Shahid Bahonar University of Kerman

Biomechanism and Bioenergy Research

Online ISSN: 2821-1855
Homepage: https://bbr.uk.ac.ir

Iranian Society of Agricultural Machinery Engineering and Mechanization

Deep Learning Framework for Joint Prediction of Energy Resources in an Industrial Building

Mohammad Hajian[1], Soha Sami [2], Tayyeb Nazghelichi [1]✉

[1] Gorgan University of Agricultural Sciences and Natural Resources, Faculty of Water and Soil Engineering, Department of Biosystems Engineering, Gorgan, Iran.
[2] Mechanics of Biosystems Engineering Department, College of Aburaihan, University of Tehran, Tehran, Iran.

✉ Corresponding author: tnazghelichi@gmail.com

ARTICLE INFO

Article type:

Research Article

Article history:

Received 11 September 2025

Received in revised form 28 October 2025

Accepted 30 November 2025

Available Online 31 December 2025

Keywords:

Photovoltaic Generation, Neural Networks, LSTM, Short-Term Energy Forecasting, Time-Series Analysis.

ABSTRACT

Renewable energy sources like solar photovoltaic (PV) systems are inherently volatile, necessitating accurate short-term forecasting for efficient grid management, energy storage, and consumption optimization. This study proposes a deep learning framework using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models to simultaneously forecast PV generation (from roof and facade panels), on-site electrical consumption, and grid imports in an industrial building, relying solely on raw historical time-series data without meteorological or temporal features. The dataset, sourced from the Open Power System Data repository, comprises minute-by-minute measurements from a facility in Konstanz, Germany, preprocessed with Min-Max scaling and linear interpolation for gaps. The models were trained on 80% of the data using a 24-hour sequence length, with architectures featuring two recurrent layers (50 units each), dropout (0.2), ReLU activation, Adam optimizer, and MSE loss. Early stopping and learning rate reduction callbacks prevented overfitting. Evaluation via MSE, RMSE, MAE, and R^2 on test data showed strong performance, with GRU achieving an average R^2 of 0.906, MAE of 0.034, MSE of 0.0038, and RMSE of 0.062. Five-fold cross-validation confirmed model stability (mean $R^2 \approx 0.89$ for both architectures), with GRU slightly outperforming LSTM. Results demonstrate the framework's ability to capture endogenous patterns, reducing data collection costs and enhancing applicability in legacy systems. Limitations include site-specific data, suggesting future enhancements via data augmentation, transfer learning, or hybrid models for broader generalizability. This approach supports sustainable energy management by minimizing fossil fuel reliance and operational costs.

Cite this article: Hajian, M., Sami, S. & Nazghelichi, T. (2025). Deep Learning Framework for Joint Prediction of Energy Resources in an Industrial Building. *Biomechanism and Bioenergy Research*, *4*(4), 95-109. https://doi.org/10.22103/bbr.2025.25960.1134

Publisher: Shahid Bahonar University of Kerman
DOI: https://doi.org/10.22103/bbr.2025.25960.1134

INTRODUCTION

Renewable energy sources, including solar and wind, are volatile and unpredictable, presenting significant challenges in ensuring a reliable supply. This necessitates the employment of advanced energy-storing technologies and the upgrading of transmission infrastructure (Ramsebner et al., 2021).

The application of artificial intelligence (AI) in energy administration infrastructure is transforming the response of power plants and other infrastructures toward addressing issues of energy production and consumption. AI allows assimilation of advanced data, modeling-based forecasting, and real-time optimization, which underpin the basis of enhancing energy efficiency and diminishing its carbon footprint. AI-based applications, for instance, can forecast energy demand through the application of deep learning architectures, namely convolutional neural networks (CNN) and long short-term memory (LSTM) networks, thus optimizing energy consumption and reducing cost (Ishaq & Kwon, 2021; Kristian et al., 2024).

Good energy administration, grid stability, and accommodation of renewable energy depend partly on short-term forecasting of photovoltaic (PV) solar energy production. Natural variation of solar energy, which depends on climatic conditions, creates challenges for energy grids and markets. Accurate short-term forecasting constitutes the basis of optimizing energy production, energy storage, and scheduling of distribution, thus diminishing the use of fossil fuel backup infrastructure) Agoua et al., 2017).

More advanced machine learning models, including Long Short-Term Memory (LSTM) networks and hybrid models uniting Convolutional Neural Networks and LSTM (CNN-LSTM), have demonstrated remarkable enhancements in the accuracy of predictions. These models make use of both past datasets and current real-time weather data in order to predict photovoltaic (PV) generation, and in this way increase the reliability of operations of the grid (Jakoplić et al., 2023; Wang et al., 2024). Such advances hold key significance in enhancing the efficiency of energy trade, lowering operational expenditures, and aiding the worldwide shift toward renewable energy infrastructure (Cao et al., 2022; Zhang et al., 2024).

Load forecasting usually refers to the electric demand forecast of a system at different time horizons. It is usually divided into four specific types depending on the horizon of forecasting: long-term, medium-term, short-term, and ultra-short-term (Xiuyun et al., 2018). Conditional modeling is exceptionally prevalent in long-term and medium-term load forecasting, respectively, for LTLF and MTLF applications. Here, a time series of physical variables, for example, moisture content and temperature, and of economic variables, for instance, gross domestic product (GDP) and consumer price index (CPI), act as inputs of the forecasting models (Oreshkin et al., 2021). In contrast, the autonomous approach is generally applied in both short-term load forecasting (STLF) and very short-term load forecasting (VSTLF), which depend exclusively on historical consumption data (Jiang et al., 2022). STLF is tasked with forecasting electricity demand for the forthcoming 24 hours or several days. The forecasting process is a key constituent involved in the running of power systems, both day-ahead and real-time, given that it aids in running and guiding day-to-day operations through releasing accurate estimates of demand loads. It can fluctuate drastically in the short term under the impact of variables ranging from weather conditions and volatility of energy prices, vacation seasons, and consumer demand (Mounir et al., 2023; Zhang, 2022).

