[24] S. Theocharides, A. Koumis, G. Makrides, G.E. Georghiou, Evaluating the weather forecasting models and the impact to PV generation forecasting, in: IEEE 50th Photovoltaic Specialists Conference (PVSC), 2023. San Juan, PR, USA.
[25] J. Shi, Y. Guo, W. Wu, H. Sun, Scenario-oriented multi-area joint market clearing of energy and reserve under uncertainty, Appl. Energy 361 (122873) (2024).
[26] M. Abumohsen, A. Owda, M. Owda, A. Abumihsan, Hybrid machine learning model combining of CNN-LSTM-RF for time series forecasting of Solar Power Generation, e-Prime - Adv. Electr. Eng. Electr. Energy (2024) 100636.
[27] S. Khadke, B. Ramasubramanian, P.e.a. Paul, Predicting active solar power with machine learning and weather data, Mater. Cir. Econ. 5 (15) (04 October 2023).
[28] D. Fernández-Muñoz, J. Pérez-Díaz, Optimisation models for the day-ahead energy and reserve self-scheduling of a hybrid wind–battery virtual power plant, Daniel Fernández-Muñoz, Juan I. Pérez-Díaz 57 (106296) (2023).
[29] A. Shayegan-Rad, A. Badri, A. Zangeneh, Day-ahead scheduling of virtual power plant in joint energy and regulation reserve markets under uncertainties, Energy 121 (2017) 114–125.
[30] E. Comission, Photovoltaic Geographical Information System [Online]. Available: https://re.jrc.ec.europa.eu/pvg_tools/en/, 03 January 2022 [Accessed 29 April 2024].
[31] S.E.S.f. Professionals, Monthly means of solar irradiance, temperature and relative humidity, SoDa Service (2014) [Online]. Available: https://www.soda-pro.com/web-services/meteo-data/monthly-means-solar-irradiance-temperature-relative-humidity [Accessed 29 April 2024].
[32] M. Celeska, M. Srbinovska, V. Andova, P. Krstevski, A. Krkoleva, Energy production forecast for the purpose of a technical virtual power plant. MedPower 2024, 2024. Athens, Greece.
[33] R. Djimasbe, A. Gyamfi, C. Iwei, B. RIbar, Development of an ARIMAX model for forecasting airport electricity consumption in Accra-Ghana: the role of weather and air passenger traffic, e-Prime- Adv. Electr. Eng. Electr. Energy 9 (2024) 100691.
[34] A. Tomar, P. Gaur, X. Jin, Prediction Techniques for Renewable Energy Generation and Load Demand Forecasting, Springer Nature Singapore, 2023.
[35] A.W. Woodward, B. Sadler, S. Robertson, Time Series for Data Science (Analysis and Forecasting), CRC Press, United Kingdom, 2022.
[36] N. Rahimi, S. Park, W.e.a. Choi, A comprehensive review on ensemble solar power forecasting algorithms, J. Electr. Eng. Technol. 18 (2023) 719–733.
[37] A. Nespoli, E. Ogliari, S. Leva, A. Pavan, A. Mellit, V. Lughi, A. Dolara, Day-ahead photovoltaic forecasting: a comparison of the most E ective techniques, Energies 1621 (12) (29 April 2019).
[38] M.P. Almeida, M. Muñoz, I. Parra, O. Perpiñán, Comparative study of PV power forecast using parametric and nonparametric PV models, Sol. Energy 155 (15 July 2017) 854–866.
[39] M. Srbinovska, S. Pechkova, A. Pechkov, M. Celeska, A. Krkoleva, P. Dimovski, V. Andova, A comprehensive study on air pollution prediction in North Macedonia: insights from LASSO modeling, in: 11th International Conference on Electrical, Electronics and Computer Engineering (IcETRAN), 2024. Nis, Serbia.
[40] M. Srbinovska, V. Andova, A. Krkoleva, M. Celeska, V. Andonovic, M. Kutirov, M. Majstorski, Assessing the impact of air pollution in North Macedonia: a meteorological and green infrastructure study, in: 18th SDEWES conference, 2023. Dubrovnik.



Maja Celeska Krstevska is an Associate Professor at Ss. Cyril and Methodius University in Skopje, where she earned her B. Sc. (2012), M.Sc. (2014), and Ph.D. (2019) degrees in Electrical Engineering. Her doctoral research was conducted in collaboration with Graz University of Technology. Dr. Celeska Krstevska has authored over 30 scientific publications and technical papers in areas including renewable energy systems, wind energy conversion, sustainable development, and electromechanical energy conversion. She has actively participated in international and national projects, including DAAD, ERASMUS+, IPA ADRION, and research initiatives supported by Ss. Cyril and Methodius University. She has been an active member of the IEEE professional organization for over a decade, contributing to advancements in her field.



Petar Krstevski received his MSc and PhD degree in electrical engineering from Ss Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies (FEEIT) in the area of power systems. His research interests include electricity markets and regulation, cross-border coordination in the operation of power systems and integration of RES generation in power systems. He has been actively involved in several EU funded projects from the FP7 and Horizon Europe programmes as well as in other international and national projects and studies related to his field of interest. He is author or co-author of more than fifty research papers. Currently he holds position of Associate Professor at Ss Cyril and Methodius University, FEEIT, working at the Power Systems Department. Associate professor

Dr. **Mare Srbinovska** received her BSc in 2003, MSc in 2009 and Ph.D. in 2015 from the Faculty of Electrical Engineering and Information Technology (FEEIT), Ss. Cyril and Methodius University in Skopje (UKIM). She is author/co-author of 10 papers in journals from SCI/SCIe list, >60 conference and symposium papers. She is coordinator of one project financed by Erasmus + program for higher education, two projects from FEEIT for air quality improvement using green walls, she has participated in TEMPUS projects and one DAAD project. Her expertise is related to metrology, sensors and sensor technologies, data acquisition and virtual instrumentation.

Vesna Andova is a full professor at the Faculty of Electrical Engineering and Information Technology at the University of Ss. Cyril and Methodius in Skopje. In 2008, she graduated from the Faculty of Natural Sciences and Mathematics at the University of Ss. Cyril and Methodius in Skopje, where she also obtained her Master's degree. In 2013, she completed her PhD in Mathematics at the University of Ljubljana, Slovenia, with a thesis entitled "Some distance and degree invariants and fullerene structures". She has authored and co-authored several research papers in graph theory or chemical graph theory and applied mathematics.

Aleksandra Krkoleva Mateska is professor at the Ss Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies (UKIM/FEIT). She works in the field of power systems, focusing on Smart Grids, renewable sources integration in distribution grids and Microgrids, electricity markets and regulation related to these areas. She has had a number of study visits to other universities, including at the University of Manchester, UK, University of Rostock, Germany, National Technical University in Athens, Greece. She is an author and co-author of more than ninety research papers published in conferences and international journals. She has participated in several research projects financed by various programs of the European Commission as a member of the UKIM/FEIT. As a representative of the CROSSBOW and R2D2 projects, she participates in the BRIDGE Regulation Working Group. She is a member of IEEE and CIGRE.