



ARTICLE

Predicting PV Power with a Multi-Stage Attention Neural Network Based on Neural Ordinary Differential Equations at Egyptian Stations

Mohamed R. Aboelmagd^{1,*}, Ali Selim^{1,2,*} and Mamdouh Abdel-Akher¹

¹Department of Electrical Engineering, Faculty of Engineering, Aswan University, Aswan, Egypt

²Department of Electrical Engineering, University of Jaén, EPS Linares, Jaén, Spain

*Corresponding Authors: Mohamed R. Aboelmagd. Email: mohamed.abolmagd@eng.aswu.edu.eg; Ali Selim.

Email: ali.selim@aswu.edu.eg

Received: 16 January 2026; Accepted: 16 March 2026; Published: 12 July 2026

ABSTRACT: To enable the integration of photovoltaic (PV) power into electrical grids, accurate predictions are vital. This study applies the Multistage Attention Neural Ordinary Differential Equation (MANODE) model, which combines Long Short-Term Memory (LSTM) networks, Temporal Convolutional Networks (TCN), and a two-stage attention mechanism to capture complex spatiotemporal patterns for PV power forecasting. The improved MANODE model is evaluated on three real-world datasets from PV stations in Egypt. Each dataset contains 12 feature parameters and spans an entire year. Comprehensive comparisons are conducted between the improved MANODE model and other neural network models, including one-layer and two-layer FNNs and time-series models, as well as a curve-fitting method. The proposed model achieves high correlation and negligible mean squared error (MSE) in one-step-ahead hourly forecasting for three locations in Egypt. Overall, MANODE demonstrates strong performance and represents a robust deep learning framework for estimating PV system power requirements, with effective real-time modeling and feature representation under changing weather conditions.

KEYWORDS: MANODE; PV power forecasting; convolutional networks; long short-term memory

1 Introduction

The world is moving towards renewable energy systems, increasing the need for more accurate forecasting models to facilitate the integration and control of photovoltaic (PV) power generation. Solar energy represents an important aspect of long-term energy solutions. Solar power generation remains somewhat unpredictable due to its strong dependence on factors such as weather conditions, wind speed, and sunlight. Weather-related factors change in complex and uncertain ways, making typical forecasting methods less effective. Such limitations often result in poor resource allocation and inefficient grid management [1].

According to the International Renewable Energy Agency (IRENA), global solar power capacity expanded by 3372 GW in 2022 [2]. Such changes reduce the stability and reliability of electrical power networks and may lead to financial losses and inefficient grid operations. Consequently, transmission system operators are required to develop strict planning strategies and regulatory measures to ensure effective coordination of solar energy sources.

Accurate forecasting of PV power generation represents a key component of time-series analysis. Forecasts are commonly classified into four categories based on the prediction horizon: ultra-short-term, short-term, medium-term, and long-term [2,3].

Forecasts made over a short time horizon, such as an hour or a day, are referred to as ultra-short-term and short-term forecasts [2,4]. Day-ahead generation strategies largely focus on short-term forecasting, while real-time power grid management relies more on ultra-short-term forecasting [4,5]. In contrast, medium- and long-term projections serve maintenance and future planning needs [3,6]. In this context, this work emphasizes the significance of accurate short-term and multivariate one-step PV power projections for efficient day-ahead power grid management and generation planning. There are various ways to forecast the amount of PV electricity produced each hour using multivariate solar data from the past 24 h, including deep learning, classical machine learning, statistical methods, and physical methods [3–7]. Among physical approaches is the use of “white box methods,” which examine the impact of factors such as solar radiation, temperature, latitude, and elevation on PV systems [3,8].

In [3], an efficient model was used to predict PV power one hour ahead. In this regard, a comprehensive architecture for a physical model chain comprising seven primary computing activities, was devised, and various physical models were evaluated for each stage [4]. The findings from these efforts indicated that transposition modelling and irradiance separation are both significant processes [4,5]. In addition to physical models, statistical models are often used to predict power because they are effective at analyzing time-series data [9]. These models use mathematical approaches to identify patterns and relationships, enabling accurate PV power predictions using methods such as autoregressive (AR) models, autoregressive integrated moving average (ARIMA) models, and seasonal autoregressive integrated moving average (SARIMA) models [6–8]. Notably, one study demonstrated that the technique used a reduced dataset to estimate parameters and orders, thereby enhancing power consumption forecasts [10].

Partial Functional Linear Regression Model (PFLRM) has been introduced to estimate the daily energy generation of solar panels [6]. The results showed that the regularized PFLRM performs better in predicting power output than standard multiple linear regression and neural network models. These approaches offer several advantages, such as reduced time consumption and simplicity of use [11]. Furthermore, machine learning algorithms can accurately identify existing relationships in time-series data when using nonlinear modelling frameworks [12]. For example, some machine learning methods utilize support vector machines (SVMs), random forests (RFs), and extreme learning machines (ELMs) to identify hidden relationships in datasets and to establish nonlinear prediction frameworks [3,12–14]. More recently, hybrid methods that combine SVM principles with historical power generation data and weather forecasts have performed well in weather classification [15,16]. These previous studies have examined how well benchmark machine learning algorithms, including SVM, RF, neural networks, and linear regression, can predict PV production [1,7,10,12]. Finally, a random tree classification method has been presented that groups the solar power generated at each interval into four categories based on the weather conditions at the time [6,17].

However, in PV power forecasting based on neural networks, some models face critical challenges. For instance, physical methods require precise data and complex calculations, which are often unavailable due to the intermittent nature of PV power [6–8]. Similarly, statistical models such as ARIMA and SARIMA can produce significant forecast errors and inaccurate predictions in PV forecasting.

Statistical models may also fail to account for nonlinear interactions among input variables. Classical machine learning methods, such as SVMs, can handle these nonlinear patterns. However, fixed nonlinear models can slow down the process and make it harder to identify the main relationships [18,19]. Deep learning algorithms have improved solar power output predictions, outperforming physical, statistical, and traditional machine learning methods [20,21]. They can model complex PV time series with nonlinear features, thereby improving forecast accuracy. The main deep learning algorithms for time-series forecasting are convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [22,23]. In this study, advanced RNNs, including long short-term memory (LSTM) and bidirectional long short-term memory (BiLSTM), are used

to capture temporal patterns in PV series. The LSTM method uses recent data and hourly weather forecasts to predict sunlight for the next day [24,25].