Despite significant developments in the application of deep neural networks in short-term photovoltaic and load forecasting, most of the current research significantly relies on the combination of meteorological variables and special temporal characteristics (e.g., irradiance, temperature, time of day, and calendar effects) for enhancing the accuracy of the models (Sami et al., 2025). Few research efforts have examined whether dependable forecasts can be generated using solely raw historical time-series data (without auxiliary weather measurements or engineered temporal inputs) (Sami et al., 2025). In this work, we address that gap: the proposed LSTM-based framework leverages endogenous patterns within historical energy records to produce simultaneous, short-term forecasts of multiple energy resources in industrial buildings, without relying on meteorological inputs or explicit time-stamp features. By avoiding external sensors and supplementary datasets, our approach minimizes data collection efforts, improves applicability in environments with incomplete or legacy instrumentation, and provides a cost-effective alternative for implementing predictive capabilities within energy management systems.

MATERIALS AND METHODS

Dataset

The dataset used in this study was established over a year by a German research consortium and obtained from the Open Power System Data repository. It conforms to ENTSO-E and ENTSO-G standards and spans varying spatial scales (i.e., micro and large) (Open Power System Data, n.d). Therefore, it is appropriate for analysis and modeling at varying system levels. It comprises minute-by-minute time-series data of photovoltaic (PV) generation and electrical consumption registered in households and small office buildings in Konstanz, Germany. It includes information corresponding to the device level measured in kilowatt-hours (kWh). The coverage and quality over time vary between units, and gaps are present, ranging from a few minutes to days; however, since meters register cumulative energy, total energy values remain accurate regardless of communication outages. All Measurements are converted to uniform one-minute intervals, and missing values are replaced by linear interpolation for better processing. Either taking over values corresponding to the preceding day. Besides, datasets aggregated over intervals of 15 and 60 minutes are processed and provided for other stages of time-series data to be compatible with other data. The subset corresponding to applied modeling purposes in this work is that regarding one industrial building and spans corresponding to PV generation, consumption onsite, and energy supplied from the public grid (kWh).

Simultaneous forecasting of PV generation, electrical consumption, and grid imports

This section utilizes data sourced from an industrial structure. Initially, the dataset underwent examination and was subjected to preliminary exploratory analysis to identify its prominent features. In order to enhance both model accuracy and efficiency, the data were standardized through the application of a Min–Max scaler, which restricts feature values within a predetermined range and reduces superfluous fluctuations caused by magnitude.

Modeling after pre-processing was conducted with a recurrent neural network of the LSTM type. It was trained so that it would predict hourly the two PV systems' output power, on-site electrical consumption, and energy imported from the grid based on the preceding 24 hours of observation. The choice of architecture was such that it would capture complex long-term patterns and forecast likely trends over subsequent time intervals.

For the simultaneous prediction of four key variables, including industrial grid import (DE_KN_industrial3_grid_import), total consumption (total_consumption), PV facade generation, and PV roof generation, a recurrent neural network model based on GRU was employed. The input data were first scaled and then converted into time sequences with a length of 24 hours (equivalent to one day). These sequences were created using the create_sequences function, where each sequence included input features and multiple output labels (4 target variables). The data were divided into training (80%) and testing (20%) sets, without temporal shuffling to preserve the chronological order.

The model was implemented using the Keras library and consisted of two GRU layers (each with 50 units) along with Dropout layers at a rate of 0.2 to prevent overfitting. The final Dense layer with 4 outputs was designed for simultaneous prediction of the variables. The model was compiled with the Adam optimizer and mean squared error (MSE) loss function. Model training was conducted for a maximum of 30 epochs with a batch size of 32, and EarlyStopping (with patience of 10 epochs) and ReduceLROnPlateau (with a reduction factor of 0.5 and patience of 5 epochs) callbacks were used to enhance performance.

To achieve a more robust evaluation and mitigate the impact of random data splitting, 5-fold cross-validation was applied to both LSTM and GRU architectures. In each fold, 80% of the data was used as the training set, and the remaining 20% as the validation set. The metrics R^2, MAE, MSE, and RMSE were calculated for each fold, and their mean and standard deviation were reported. The final evaluation of the GRU model was also performed based on separate test data (time-series split).

Performance was measured using the dataset divided into training and test sets so that 80% of the samples were used to train the model and the other 20% for testing purposes. In the training phase, hyperparameters such as the number of layers, the learning rate, and the number of hidden units were adjusted with extreme care to optimize the results. These values were selected through preliminary experiments: smaller architectures underfit the time-series patterns, larger ones caused overfitting, and a learning rate of 0.001 provided stable convergence. After training the model, it was used to produce the forecast, and the final predictions were measured using four evaluation criteria: MSE, MAE, R^2, and RMSE.

Implementation details of the deployed model

The implemented deep-learning solution is an LSTM-based model for simultaneous prediction of energy-related outputs: PV generation, electrical consumption, and grid imports. The workflow consists of standard stages including data preprocessing, sequence construction, model definition, training, and evaluation.

Sequences were created to capture temporal dependencies in the dataset. A create_sequences function processes the input dataframe and generates sequences of a fixed length (24 hours), extracting the corresponding multivariate prediction targets. The selected target variables for prediction are:

- DE_KN_industrial3_grid_import
- total_consumption
- PV facade generation
- PV Roof generation

Each sequence contains multiple time steps; the function returns NumPy arrays of the processed sequences and their associated labels. The dataset was split into training and test sets using an 80–20 ratio while maintaining chronological order (no shuffling) to preserve temporal continuity.

The LSTM layers were set to 50 units because this size provided a good balance between model capacity and computational efficiency in preliminary experiments. Smaller sizes (e.g., 16 or 32 units) were unable to capture the temporal dynamics of the multivariate time-series data, while larger sizes (e.g., 64 or 128 units) increased training time and led to over fitting without notable improvements in accuracy.

ReLU was chosen as the activation function because it accelerates convergence, avoids vanishing-gradient issues commonly seen with tanh in deeper architectures, and performed better in initial tuning rounds.

Dropout with a rate of 0.2 was selected to reduce over fitting. Rates below 0.1 had little regularization effect, while higher values (0.3–0.5) significantly degraded the model's ability to learn temporal dependencies by removing too much information from the recurrent layers.

The dense output layer contains four neurons because the model simultaneously predicts four target variables (PV generation, consumption, grid import, etc.), making a four-dimensional output layer necessary.

