This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI10.1109/UPEC65436.2025.11279840, 2025 60th International Universities Power Engineering Conference (UPEC)

Hybrid Deep Learning Framework for Long-Term Forecasting of PV Generation

Ljubomir Babic
Department of Electronic and Electrical Engineering
Brunel University of London
Uxbridge, United Kingdom
2438022@brunel.ac.uk

Marko Aunedi
Department of Electronic and Electrical Engineering
Brunel University of London
Uxbridge, United Kingdom
marko.aunedi@brunel.ac.uk

***Abstract*—Accurate forecasting of photovoltaic (PV) generation is critical for renewable energy integration and maintaining system reliability. This paper demonstrates the application of a hybrid deep learning framework based on Convolutional Neural Network (CNN) and Long-Short-Term Memory (LSTM) for forecasting PV generation in the UK. The proposed model uses a 45-year dataset (1980-2024) of simulated hourly PV output values, allowing it to capture local temporal patterns through convolutional layers and long-term dependencies via bidirectional LSTM layers. Cyclical temporal features (hour of the day, day of the week, and month of the year) are feature engineered to encode periodicity. Experimental results show that the predictions made by a hybrid CNN+LSTM are very close to the actual PV generation values, with a Mean Absolute Error (MAE) of 0.0061, Root Mean Squared Error (RMSE) of 0.0121, and Mean Absolute Percentage Error (MAPE) of 4.86% on the test set.**

Keywords—deep learning, photovoltaics, forecasting, convolutional neural networks, long-short-term memory.

I. Introduction

Solar photovoltaic (PV) deployment is rapidly accelerating worldwide, driven by decarbonization goals and decreasing technology costs. As the penetration of PV and other variable renewable sources increases, accurate forecasting of PV output has become crucial for the efficient operation and design of low-carbon electricity systems. Robust PV output forecasts are essential for both short-term operational planning and long-term, cost-efficient investment planning.

Recent advances in machine learning (ML) have led to the development of new approaches for forecasting renewable output and electricity demand. These approaches can extract and replicate key features from large datasets, enabling the creation of more advanced forecasting models.

Time series forecasting (TSF) models using deep learning [1,5] rely heavily on open-source benchmark datasets for training and testing. According to [1], forecasting the industrial energy demand most frequently use the open-source datasets Electricity Transformed Temperature (ETT) [2,4] and Solar Energy and Electricity [3]. The former comprises two years' worth of load and oil temperature data from power transformers, collected across two Chinese regions, while the latter provides 2006 data on solar PV output from 137 PV stations in Alabama and electricity demand from 321 customers between 2011 and 2014.

Before the rise of ML and deep learning, the TSF research was dominated by traditional statistical methods such as autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) [1,5]. One example of using an adaptive ARMA model [10] showed that conventional ARMA models based on the Box-Jenkins method lack real-time error adaptation, leading to less accurate forecasts for 24-hour and one week ahead predictions. To address this limitation, the authors proposed an adaptive ARMA algorithm that updates forecasts using one step ahead errors, and outperformed the conventional Box-Jenkins approach. The main disadvantage of statistical methods, when compared with ML and deep learning approaches, is that they depend on fixed, linear equations based on predefined assumptions for forecasting. These dependencies on assumptions limit the ability of statistical models to learn complex patterns in time series data [5]. By contrast, ML and deep learning models can outperform statistical models because they do not rely on predefined assumptions and instead learn nonlinear relationships directly from time series data [1,5].

Wen et al. [11] proposed a hybrid deep learning model that combines recurrent neural networks (RNNs) with long-short-term memory units (LSTM) to forecast residential power demand and photovoltaic (PV) generation in community microgrids. Their model outperformed traditional ML methods such as support vector machines (SVMs) and multi-layer perceptron (MLP) achieving a mean absolute percentage error (MAPE) of 73% for load forecasting. Yan *et al.* [14] proposed a hybrid deep learning framework employing stationary wavelet transform (SWT) with an ensemble LSTM architecture to forecast energy consumption in individual households, outperforming several benchmark models. The authors of [16] analysed short-term load forecasting for power system operation using a hybrid algorithm that is based on similar days (SD) selection, empirical model decomposition (EMD), and LSTM. A physics-constrained LSTM framework for forecasting day-ahead PV generation was proposed in [24], incorporating domain-specific knowledge, such as zero generation at night and non-negative power outputs, achieving superior performance to a standard LSTM model.