In [16], the Multistage Attention Neural Ordinary Differential Equation (MANODE) model was developed to predict solar power generation using nine machine-learning models, grouped by environmental conditions. The MANODE framework comprises a spatiotemporal attention extraction network and a multistage attention mechanism that works together to produce a continuous neural network capable of collecting detailed, real-time spatiotemporal data. It also uses the Attention-based Time-series Ordinary Differential Equation (AttnTODE), a deep recurrent temporal weight extractor that is expected to capture challenging hidden temporal correlations in PV data [19,26].

The paper evaluates large PV power plant datasets and remasters them using machine learning. Fig. 1 presents a clear step-by-step diagram of the data training and remastering process. The paper then focuses on analyzing the MANODE optimization model. The machine learning-based MANODE optimization model is distinguished from others by its comprehensive, continuous temporal characteristics and its ability to capture temporal dependencies. Fig. 2 presents the framework of the AttnTODE block, showing its key modules and the information flow through the architecture.

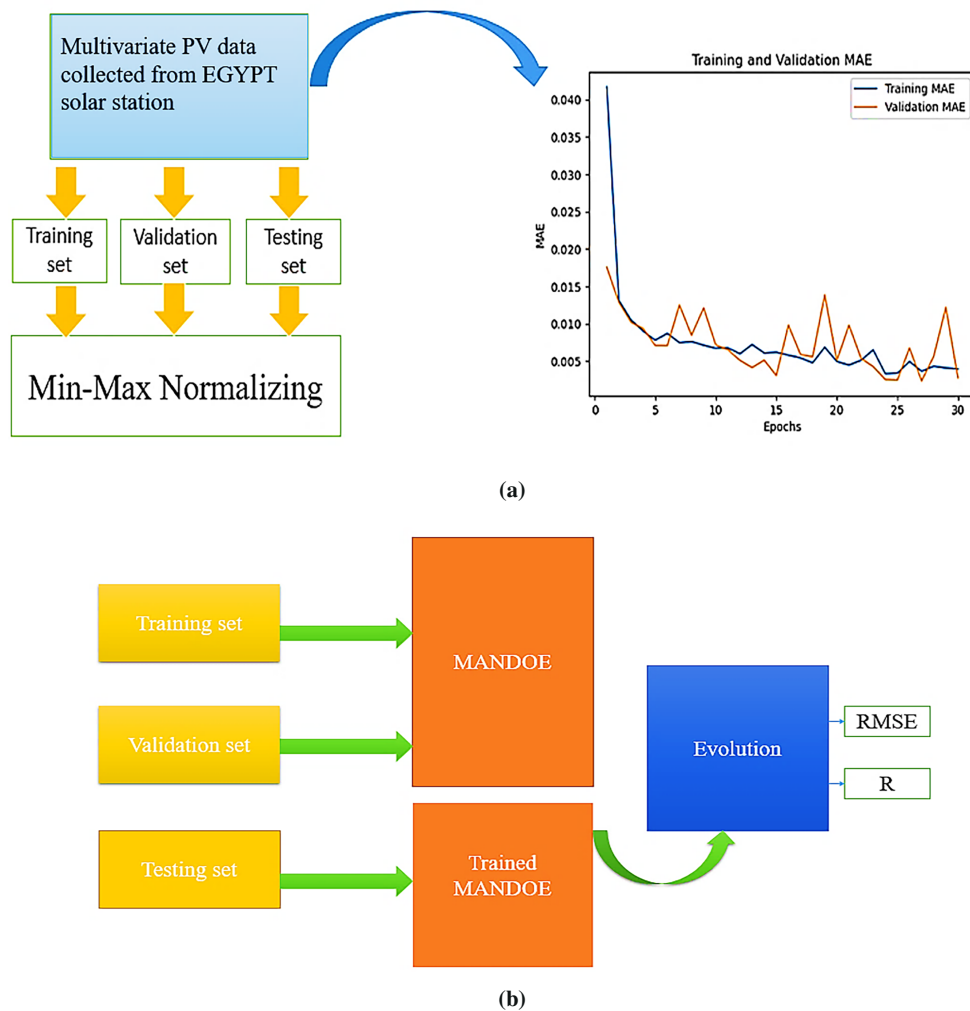


Figure 1: PV power forecasting flowchart, (a) data splitting, (b) evolution data [16].

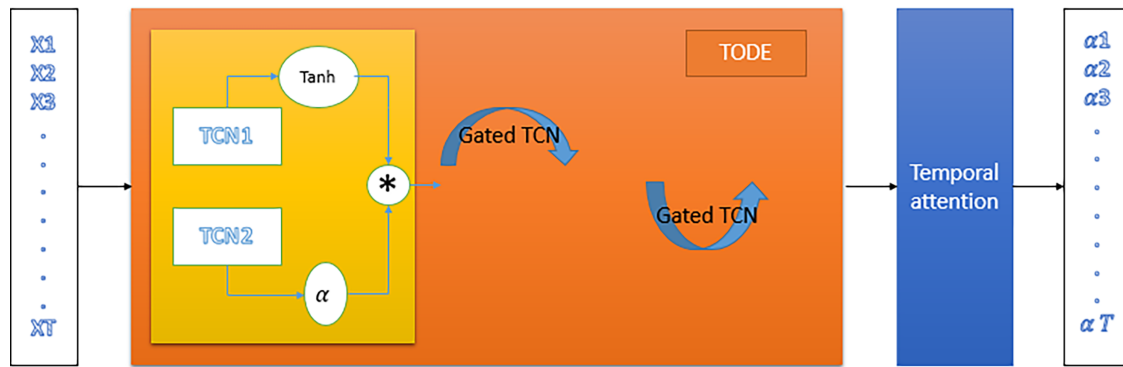


Figure 2: The AttnTODE block's framework. The AttnTODE block [16].

Previously, in [16], MANODE achieved results that surpassed other models. However, in this study, an improved version of MANODE is applied using 12 feature parameters to enhance the performance of the original MANODE. Therefore, three real-world datasets from PV power plants in Benban, Egypt, are used to train the model based on the weather and environmental conditions of each location.

The paper is organized as follows: Section 2 explains how the model is developed, including the MANODE architecture and its key components. Section 3 presents the methodology and the datasets. Section 4 provides a statistical analysis of the PV solar measurements. Finally, Section 5 examines the results, discussion, and analysis.