Table 1. Neural network architecture for simultaneous forecasting of generation and consumption

Layer type	Units / Rate	Activation
LSTM	50	ReLU
Dropout	0.2	—
LSTM	50	ReLU
Dropout	0.2	—
Dense (output)	4	—

The model was compiled with the Adam optimizer and the mean squared error (MSE) loss function, which is appropriate for regression tasks. Training was performed for 30 epochs with a batch size of 32. Two callbacks were employed to optimize training:

- **EarlyStopping:** Monitors validation loss and halts training if no improvement is observed for 10 consecutive epochs.
- **ReduceLROnPlateau:** Reduces the learning rate by a factor of 0.5 if the validation loss remains stagnant for five epochs, improving convergence behavior.

2.1 Evaluation metrics

Evaluating predictive models requires robust performance metrics to quantify accuracy and reliability. This report focuses on four widely used metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These metrics are central to model assessment across diverse fields, offering unique insights into error magnitude, variance explanation, and prediction quality.

Mean Squared Error (MSE) and Root Mean Squared Error (RMSE)

MSE and RMSE are widely used metrics for evaluating predictive models. MSE measures the average squared difference between predicted and observed values, defined mathematically as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (1)$$

Where y_i represents observed values, $\hat{y}_i$ represents predicted values, and n represents the number of observations. RMSE is the square root of MSE, expressed as:

$$RMSE = \sqrt{MSE} \quad (2)$$

RMSE retains the same unit as the original data, making it more interpretable in practical contexts (Chai & Draxler, 2014; Hodson, 2022).

Both metrics are sensitive to significant errors due to the squaring operation amplifying outliers. This property makes them particularly useful in applications where penalizing significant deviations is critical, such as forecasting solar and wind energy outputs or air pollution levels (Kumar et al., 2025; Mazen et al., 2023).

Mean Absolute Error (MAE): MAE is a widely used metric that measures the average magnitude of prediction errors without considering their direction, making it a straightforward and interpretable measure of model performance**.** It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (3)$$

Coefficient of Determination (R^2): R^2 quantifies the proportion of variance in the dependent variable explained by the independent variables in a regression model. Mathematically, it is expressed as:

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} \quad (4)$$

Where $SS_{residual}$ is the sum of squared residuals and SS_{total} is the total sum of squares. An R^2A value of 1 indicates perfect prediction, while 0 implies no explanatory power (Li, 2017; Zhou et al., 2019).

However, R^2 has limitations. It always increases when additional predictors are added, even if irrelevant, leading to overfitting in complex models (Bar-Gera, 2017).

RESULTS AND DISCUSSION

This section will review and analyze the results, and based on these results, the implemented model will be verified separately.

Examining the model from the perspective of over fitting

An indication of over fitting when the results on the training and test data are not the same. The main way to check for over fitting is to compare the model's training results with its performance on test data (Sami et al., n.d.). In this method, if the model performs well on test data, it indicates that it was adequately trained and overfitting has been avoided. As shown in**Error! Reference source not found.**, the average results of the model on both training and test data were very similar. This close match suggests that the model is in good shape and shows no signs of overfitting. In other words, the model has not become too familiar with the training data and has learned the overall features of the data correctly. Therefore, the hypothesis of overfitting is rejected, and it can be concluded that the model is in a favorable condition regarding generalizability and performance on new data.

Table 2. Average results of the implemented model based on the evaluation criteria

Error and accuracy parameters	Training data	Evaluation data
R^2	0.918	0.906
MAE	0.028	0.033
MSE	0.003	0.004
RMSE	0.053	0.061

Overall validation of the implemented model based on four defined error parameters

After running the model and completing the data processing, the results were carefully reviewed and analyzed. The best result occurred at epoch 29, where the lowest error was recorded. The overall error was averaged for predicting the output power of two photovoltaic panels installed in the building, the electricity imported as recorded by the meter, and the total building consumption on an hourly basis. These predictions were generated using a model designed for simultaneous forecasting based on historical data from the systems, and the results were separately recorded in Table 3**Error! Reference source not found.**, according to the desired accuracy and error criteria.

All prediction parameters were assessed using the R^2 index on experimental data, and the results are presented in Table 3.These findings illustrate the model's prediction accuracy with the R^2 index and support performance evaluation. Additionally, the table lists the values for each parameter separately, aiding in a more detailed examination of the model's performance and a comparison between predictions and actual values. After conducting a detailed model analysis, the next step was to review the model outputs and assess their accuracy. This process was carefully and separately carried out for each output to evaluate their correctness and error rates.

Table 2. Results of the implemented model based on error criteria and the coefficient of determination for all parameters

Validation	**Test data**				**Training data**			
Validation parameters	Import from the network	Total consumption	First photovoltaic panel production	Second photovoltaic panel production	Import from the network	Total consumption	First photovoltaic panel production	Second photovoltaic panel production
R^2 Score	0.899	0.896	0.906	0.934	0.958	0.903	0.898	0.932
MAE	0.029	0.028	0.035	0.035	0.023	0.027	0.028	0.027
MSE	0.002	0.002	0.005	0.004	0.001	0.002	0.004	0.003
RMSE	0.048	0.049	0.072	0.068	0.040	0.045	0.063	0.055

As shown in Table 3**Error! Reference source not found.**, the model was validated using two different datasets, including training and test data. The results showed that the model outputs closely match the values of both datasets, helping to rule out over fitting by examining all errors. In summary, the model can generalize learned patterns to new data, improving its validity.

Furthermore, because the model makes hourly predictions and exhibits large fluctuations, it can be said that it has effectively captured these rapid and variable changes. The review of the model's variance and bias also yielded very positive results. Specifically, despite significant changes in the data, the model has been able to achieve high accuracy and precision in its predictions, demonstrating its ability to handle various challenges.

Checking the model performance on the original values

After evaluating the model's error and accuracy parameters, 10 samples were randomly selected from the data set to further examine the level of agreement between the predicted and actual values. These samples underwent detailed analysis to assess the model's performance in predicting the target variables. The results, shown in Table 4, where the values of each feature are provided in kilowatt-hours, indicate that the predicted values align acceptably with the actual values. This agreement demonstrates the model's desired accuracy and generalization ability to unseen data. Therefore, the model can identify trends in the data and offer reliable predictions with an acceptable level of error.