While hybrid architectures [11,14,16] demonstrate advantages of deep learning over statistical methods in capturing nonlinear patterns for TSF, they remain constrained by their need for task-specific architectural adjustments, like statistical methods. In contrast, Yue *et al.* [2] proposed TS2Vec, a universal representation learning framework that leverages hierarchical contrastive learning to generate task-agnostic, timestamp-level embeddings. Unlike supervised hybrid models which optimise architectures for specific TSF tasks, TS2Vec extracts generalisable features from raw time series data, eliminating the need for domain-specific modification. Experimental results have shown a 32.6% decrease of average Mean Squared Error (MSE) compared with specialised models such as Informer [4] and TCN [15].

The most common metrics used for evaluating the performance of deterministic forecasting models are Mean Absolute Error (MAE) and Mean Squared Error (MSE) [1,2,4,5,17]. MAE and MSE are well-suited for evaluating deterministic models because they directly measure the

Copyright 'For the purpose of open access, the author has applied a 'Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.'

deviation between the unambiguous predictions and the actual observed values. In contrast to deterministic models, probabilistic models account for uncertainty by generating distributions of possible outcomes instead of single-value predictions. These models are typically evaluated using Continuous Ranked Probability Score (CRPS), which quantifies how close the predicted distribution aligns with the true observed value [1,5]. While MAE, MSE, and CRPS are effective for calculating the magnitude of errors, they cannot evaluate whether the model has effectively learned temporal patterns (e.g., seasonality, trends, or dependencies) and features from time series data.

To determine how well the model captures seasonal or periodic patterns, metrics like Mean Absolute Percentage Error (MAPE) [11,14,18] and Symmetric Mean Absolute Percentage Error (sMAPE) are often employed [1,5,11]. MAPE calculates the mean absolute percentage difference between predictions and observations, while sMAPE addresses MAPE's asymmetry by balancing the error relative to both the actual and predicted values. Furthermore, the coefficient of determination (R^2) measures the proportion of variance of in the observed data explained by the model's predictions [1], making this metrics useful for evaluating how well a model captures long-term trends in time series data [1,5,18]. Time series data often exhibits irregularities (e.g., sudden spikes or drops) that significantly impact metrics like Root Mean Squared Error (RMSE), which makes it a robust choice for assessing a model's resilience to outliers [5,14,17,18].

The main contributions of this study are summarized below:

1. Implementing a hybrid CNN+LSTM deep learning framework to forecast hourly PV generation in Brighton, UK, using 45 years of data (from January 1980 to December 2024) from MERRA-2 reanalysis dataset [19].
2. Explicitly modelling periodicity in solar patterns with cyclical features (hourly, weekly, and seasonal) using sine and cosine transforms.
3. Evaluating the model performance using MAE, RMSE, and MAPE metrics over a nine-year period (from January 2016 to December 2024), to determine the robustness of the model.

The remainder of this paper is organized as follows. Section II describes the methodology. The experimental results are presented in Section III. Finally, Section IV concludes this paper.

II. Methodology

A. Dataset

The solar PV data was obtained from Renewables.ninja website [20,21,22] which provides hourly power output from wind and solar power plants located anywhere in the world, and collected from January 1980 to December 2024. Renewables.ninja website uses two datasets to simulate renewable energy outputs: the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) dataset [19] and the Solar surfAce RAdiation Heliostat (SARAH) dataset [23].

The MERRA-2 dataset is a global meteorological reanalysis dataset produced by NASA's Global Modeling and Assimilation Office (GMAO). This dataset combines historical observational data (e.g., temperature, wind, humidity, precipitation) with advanced modelling techniques to generate a consistent, multidecade record or Earth's climate. In contrast, the SARAH dataset focuses on solar surface irradiance, direct irradiance (horizontal and direct normal) and clear sky irradiance, collected from 1980 to 2013. The SARAH dataset covers Africa, Europe, and parts of Atlantic, and Indian Oceans at high spatial resolution with hourly, daily, and monthly averages.