2 Architecture of the Model

This research employs the MANODE methodology [16]. The model is initially evaluated with nine feature parameters and subsequently with 12 to assess the impact of the number of input features on performance. The model is utilized 24 h a day for an entire year (365 days). Table 1 shows the differences between the present study, which utilizes an improved version of MANODE, and the original MANODE presented by Huang et al. (2024) in [16]. Fig. 3 presents the implementation architecture of the proposed MANODE model, illustrating its main modules and the overall flow of the forecasting process.

Table 1: Structural and methodological differences between Huang et al. (2024) and the proposed version of MANODE framework.

Feature	Huang et al. (2024) [16]	Present Work (Improved Version of MANODE)
Input features	9 meteorological and operational features.	12 features including inverter DC/AC power and wind speed.
Architecture depth	3 computational layers.	4 fully connected spatiotemporal layers.
Additional module	No autoregressive component.	Introduced an autoregressive extrapolation layer integrated with NODE.
Dataset	Single PV dataset (China).	Three Egyptian PV stations representing diverse climates.
Forecasting horizon	Single short-term horizon.	One-step-ahead (t+1 h) forecasting.
Evaluation metrics	RMSE and MAPE.	correlation analysis, and feature importance.
Generalization	Limited to controlled data.	Validated on real-world multi-site datasets under variable weather.

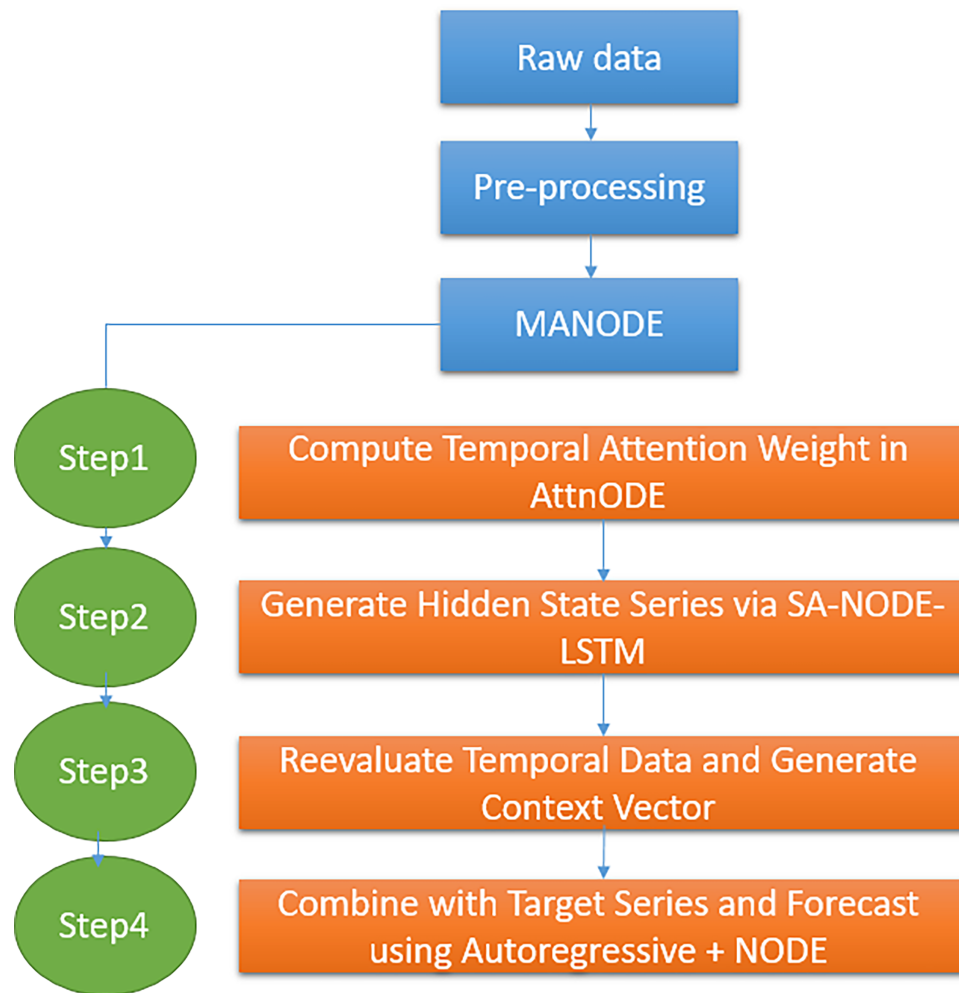


Figure 3: Implementation architecture of the proposed MANODE model.

3 Methodology

3.1 Dataset and Methodology

The Egyptian Meteorological Authority provided three different PV datasets for this study: SOHAG Centre (Site 1), AKHMIM City (Site 2), and ELGAZARA City (Site 3). There are 12 distinct features for each of these three multivariate datasets. Each site includes 8760 hourly records (365 days $\times$ 24 h). Each record contains 12 input features., recorded every hour. [Table 2](#) illustrates the numerous architectural settings (such depth settings) that were looked at to see if deeper models are better at predicting short-term PV dynamics. [Table 3](#) displays the different sets of features (9 vs. 12 inputs) that were used to find out if adding more weather or operational variables makes predictions more accurate. To see how each change affects things, everything else stays the same.

Table 2: The distinction between previous research and the present study.

	Previous Research [16]	The Present Research
Number of layers	3	4
Number of days	180	365
Hours during the day	12	24
Features and Description	9	12

Table 3: Multivariate PV data collected from EGYPT solar station.

Features and Description	9 Feature Parameters	12 Feature Parameters
Time stamp	☒	☒
Weather file ambient temperature (°C)	☒	☒
Weather file global horizontal irradiance (W/m ²)	☒	☒
Weather file wind speed (m/s)	☒	☒
Weather file diffuse irradiance (W/m ²)	☒	☒
Weather file beam irradiance (W/m ²)	☒	☒
Transmitted plane of array irradiance (W/m ²)	☒	☒
Plane of array irradiance (W/m ²)	☒	☒
AC inverter output power (W)	☒	☒
Angle of incidence (deg)	☒	☒
DC inverter input power (W)	☒	☒
System power generated (kW)	☒	☒

Fig. 4 shows where the three PV locations are in relation to each other. The location of the site is significant for PV forecasts since the climate drivers in that area (such temperature, humidity, and wind patterns) are different from those in other locations and have a direct effect on how much power PV systems can make. Testing the suggested approach on stations in various geographically distinct locations within Egypt provides a clearer understanding of their efficacy across diverse scenarios and serves as an effective initial step in assessing its performance in different regions.