Table 3. Actual values and model predictions in 10 random data points

The actual amount of network imports	Network import forecast value	The actual amount of total consumption	Predicted total consumption amount	Actual amount of display panel production	Projected amount of display panel production	Actual ceiling panel production amount	Projected amount of ceiling panel production
0.721	0.57	0.778	0.603	0.895	0.81	0.776	0.674
0.252	0.26	0.272	0.274	0	0.006	0	0.0001
0.251	0.278	0.284	0.294	0.102	0.136	0.197	0.191
0.631	0.535	0.628	0.565	0.512	0.271	0.499	0.375
0.202	0.249	0.278	0.278	0.832	0.407	0.709	0.413
0.269	0.274	0.292	0.293	0	0.0004	0.003	0.0086
0.482	0.471	0.526	0.531	0.279	0.323	0.726	0.631
0.248	0.248	0.272	0.259	0.001	0.012	0.005	0.006
0.55	0.499	0.524	0.512	0.116	0.201	0.564	0.47
0.243	0.279	0.27	0.292	0.063	0.152	0.13	0.192

Investigating the performance of the model to predict photovoltaic generation

The prediction of photovoltaic energy generation in a grid equipped with two types of solar panels, including rooftop and ground-mounted panels, was investigated. To evaluate the performance of the prediction model, the results were analyzed on the entire test data and 300 test samples separately. The findings are presented in the form of graphs shown in **Error! Reference source not found.**, **Error! Reference source not found.Error! Reference source not found.**, **Error! Reference source not found.**, and Figure 4.

.

The studies for both types of panels showed that, despite significant fluctuations in the recorded data, the prediction model could produce satisfactory and reliable results. These findings demonstrate the model's ability to adapt to changes in energy output for both panel types and to estimate photovoltaic production under various conditions accurately

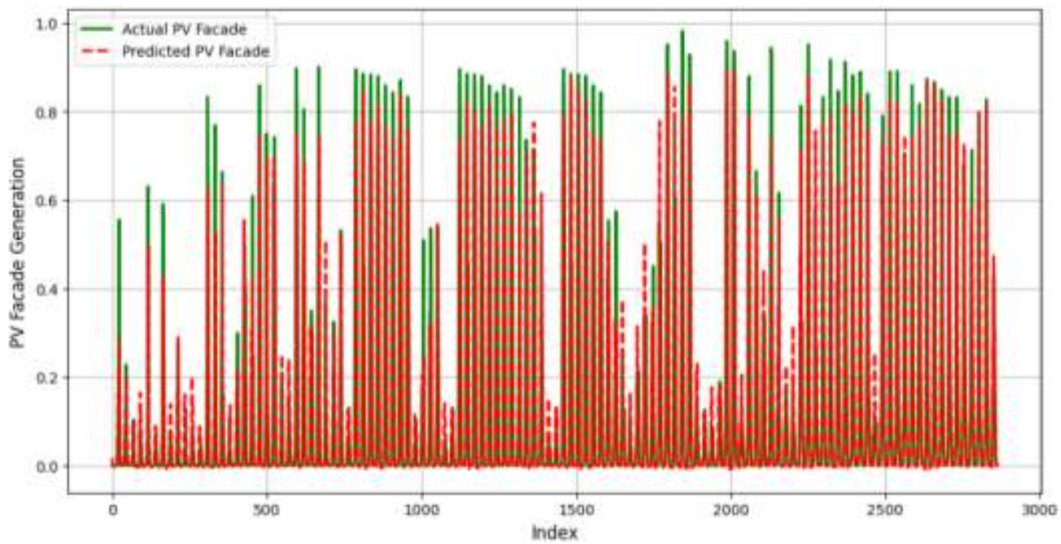


Figure 1. Comparison of actual and predicted values in the test data for the first panel.

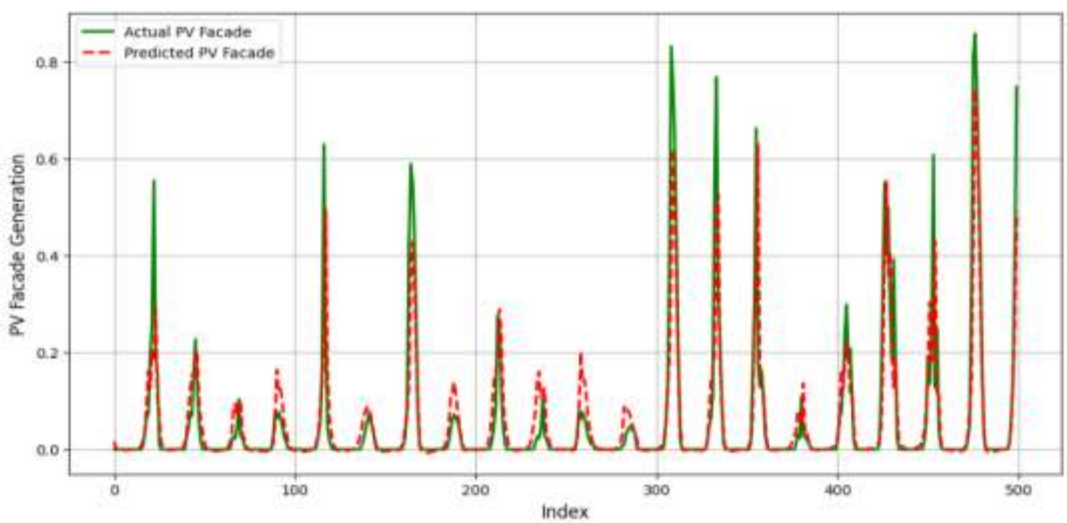
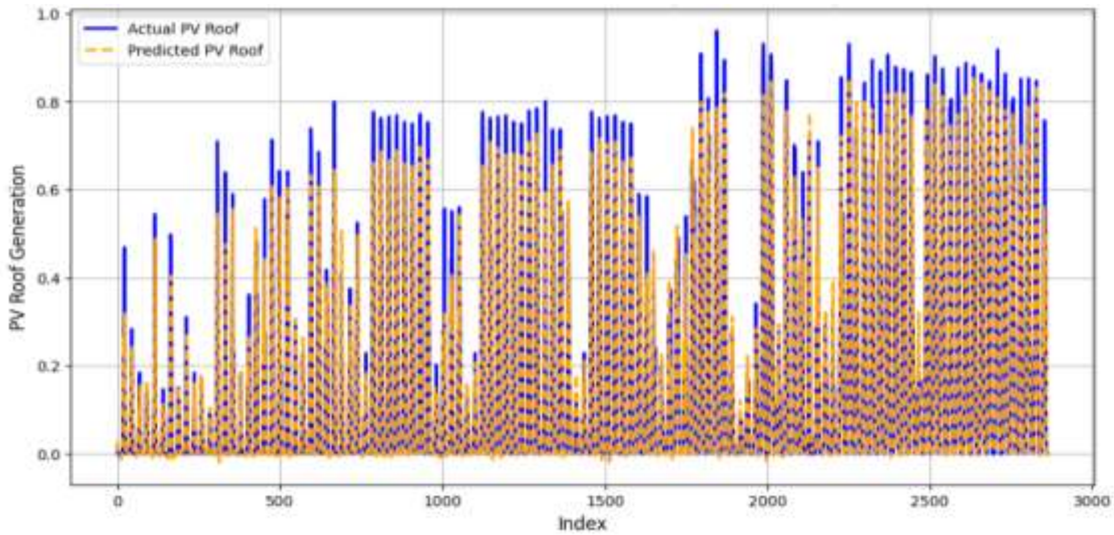