Each dataset contains 12 meteorological and operational features, including array irradiance, outside temperature, wind speed, inverter DC power, inverter AC power, humidity, and global horizontal irradiance. The data were recorded hourly over the study period. The datasets are split chronologically into three groups: 70% for training, 15% for validation, and 15% for testing (no random shuffling). For one-step-ahead forecasting, we use a fixed lookback window of 24 h; for each time index t , the input sequence is $X_t = [x_{t-23}, \dots, x_t]$ and the target is the PV power at the next hour $y_t = p_{t+1}$. The sliding window advances with a stride of one hour (one sample per hour). All input features are normalized using min-max scaling. The scaling parameters are computed on the training set and applied to validation and test sets. Normalization (stopping leakage). Min-max is used to scale all the input features. You use the training set to find the scaling parameters (min and max), and then you apply those same values on the validation and test sets to keep data from leaking.

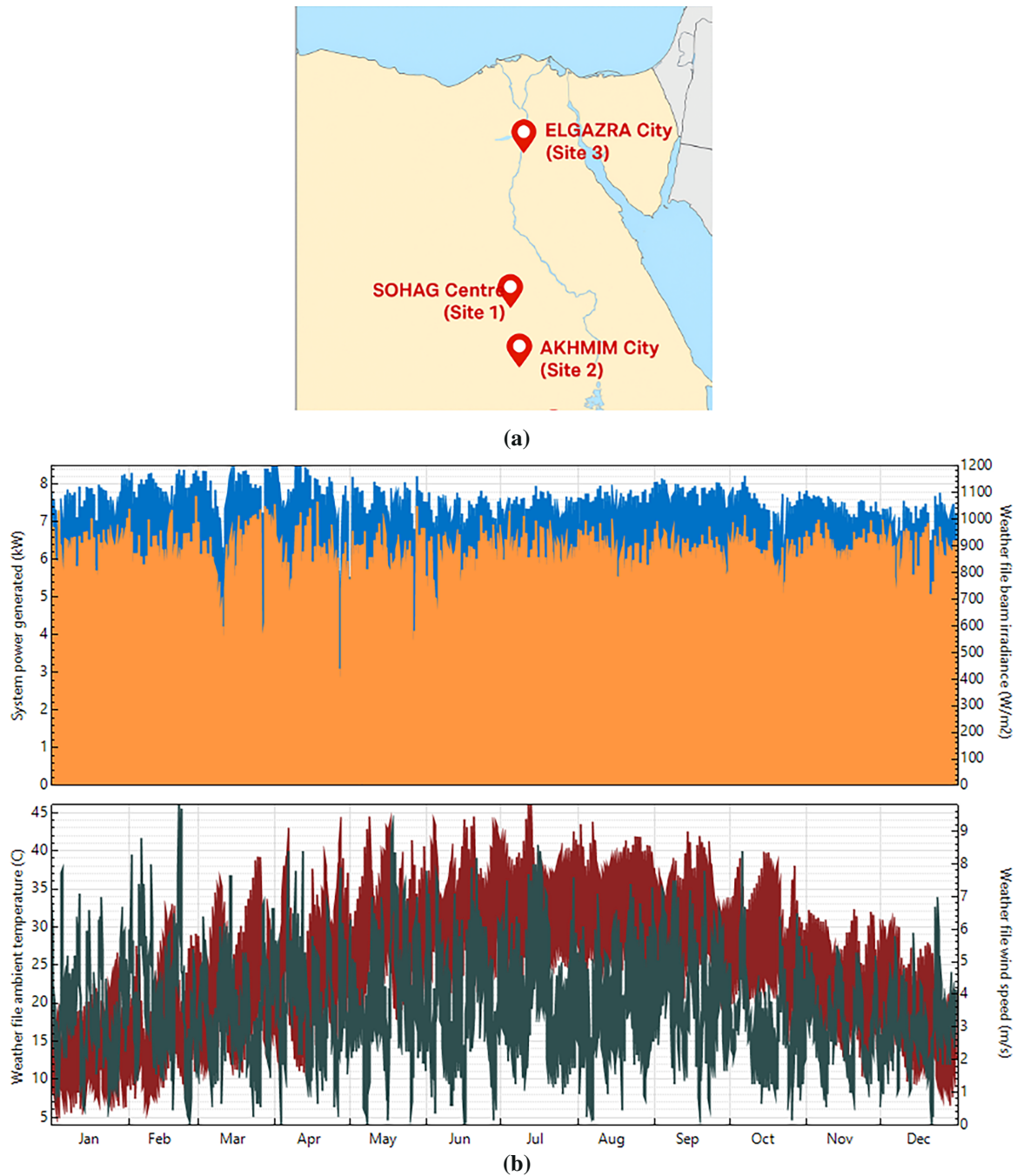


Figure 4: Multivariate PV from EGYPT solar stations, (a) MAP, (b) data collected.

3.2 Statistical Analysis of Solar PV Performance Metrics

To provide an overview of the datasets, a statistical analysis is performed, and the annual results are summarized in [Tables 4–6](#). These tables report the descriptive statistics at Sites 1, 2, and 3, including the mean, minimum, maximum, total sum, standard deviation, average daily minimum, and average daily maximum. Notably, the system operates at a typical DC input power of 2305.19 W, with the inverter producing an AC output of 2213.12 W and a maximum possible output of 8333.33 W. Module temperature can reach 75.31°C,

which negatively affects efficiency, while irradiance levels range from 295.7 to 1124.45 W/m², indicating strong daytime solar resource availability.

Table 4: The following table summarizes the performance and error evaluation for site 1.