Figure 2. Comparison of actual values and model predictions on test data in the second panel.



.

Figure 3. Comparison of actual values and model predictions on test data in the second panel

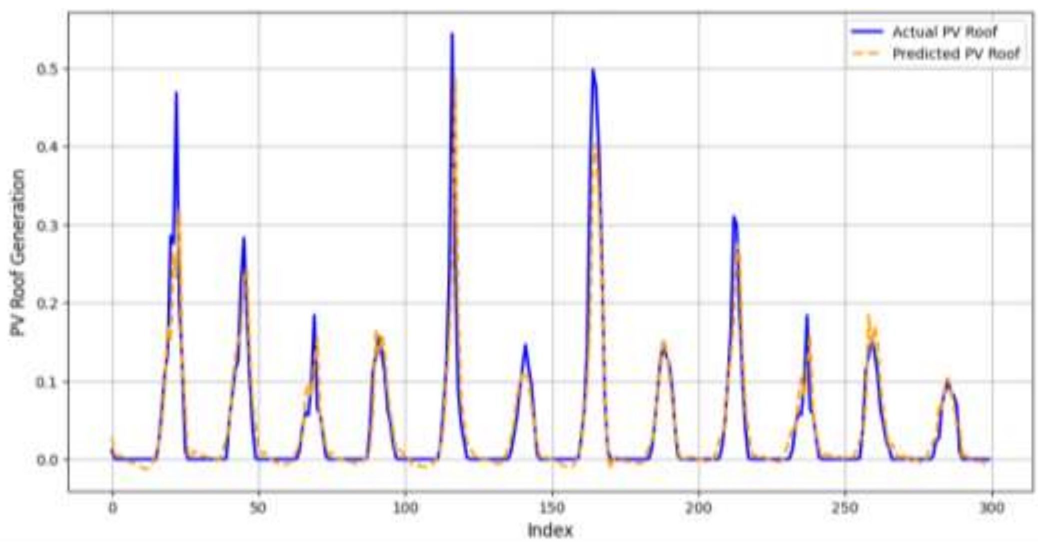


Figure 4. Comparison of predicted and actual values in 300 test data in the first panel.

Given the problem studied in this research, it is important to evaluate the model's accuracy. The model is designed and implemented with four inputs and four outputs. The results show that the model's accuracy has decreased somewhat compared to predicting photovoltaic production independently.

The accuracy of the rooftop photovoltaic panel was higher, likely because of consistent sunlight exposure and the lack of shadows and other experimental factors affecting the building under study.

However, this reduction in accuracy is understandable because the model outputs include highly variable elements, such as energy consumption and grid imports. Therefore, the model's performance remains quite acceptable despite its complexity, and the results are still valid and widely applicable.

Examining the efficiency of the model to predict the total building consumption

The partial consumption of various building equipment, including dishwashers, electric vehicle charging stations, air conditioning and heating devices, and other low-power appliances, was examined. These values were available in the corresponding columns of the dataset, and the prediction model was run on the total of these partial consumptions to assess its effectiveness in estimating energy use in different parts of the building.

To evaluate the model's accuracy, predicted values were compared with actual values at two levels: the entire dataset and a sample of 300 test data points. This comparison aimed to assess how well the predictions matched the real data. The results are shown in graphs in Figure 5 and Figure 6, enabling a detailed review of the model's accuracy in predicting these costs.

The results of the studies showed that the presented model, considering the type of problem, was able to provide predictions with acceptable accuracy close to the actual values. Given the variable nature of energy consumption and its unpredictable patterns, along with the multi-output nature of the model, assessing the prediction accuracy is highly important. Analysis of the results indicates that the developed model, despite existing challenges, has performed well and its prediction accuracy is at a very acceptable level.

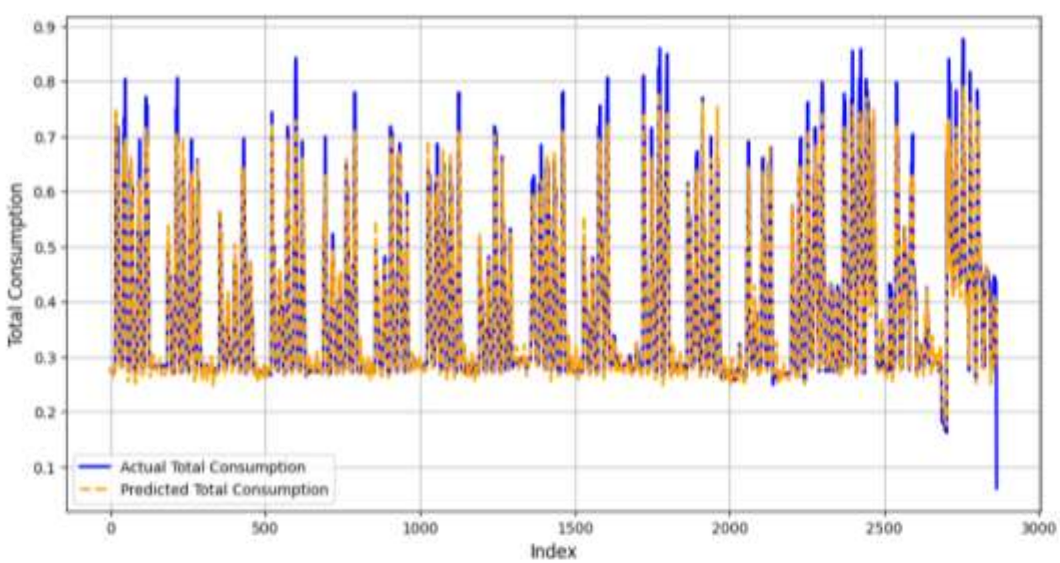


Figure 5. Compare actual values and model predictions to predict consumption across the entire test data.

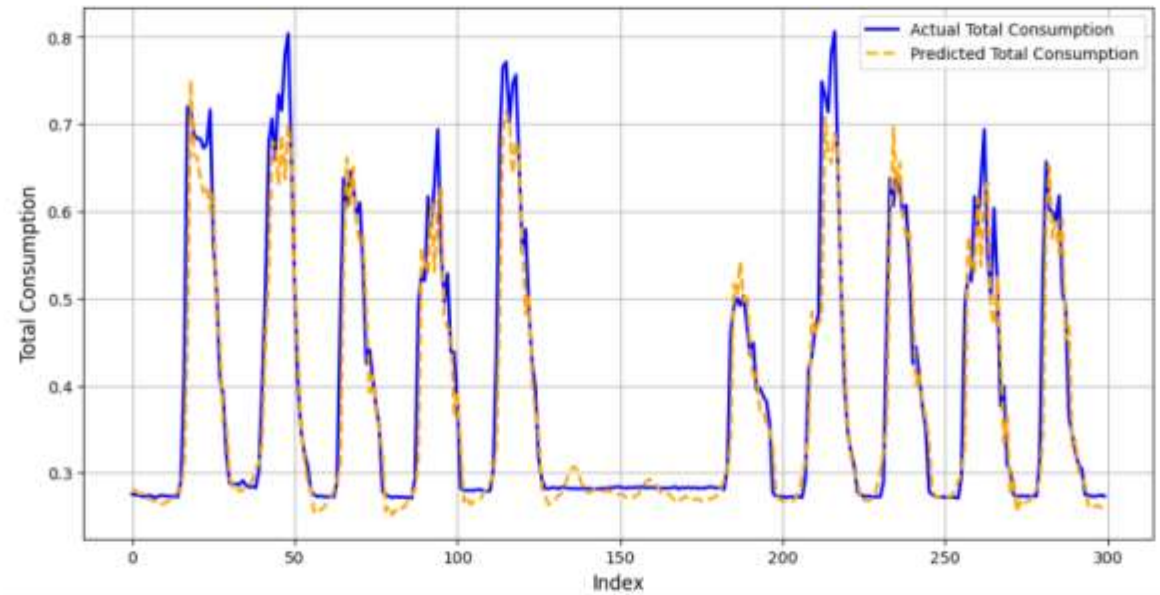


Figure 6. Comparison of actual values and model predictions for consumption prediction on 300 test data.

Model results for predicting electricity imports from the grid

The model's accuracy in predicting network traffic was also evaluated in line with the previous two features. This assessment was carried out on the entire test data set as well as 300 individual test cases to enable more detailed analysis. The results of this assessment are shown in Figure 7 and 8.

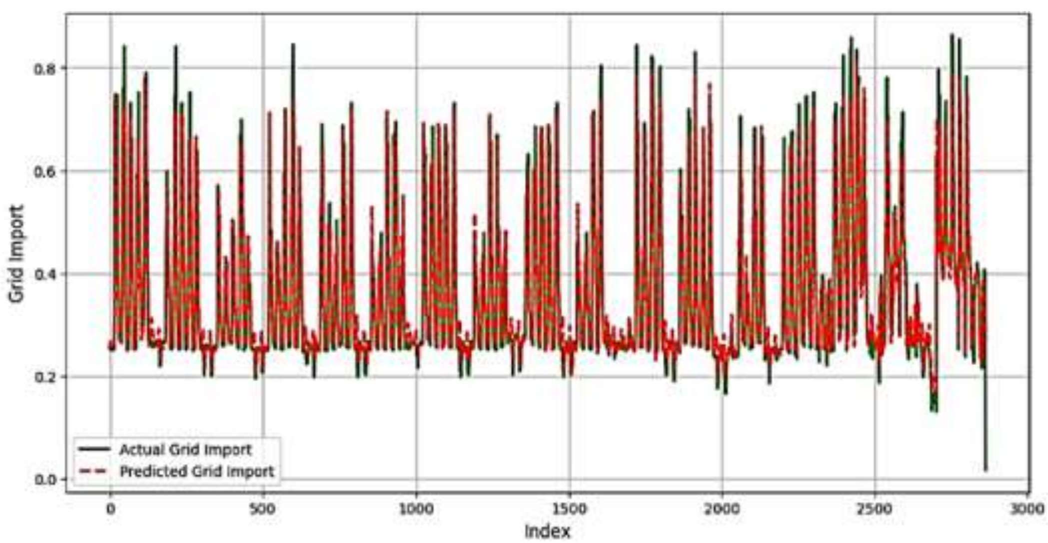


Figure 7. Compare actual values and model predictions on all test data.

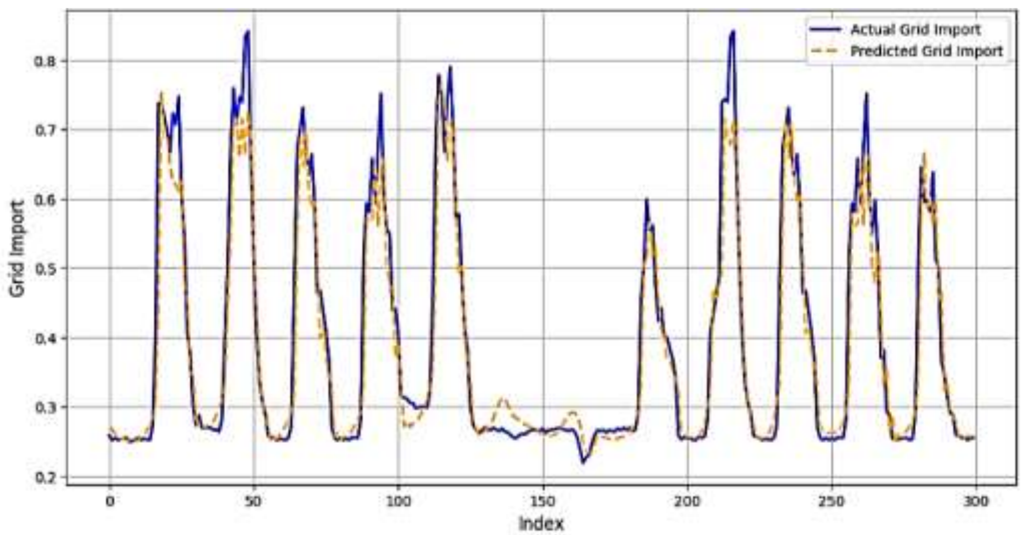


Figure 8. Comparison of actual values and model-predicted data on 300 test data.