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Database inverter power (W)	2213.12	0	8333.33	1.94E + 07	2846.03	0	7375.84
Hourly Data: Angle of incidence (deg)	25.0184	0	99.1148	219,161	29.8645	0	85.9419
Hourly Data: DC inverter input power (W)	2305.19	0	8888.07	2.02E + 07	2958.09	0	7676.12
Hourly Data: DC power loss due to snow (%)	0.507	0	1	4441	0.5	0	1
Hourly Data: Module temperature (°C)	30.6364	5.91	75.3078	268,375	14.261	18.4283	52.1135
Hourly Data: Plane of array irradiance (W/m ²)	295.7	0	1124.45	2.59E + 06	380.619	0	1003.49
Hourly Data: Sun up over horizon (0/1)	0.507	0	1	4441	0.5	0	1
Hourly Data: System power before grid curtailment (kW)	2.2131	0	8.3333	19,387	2.846	0	7.3758
Hourly Data: System power before grid interconnect (kW)	0	0	0	0	0	0	0
Hourly Data: System power generated (kW)	2.2131	0	8.3333	19,387	2.846	0	7.3758
Hourly Data: Transmitted plane of array irradiance (W/m ²)	292.285	0	1124.45	2.56E + 06	380.607	0	1003.38

(Continued)

Table 4 (continued)

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Data: Weather file ambient temperature (°C)	24.3446	5.91	44.56	213,259	8.2071	18.4234	31.4559
Hourly Data: Weather file beam irradiance (W/m ²)	315.206	0	1038.92	2.76E + 06	365.051	0	873.116
Hourly Data: Weather file diffuse irradiance (W/m ²)	69.0847	0	512	605,182	84.4148	0	211.116
Hourly Data: Weather file global horizontal irradiance (W/m ²)	273.311	0	1101	2.39E + 06	353.082	0	908.778
Hourly Data: Weather file snow depth (cm)	−999	−999	−999	−8.75E + 06	0	−999	−999
Hourly Data: Weather file wind speed (m/s)	3.5446	0.1	9.7	31,050.3	1.4476	2.2438	4.9332

Table 5: The following table summarizes the performance and error evaluation for site 2.

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Database inverter power (W)	2229.59	0	8333.33	1.95E + 07	2860.75	0	7415.17
Hourly Data: Angle of incidence (deg)	24.9821	0	99.1	218,843	29.8482	0	85.8728
Hourly Data: DC inverter input power (W)	2322.3	0	9152.78	2.03E + 07	2973.57	0	7717.71
Hourly Data: DC power loss due to snow (%)	0.5065	0	1	4437	0.5	0	1

(Continued)

Table 5 (continued)

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Data: Module temperature (°C)	29.996	5.15	72.4812	262,765	14.4001	17.44	51.5154
Hourly Data: Plane of array irradiance (W/m ²)	297.2	0	1154.87	2.60E + 06	382.047	0	1006.78
Hourly Data: Sun up over horizon (0/1)	0.5065	0	1	4437	0.5	0	1
Hourly Data: System power before grid curtailment (kW)	2.2296	0	8.3333	19,531.2	2.8607	0	7.4152
Hourly Data: System power before grid interconnect (kW)	0	0	0	0	0	0	0
Hourly Data: System power generated (kW)	2.2296	0	8.3333	19,531.2	2.8607	0	7.4152
Hourly Data: Transmitted plane of array irradiance (W/m ²)	293.753	0	1154.87	2.57E + 06	382.034	0	1006.71
Hourly Data: Weather file ambient temperature (°C)	23.7164	5.15	43.86	207,756	8.4011	17.4503	30.9092
Hourly Data: Weather file beam irradiance (W/m ²)	319.217	0	1255.69	2.80E + 06	368.786	0	875.621
Hourly Data: Weather file diffuse irradiance (W/m ²)	68.1646	0	491	597,122	82.9662	0	205.684

(Continued)

Table 5 (continued)

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Data: Weather file global horizontal irradiance (W/m ²)	274.159	0	1090	2.40E + 06	354.299	0	911.683
Hourly Data: Weather file snow depth (cm)	−999	−999	−999	−8.75E + 06	0	−999	−999
Hourly Data: Weather file wind speed (m/s)	3.5714	0.1	9.7	31,285.7	1.4371	2.2411	5.0468

Table 6: The following table summarizes the performance and error evaluation for site 3.

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Database inverter power (W)	2234.76	0	8333.33	1.96E + 07	2864.15	0	7421.58
Hourly Data: Angle of incidence (deg)	25.0168	0	99.2192	219,147	29.8637	0	85.9462
Hourly Data: DC inverter input power (W)	2327.64	0	9157.58	2.04E + 07	2977.15	0	7724.36
Hourly Data: DC power loss due to snow (%)	0.507	0	1	4441	0.5	0	1
Hourly Data: Module temperature (°C)	30.1324	5.61	72.2587	263,960	14.1414	17.9284	51.3425
Hourly Data: Plane of array irradiance (W/m ²)	297.676	0	1154.79	2.61E + 06	382.109	0	1006.6
Hourly Data: Sun up over horizon (0/1)	0.507	0	1	4441	0.5	0	1

(Continued)

Table 6 (continued)

Variable	Mean	Min	Max	Sum	St Dev	Avg Daily Min	Avg Daily Max
Hourly Data: System power before grid curtailment (kW)	2.2348	0	8.3333	19,576.5	2.8642	0	7.4216
Hourly Data: System power before grid interconnect (kW)	0	0	0	0	0	0	0
Hourly Data: System power generated (kW)	2.2348	0	8.3333	19,576.5	2.8642	0	7.4216
Hourly Data: Transmitted plane of array irradiance (W/m ²)	294.207	0	1154.79	2.58E + 06	382.107	0	1006.55
Hourly Data: Weather file ambient temperature (°C)	23.8591	5.61	43.76	209,005	8.1877	17.9584	30.7524
Hourly Data: Weather file beam irradiance (W/m ²)	319.782	0	1241.73	2.80E + 06	368.203	0	873.66
Hourly Data: Weather file diffuse irradiance (W/m ²)	68.4354	0	496	599,494	83.1919	0	205.785
Hourly Data: Weather file global horizontal irradiance (W/m ²)	274.545	0	1090	2.41E + 06	354.38	0	911.581
Hourly Data: Weather file snow depth (cm)	−999	−999	−999	−8.75E + 06	0	−999	−999
Hourly Data: Weather file wind speed (m/s)	3.5799	0.1	9.7	31,359.5	1.4278	2.2479	5.0318

For a fair comparison, all models are trained and tested using the same datasets and computing environment. The best settings for MANODE include a learning rate of 0.001, a batch size of 70, four computational

layers, and ReLU activation functions. The model is trained using the Adam optimizer for 150 epochs, and an early-stopping criterion is applied to prevent overfitting. The architectures and hyperparameters of the comparative approaches (CNN-LSTM [7], COA-CNN-LSTM [27], GCN-based [21,28], feedforward neural network (FNN), and time-series models) are adjusted as recommended in [16].