The results of predicting the amount of electricity imported from the grid showed that the model performed acceptably, given the nature of the problem. The model was able to capture a significant portion of the data variance and accurately identify patterns in the changes of electricity imports. These findings suggest that, relative to the complexity of the problem, the presented model provides predictions with the desired accuracy, and its ability to estimate the amount of imported electricity has been confirmed.

Examining model hyperparameters on prediction

The sequence length (SEQ_LENGTH) determines how much historical data the model uses to predict future values. The value of 24 was selected as the optimal choice because of the data's periodic nature and daily patterns. Values lower than this miss important information, while higher values make the model more complex and reduce performance.

The number of LSTM units in each layer is set to 50. After testing various values, this number strikes a good balance between model accuracy and computational efficiency. Values below this reduce prediction accuracy, while higher values increase training time and the risk of overfitting.

To reduce overfitting, a dropout rate of 0.2 was used in both LSTM layers. This rate was adjusted to prevent the model from relying too heavily on the training data while maintaining important information. Values higher than 0.3 lowered the model's accuracy, and lower values did not effectively prevent overfitting.

The cost function used is Mean Squared Error (MSE), which is appropriate because of its suitability for regression problems. Additionally, the Adam optimizer was chosen as the best option after comparing it with other methods like SGD and RMSprop, given its fast convergence and ability to adjust the learning rate.

The model is set for 30 training sessions with a Batch Size of 32. The optimal number of sessions is chosen to ensure the model learns adequately without overfitting. The Batch Size was tested with values of 16, 32, and 64, and 32 provided the best balance between training speed and model convergence.

To prevent overtraining and optimize the learning rate, two mechanisms, Early Stopping and ReduceLROnPlateau, are used. The Early Stopping mechanism causes the model training to halt at the optimal point and prevents unnecessary increases in training time, which is crucial for improving the model's efficiency.

GRU model results

The proposed GRU model demonstrated excellent performance in the simultaneous prediction of the four target variables (industrial grid import, total consumption, PV facade generation, and PV roof generation). The final evaluation on the separate test set (using time-series split) indicated high accuracy: an R^2 score of 0.906, mean absolute error (MAE) of 0.034, mean squared error (MSE) of 0.0038, and root mean squared error (RMSE) of 0.062. The model's performance on the training data, with an R^2 of 0.917, MAE of 0.028, MSE of 0.0028, and RMSE of 0.053, reflects effective learning of temporal patterns and the absence of significant overfitting.

To assess model stability, 5-fold cross-validation was applied to both LSTM and GRU architectures. The results of this cross-validation are presented in Table 5. The mean R^2 score for both models was approximately 0.89, indicating high consistency across different folds.

Table 5. Results of 5-fold cross-validation for LSTM and GRU models

Fold	LSTM R^2	GRU R^2	LSTM MAE	GRU MAE	LSTM RMSE	GRU RMSE
1	0.887	0.892	0.036	0.034	0.065	0.062
2	0.894	0.890	0.035	0.035	0.063	0.063
3	0.882	0.885	0.037	0.036	0.067	0.064
4	0.896	0.893	0.034	0.034	0.062	0.061
5	0.889	0.891	0.036	0.035	0.065	0.062
Mean ± Std	0.890 ± 0.005	0.890 ± 0.003	0.036 ± 0.001	0.035 ± 0.001	0.064 ± 0.002	0.062 ± 0.001

The obtained results demonstrate that the GRU-based model can effectively learn and predict complex temporal patterns in energy consumption and solar generation data with high accuracy. The R^2 score exceeding 0.90 on the separate test set confirms the model's strong generalization capability, making it suitable for real-world applications. The small difference between training and test performance metrics (e.g., R^2 of 0.917 vs. 0.906) indicates that overfitting was successfully mitigated through Dropout layers and early stopping mechanisms.

The 5-fold cross-validation confirmed the high stability of both LSTM and GRU architectures, with a mean R^2 of approximately 0.89 and very low standard deviations (less than 0.005), reflecting low sensitivity to variations in data splitting. Nevertheless, the GRU model exhibited slightly superior performance in most metrics (particularly RMSE and MAE), which can be attributed to its simpler structure (one fewer gate compared to LSTM), leading to faster training and reduced risk of overfitting. This advantage led to the selection of GRU as the final architecture.

Overall, the findings suggest that recurrent neural networks based on GRU represent an efficient and reliable approach for multi-variable time series forecasting in the energy domain. This can contribute to improved energy management and better integration of renewable sources. Future studies could further enhance model accuracy by incorporating additional external features (such as more detailed weather data) or hybrid architectures (e.g., combining GRU with attention mechanisms).

Limitations and Suggestions

Although the GRU model shows strong performance on the current dataset, several limitations should be noted. The dataset is derived from a specific industrial site in Germany (DE_KN_industrial3), which may limit its applicability to other geographic regions, climates, or industrial facilities. For example, the data primarily capture seasonal variations in the temperate European climate and potentially ignore extreme weather patterns in tropical or arid regions that could affect PV production. Furthermore, the time span of the dataset may not capture long-term trends such as equipment degradation or evolving energy policies, leading to reduced generalizability. The scaling and sequencing processes assume stationary in the time series, which may not hold in dynamic real-world scenarios with abrupt changes (e.g., due to economic changes).

To increase generalizability, future strategies could include data augmentation techniques, such as combining synthetic data generated through Generative Adversarial Networks (GANs) to simulate diverse environmental conditions. Transfer learning from pre-trained models on broader energy datasets (e.g., from public repositories such as UCI or Kaggle) could adapt the model to new domains. Furthermore, ensemble methods that combine GRU with other architectures (e.g., transformers) may improve robustness. Cross-validation across multiple sites or federated learning approaches could further validate the model's performance in heterogeneous environments.