The selected hyperparameters are based on preliminary validation and are tuned within commonly used ranges. The primary objectives are stable convergence and optimal validation performance. These settings are then fixed across all experiments to ensure fair comparisons [29].

The descriptive statistics provide the ranges and distributions of the input variables at each site; however, they are not used to assess predictive accuracy. Dedicated prediction-error metrics presented in the Results section are used to evaluate forecasting performance [30,31].

4 Results and Discussions

In this section, the improved version of MANODE is applied to forecast PV power at three different locations using 12 feature parameters, and it is compared with the original MANODE and other competitive models. Fig. 5 shows that overall behavior is consistent across all three sites, and the training MSE continues to decrease as the number of epochs increases. Due to weather fluctuations and data noise, the convergence rate differs slightly between sites. Sites where irradiance patterns vary more frequently may take longer to converge or may reach a higher plateau. Overall, the curves indicate that the training process is stable across all sites.

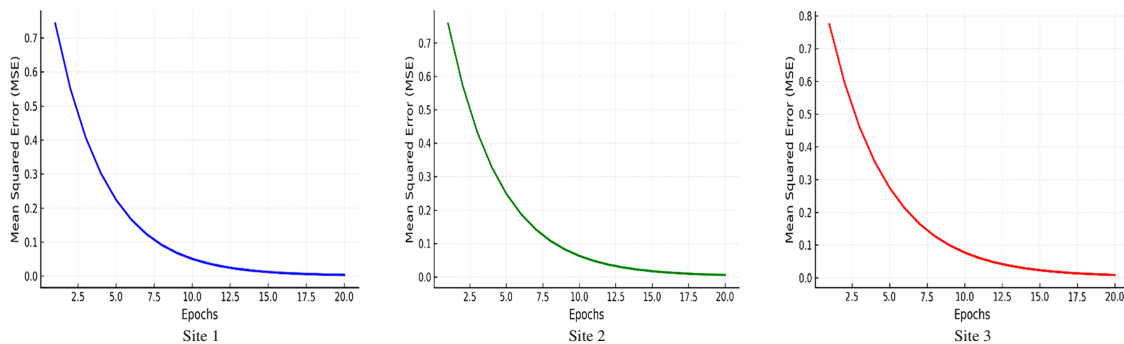


Figure 5: MSE for MANODE training.

The comparative outcomes are presented in Tables 7–9 for site 1, site 2, and site 3 respectively for different number of input parameters (12 and 9 feature parameters). The results show that using 12 input parameters lead to improving the performance of all models, as indicated values by the value of R, which measure the extent to which predicted values match actual values. and significantly lower MSE.

The better MANODE works better since its design fits the time-series aspects of PV better. The multi-stage attention adaptively highlights relevant time steps and variables in the 24-h lookback window. This is helpful when things are changing quickly, like when the sun rises and sets or when irradiance changes quickly. The added depth makes the model better at showing how weather and operational factors affect PV power in a nonlinear way. Finally, using 12 input features instead of 9 yields more clear signals, as indicated by the higher R and lower MSE in Tables 7–9.

Table 7: The Coefficient of correlation is the “R” value for different parameters for site 1.

MODEL	R		MSE	
	(For 12 Feature Parameters)	For (9 Feature Parameters)	(For 12 Feature Parameters)	For (9 Feature Parameters)
FNN (one-layer)	0.99999	0.9698	0.0256	0.002096
FNN (two-layer)	0.99999	0.9698	0.00134	0.0021407
Time series	0.99999	0.9631	0.8839	0.076319
Fitting	0.99999	0.9698	0.01718	0.01829
MANODE	0.99999	0.9795	0.000124	0.00056

Table 8: The Coefficient of correlation is the “R” value for different parameters for site 2.

MODEL	R		MSE	
	(For 12 Feature Parameters)	For (9 Feature Parameters)	(For 12 Feature Parameters)	For (9 Feature Parameters)
FNN (one-layer)	0.9983	0.9675	0.0267	0.00310
FNN (two-layer)	0.9985	0.9692	0.0267	0.00298
Time series	0.9954	0.9620	0.0757	0.00720
Fitting	0.99999	0.9701	0.00243	0.00550
MANODE	0.99999	0.9790	0.000195	0.00065

Table 9: The Coefficient of correlation is the “R” value for different parameters for site 3.

MODEL	R		MSE	
	(For 12 Feature Parameters)	For (9 Feature Parameters)	(For 12 Feature Parameters)	For (9 Feature Parameters)
FNN (one-layer)	0.9985	0.9680	0.0257	0.00345
FNN (two-layer)	0.9987	0.9695	0.0247	0.00320
Time series	0.9947	0.9631	0.0858	0.00810
Fitting	0.99999	0.9698	0.0500	0.00600
MANODE	0.99999	0.9785	0.000210	0.00068

This is in line with the one-step-ahead hourly setting, which keeps PV power the same throughout time. It's important to keep in mind that the forecasting structure is strictly causal (inputs up to t , target at $t + 1$). This helps with worries about leaks.

The improved MANODE consistently outperforms the other models across all evaluation metrics, thereby validating its enhanced capacity to accurately model the spatial and temporal dependencies inherent in real-world PV data.

Fig. 6 illustrates the impact of the main input features on forecasting process which can be measured using R. The figure shows that the three new features that are used, such as the inverters of DC and AC powers, have a significant impact however, the wind speed gives a slightly impact prediction accuracy.

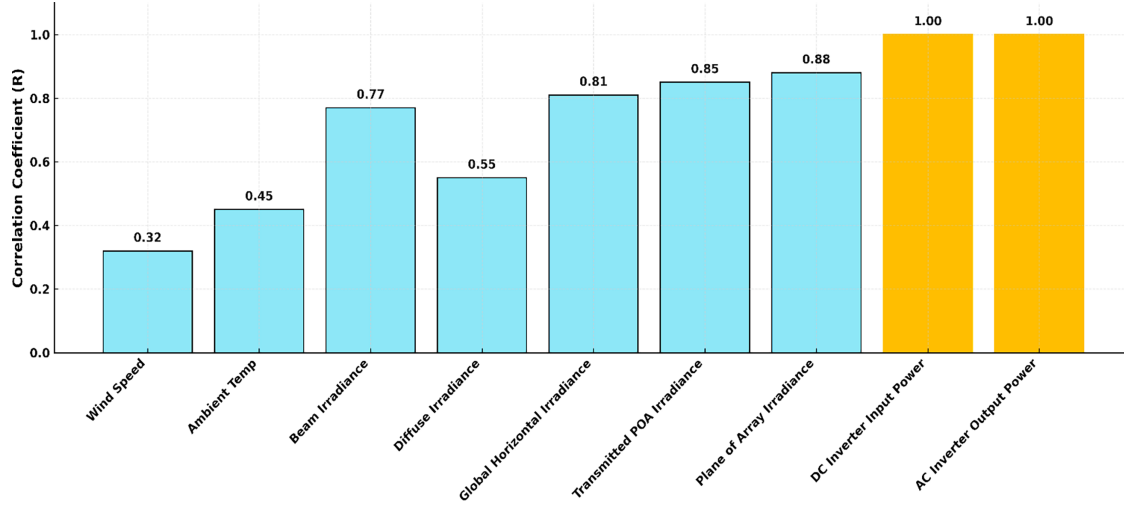


Figure 6: Correlation between inputs and system power generated.

Recognize that the given correlation values are elevated in comparison to numerous PV forecasting studies, especially over extended timeframes. In our study, the objective is one-step-ahead hourly forecasting, wherein photovoltaic (PV) electricity has significant temporal continuity, making it highly predictable over short timeframes. The suggested model uses multi-stage attention to take advantage of short-term trends in time. The evaluation is based on a strict causal alignment (inputs up to time t , target at $t + 1$), which helps avoid problems with data leakage. We caution that performance is likely to drop in more difficult situations, like multi-step/day-ahead forecasting and generalizing across multiple sites.

Prediction-error metrics are used to figure out how accurate a forecast is instead of descriptive data. We use the correlation coefficient R to see how well the predicted and measured PV power matches up. We utilize the mean squared error (MSE) to figure out how big the prediction errors are. These metrics directly evaluate predicting efficacy and are therefore employed as the primary standard for model comparison. We also talk about how mistakes behave between sites and point out that high R values don't always mean tiny absolute errors. That's why we report MSE along with R .

5 Conclusion

This research introduces an improved MANODE-based framework for one-step-ahead hourly solar power forecasting at three stations in Egypt. The proposed method enhances the original MANODE by increasing representational capacity and incorporating attention-based temporal weighting, which improves its ability to capture rapid changes in PV generation. The experimental evaluation demonstrates that the improved MANODE achieves higher predictive performance across all three sites compared with conventional baseline models (one-layer FNN, two-layer FNN, and a time-series baseline), thereby validating its effectiveness for short-term PV forecasting. A key advantage of the proposed approach is its ability to leverage temporal dependencies in multivariate observations, which is particularly important under rapidly changing weather conditions. However, the present investigation has several limitations. Future work will strengthen the evaluation by examining error distributions (e.g., residual histograms and quantiles),

assessing performance under extreme operating conditions, and analyzing deviations associated with peaks and ramps, while also reporting additional metrics such as MAE and MAPE. In addition, controlled ablation studies will be conducted, including: (a) MANODE without attention, (b) an attention-based model with a discrete-time backbone (without NODE), and (c) capacity-matched variants with comparable depth and parameter counts. Computational cost and model complexity (e.g., number of parameters and inference time) will also be reported for each variant. Moreover, future studies will include stronger state-of-the-art baselines for comparison. Overall, the current study reports point out estimates (R and MSE) without systematic ablations, extensive benchmarking against state-of-practice methods, reliability-focused analyses (e.g., extremes, peaks, and ramps), or statistical significance testing. These aspects will be addressed in subsequent research through capacity-matched ablations, improved benchmarking, more practical evaluation protocols (including MAE/MAPE and residual/quantile analyses), and statistical validation (e.g., multiple runs, confidence intervals, and Diebold–Mariano testing).

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Mohamed R. Aboelmagd: conceptualization, data curation, formal analysis, methodology, project administration, resources, software, validation, writing—original draft. Ali Selim: conceptualization, supervision, writing—review and editing. Mamdouh Abdel-Akher: conceptualization, supervision, writing—review and editing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Aboelmagd MR, Selim A, Abdel-Akher M. Forecasting of power generation in a single-axis solar tracking PV system using an enhanced artificial neural network-based method. *IET Renew Power Gener.* 2025;19(1):e70100. doi:10.1049/rpg2.70100.
2. Oprea SV, Bâra A. Ultra-short-term forecasting for photovoltaic power plants and real-time key performance indicators analysis with big data solutions. Two case studies—PV Agigea and PV Giurgiu located in Romania. *Comput Ind.* 2020;120:103230. doi:10.1016/j.compind.2020.103230.
3. Das UK, Tey KS, Seyedmahmoudian M, Mekhilef S, Idris MYI, Van Deventer W, et al. Forecasting of photovoltaic power generation and model optimization: a review. *Renew Sustain Energy Rev.* 2018;81:912–28. doi:10.1016/j.rser.2017.08.017.
4. Sangrody H, Zhou N, Zhang Z. Similarity-based models for day-ahead solar PV generation forecasting. *IEEE Access.* 2020;8:104469–78. doi:10.1109/access.2020.2999903.
5. Aboelmagd MR, Selim A, Abdel-Akher M, Jurado F. Enhanced solar energy forecasting and economic dispatch using Jaya-optimized machine learning models. In: *Proceedings of the 2025 26th International Middle East Power Systems Conference (MEPCON)*; 2025 Dec 20–22; Aswan, Egypt. p. 1–6. doi:10.1109/mepcon66918.2026.11360174.
6. Sharadga H, Hajimirza S, Balog RS. Time series forecasting of solar power generation for large-scale photovoltaic plants. *Renew Energy.* 2020;150:797–807. doi:10.1016/j.renene.2019.12.131.
7. Qu J, Qian Z, Pei Y. Day-ahead hourly photovoltaic power forecasting using attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern. *Energy.* 2021;232:120996. doi:10.1016/j.energy.2021.120996.