CONCLUSIONS

In conclusion, this study successfully demonstrates the efficacy of LSTM and GRU-based models for simultaneous short-term forecasting of multiple energy variables—PV generation, electrical consumption, and grid imports—using only historical time-series data from an industrial building. By eschewing external meteorological inputs and engineered features, the proposed framework offers a cost-effective, scalable solution that simplifies implementation in environments with limited instrumentation. The GRU model, in particular, exhibited superior performance with high R^2 scores (up to 0.906 on test data) and low error metrics, validated through rigorous cross-validation and hyperparameter tuning, confirming its robustness against overfitting and ability to handle data variability.

These findings underscore the potential of recurrent neural networks in advancing renewable energy integration, optimizing grid operations, and reducing carbon footprints without additional sensor dependencies. However, the model's reliance on a single-site dataset highlights limitations in generalizability across diverse climates or facilities. Future research could incorporate multi-site data, advanced techniques like attention mechanisms or GAN-based augmentation, and ensemble methods to enhance adaptability and accuracy, ultimately contributing to more resilient and sustainable energy systems worldwide.

REFERENCES

Agoua, X. G., Girard, R., & Kariniotakis, G. (2017). Short-term spatio-temporal forecasting of photovoltaic power production. *IEEE Transactions on Sustainable Energy*, *9*(2), 538-546. https://doi.org/10.1109/TSTE.2017.2747765

Bar-Gera, H. (2017). The target parameter of adjusted R-squared in fixed-design experiments. *The American Statistician*, *71*(2), 112-119. https://doi.org/10.1080/00031305.2016.1200489

Cao, H., Wang, T., Chen, P., Cheng, W., Cao, Y., & Liu, Z. (2022). Solar energy forecasting in short term based on the aso-bpnn model. *Frontiers in Energy Research*, *10*, 902486. https://doi.org/10.3389/fenrg.2022.902486

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014

Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development Discussions*, *2022*, 1-10. https://doi.org/10.5194/gmd-15-5481-2022

Ishaq, M., & Kwon, S. (2021). Short-term energy forecasting framework using an ensemble deep learning approach. *IEEE Access*, *9*, 94262-94271. https://doi.org/10.1109/ACCESS.2021.3093053

Jakoplić, A., Franković, D., Havelka, J., & Bulat, H. (2023). Short-term photovoltaic power plant output forecasting using sky images and deep learning. *Energies*, *16*(14), 5428. https://doi.org/10.3390/en16145428

Jiang, Y., Gao, T., Dai, Y., Si, R., Hao, J., Zhang, J., & Gao, D. W. (2022). Very short-term residential load forecasting based on deep-autoformer. *Applied Energy, 328*, 120120. https://doi.org/10.1016/j.apenergy.2022.120120

Kristian, A., Goh, T. S., Ramadan, A., Erica, A., & Sihotang, S. V. (2024). Application of ai in optimizing energy and resource management: Effectiveness of deep learning models. *International Transactions on Artificial Intelligence*, *2*(2), 99-105. https://doi.org/10.33050/italic.v2i2.530

Kumar, S., Kour, V., Raj, A., Tapung, T., Mishra, S., Misra, R., & Singh, T. (2025). Optimizing Air Pollution Forecasting Models Through Knowledge Distillation: A Novel GCN and TRANS_GRU Methodology for Indian Cities. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3546504

Li, J. (2017). Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? *PloS one*, *12*(8), e0183250. https://doi.org/10.1371/journal.pone.0183250

Mazen, F. M. A., Shaker, Y., & Abul Seoud, R. A. (2023). Forecasting of solar power using GRU–temporal fusion transformer model and DILATE loss function. *Energies*, *16*(24), 8105. https://doi.org/10.3390/en16248105

Mounir, N., Ouadi, H., & Jrhilifa, I. (2023). Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system. *Energy and Buildings*, *288*, 113022. https://doi.org/10.1016/j.enbuild.2023.113022

Open Power System Data. (n.d.). *A platform for open data of the European power system*. https://open-power-system-data.org/

Oreshkin, B. N., Dudek, G., Pełka, P., & Turkina, E. (2021). N-BEATS neural network for mid-term electricity load forecasting. *Applied Energy*, *293*, 116918. https://doi.org/10.1016/j.apenergy.2021.116918

Ramsebner, J., Haas, R., Auer, H., Ajanovic, A., Gawlik, W., Maier, C., Nemec-Begluk, S., Nacht, T., & Puchegger, M. (2021). From single to multi-energy and hybrid grids: Historic growth and future vision. *Renewable and Sustainable Energy Reviews*, *151*, 111520. https://doi.org/10.1016/j.rser.2021.111520

Sami, S., Hassan-Beygi, S. R., & Massah, J. (2025). Accurate PV Power Forecasting Using a Lightweight LSTM Model: A Case Study. *Biomechanism and Bioenergy Research*, *4*(3), 68-80. https://doi.org/10.22103/bbr.2025.25811.1132

Wang, R., Liu, X., Chang, Y., Liu, D., & Yao, H. (2024). Short-Term Photovoltaic System Output Power Prediction Based on Integrated Deep Learning Algorithms in the Clean Energy Sector. *International Journal of e-Collaboration (IJeC)*, *20*(1), 1-15. https://doi.org/10.4018/IJeC.346979

Xiuyun, G., Ying, W., Yang, G., Chengzhi, S., Wen, X., & Yimiao, Y. (2018). Short-term load forecasting model of GRU network based on deep learning framework. 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2),

Zhang, J., Huang, K., & Yang, Y. (2024). Short-term photovoltaic power prediction based on improved EMA variable weight combination. 2024 6th International Conference on Energy, Power and Grid (ICEPG),

Zhang, W. (2022). Short-term load forecasting of power model based on CS-Catboost algorithm. 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC),

Zhou, Z.-R., Wang, W.-W., Li, Y., Jin, K.-R., Wang, X.-Y., Wang, Z.-W., Chen, Y.-S., Wang, S.-J., Hu, J., & Zhang, H.-N. (2019). In-depth mining of clinical data: the construction of clinical prediction model with R. *Annals of translational medicine*, *7*(23), 796. https://doi.org/10.21037/atm.2019.08.63