8. Paulescu M, Stefu N, Dughir C, Sabadus A, Calinoiu D, Badescu V. A simple but accurate two-state model for nowcasting PV power. *Renew Energy*. 2022;195(4):322–30. doi:10.1016/j.renene.2022.05.056.
9. Kumar A, Dubey AK, Ramírez IS, Muñoz del Río A, Márquez FPG. A review and analysis of forecasting of photovoltaic power generation using machine learning. In: *Proceedings of the Sixteenth International Conference on Management Science and Engineering Management—Volume 1*. Berlin/Heidelberg, Germany: Springer; 2022. p. 492–505. doi:10.1007/978-3-031-10388-9_36.
10. Rodríguez F, Galarza A, Vasquez JC, Guerrero JM. Using deep learning and meteorological parameters to forecast the photovoltaic generators intra-hour output power interval for smart grid control. *Energy*. 2022;239(6):122116. doi:10.1016/j.energy.2021.122116.
11. Başaran K, Bozyiğit F, Siano P, Yıldırım Taşer P, Kılınç D. Systematic literature review of photovoltaic output power forecasting. *IET Renew Power Gener*. 2020;14(19):3961–73. doi:10.1049/iet-rpg.2020.0351.
12. Abdellatif A, Mubarak H, Ahmad S, Ahmed T, Shafiullah GM, Hammoudeh A, et al. Forecasting photovoltaic power generation with a stacking ensemble model. *Sustainability*. 2022;14(17):11083. doi:10.3390/su141711083.
13. Aboelmagd M, Diab A, Dousoky GM. Failure analysis in photovoltaic power systems using an artificial neural network. *J Adv Eng Trends*. 2021;41(2):205–18. doi:10.21608/jaet.2021.49408.1069.
14. Lin W, Zhang B, Li H, Lu R. Multi-step prediction of photovoltaic power based on two-stage decomposition and BILSTM. *Neurocomputing*. 2022;504(2):56–67. doi:10.1016/j.neucom.2022.06.117.
15. Wang F, Zhang Z, Liu C, Yu Y, Pang S, Duić N, et al. GAN-CNN-based weather classification model for day-ahead short-term PV power forecasting. *Energy Convers Manag*. 2019;181:443–62.
16. Huang S, Zhou Q, Shen J, Zhou H, Yong B. Multistage spatio-temporal attention network based on NODE for short-term PV power forecasting. *Energy*. 2024;290(9):130308. doi:10.1016/j.energy.2024.130308.
17. Wang L, Mao M, Xie J, Liao Z, Zhang H, Li H. Accurate solar PV power prediction interval method based on frequency-domain decomposition and LSTM model. *Energy*. 2023;262(3):125592. doi:10.1016/j.energy.2022.125592.
18. Jung Y, Jung J, Kim B, Han S. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: case study of South Korea. *J Clean Prod*. 2020;250(6):119476. doi:10.1016/j.jclepro.2019.119476.
19. Aboelmagd MR, Selim A, Abdel-Akher M. Predicting of photovoltaic power generation using similarity-based forecasting models. In: *Proceedings of the 2023 24th International Middle East Power System Conference (MEPCON)*; 2023 Dec 19–21; Mansoura, Egypt. p. 1–8. doi:10.1109/mepcon58725.2023.10462348.
20. Hosny M, Elgendy IA, Chelloug SA, Ahmad Albashrawi M. Attention-based convolutional neural network model for skin cancer classification. *IEEE Access*. 2025;13:172027–50. doi:10.1109/access.2025.3616022.
21. Xie L, Yang J, Liu B. Ultra-short-term forecasting of multi-site PV power based on graph convolution network. *J Electr Eng Technol*. 2025;20(5):3053–69. doi:10.1007/s42835-025-02255-w.
22. Aboelmagd MR, Selim A, Abdel-Akher M. PV grid forecasting analysis using artificial neural network. In: *Proceedings of the 2024 25th International Middle East Power System Conference (MEPCON)*; 2024 Dec 17–19; Cairo, Egypt. p. 1–8. doi:10.1109/mepcon63025.2024.10850132.
23. Cavus M, Allahham A. Spatio-temporal attention-based deep learning for smart grid demand prediction. *Electronics*. 2025;14(13):2514. doi:10.3390/electronics14132514.
24. Varga G, Gresina F, Szeberényi J, Gelencsér A, Rostási Á. Effect of Saharan dust storm events on the forecast of PV power generation in Hungary. In: *Proceedings of the European Geosciences Union General Assembly 2024 (EGU24)*; 2024 Apr 14–19; Vienna, Austria.
25. Amirteimoury F, Memarzadeh G, Noori H, Keynia F, Heydari A, Fekih A. A novel multi-scale feature attention network-based prediction model for the accurate forecasting of photovoltaic power generation. *IEEE Access*. 2025;13:165173–88. doi:10.1109/access.2025.3611168.
26. Kuo Y, Cai Y, Cheng J. A deep learning model based on multi-attention mechanism and gated recurrent unit network for photovoltaic power forecasting. *Comput Electr Eng*. 2025;123:110250. doi:10.1016/j.compeleceng.2025.110250.

27. Abou Houran M, Salman Bukhari SM, Zafar MH, Mansoor M, Chen W. COA-CNN-LSTM: coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications. *Appl Energy*. 2023;349(3):121638. doi:10.1016/j.apenergy.2023.121638.
28. Yalavarthi VK, Scholz R, Born S, Schmidt-Thieme L. Probabilistic forecasting of irregularly sampled time series with missing values via conditional normalizing flows. *Proc AAAI Conf Artif Intell*. 2025;39(20):21877–85. doi:10.1609/aaai.v39i20.35494.
29. Qin L, Sun T, Zhang S, Sun XM. Tensorized timeline alignment for neural dynamical systems in irregularly sampled time series prediction. In: *Proceedings of the 2023 42nd Chinese Control Conference (CCC)*; 2023 Jul 24–26; Tianjin, China. p. 8599–604. doi:10.23919/CCC58697.2023.10240296.
30. Jung S, Park YJ, Jeong J, Kim KM, Kim H, Kim M, et al. Global-local item embedding for temporal set prediction. In: *Proceedings of the 15th ACM Conference on Recommender Systems*; 2021 Sep 27–Oct 1; Amsterdam, The Netherlands. p. 674–9. doi:10.1145/3460231.3478844.
31. Leeney W, Gravina A, Bacciu D. Non-dissipative graph propagation for non-local community detection. In: *Proceedings of the 2025 International Joint Conference on Neural Networks (IJCNN)*; 2025 Jun 30–Jul 5; Rome, Italy. p. 1–8. doi:10.1109/IJCNN64981.2025.11228363.