For the solar PV forecasting, Brighton, UK was chosen as the location and the data from Renewables.ninja website collected from the MERRA-2 dataset for was used for the experiments.

B. Data Pre-Processing Details

Hourly PV generation data from January 1980 to December 2024 was first downloaded from Renewables.ninja website as yearly .csv files (one .csv file per year). Each .csv file contains six features: Time, Local Time, Electricity (in kW), Direct Irradiance, Diffuse Irradiance, and Temperature. Then, multiple annual .csv files were merged into a single dataset using the Pandas Python library, skipping metadata rows to retain structured time-series values. After that, the data was concatenated chronologically to ensure a continuous hourly time series. Finally, the combined dataset was sorted by timestamp to avoid potential gaps and overlaps. Fig. 1(a) presents the hourly PV generation dataset for Brighton from 1980 to 2024, while Fig. 1(b) presents the zoom-in view of the hourly PV generation dataset for Brighton for 72 hours.

When preparing the data to train a neural network, it is important to scale the features. This data pre-processing step is important because it ensures that all input features contribute equally during the training process. As a result of that, the electricity features were normalised to a range [0,1] using Min-Max normalisation, as defined by Eq. (1):

$$x_{scaled} = \frac{x - x_{min}}{x_{max} + x_{min}} \quad (1)$$

where x_{scaled} represents the normalised electricity data in the range [0,1], x is the original electricity data, x_{min} and x_{max} represent the minimum and maximum value of the electricity feature in the dataset, respectively.

C. Feature Engineering

Fig. 2 presents the monthly averages for normalised PV output across the period 1980-2024, while Fig. 3 show variations in normalised monthly outputs for all 45 years included in the dataset. Fig. 1(b), Fig. 2 and Fig. 3 show that there are clear daily and yearly periodicities of the PV output (i.e., day versus night, and seasonal changes). To capture these patterns, cyclical temporal features (hours of the day, day of the week, and month of the year) were introduced based on the sine and cosine transforms [25].

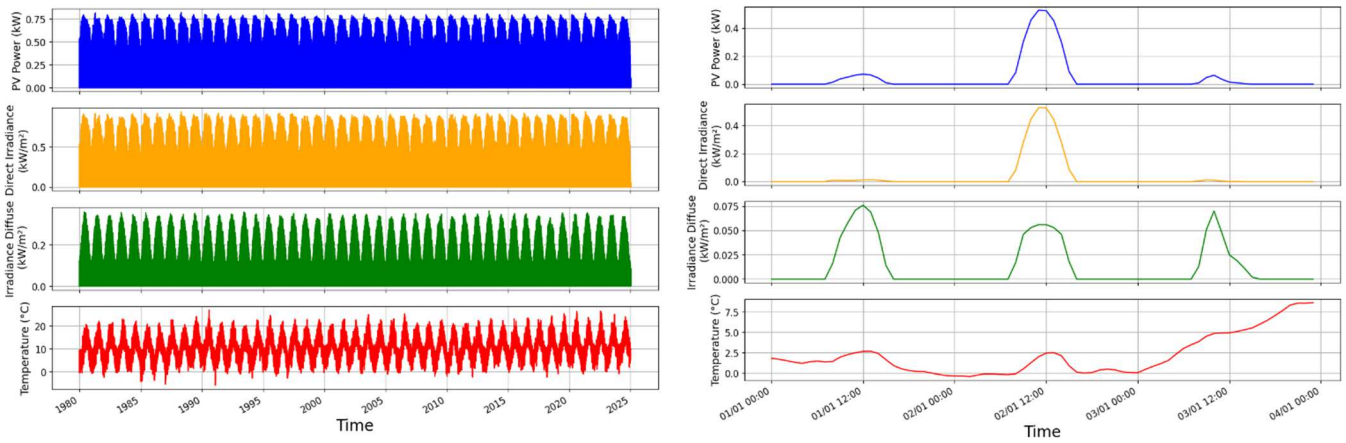


Fig. 1. Combined hourly PV generation dataset for Brighton from January 1980 to December 2024.

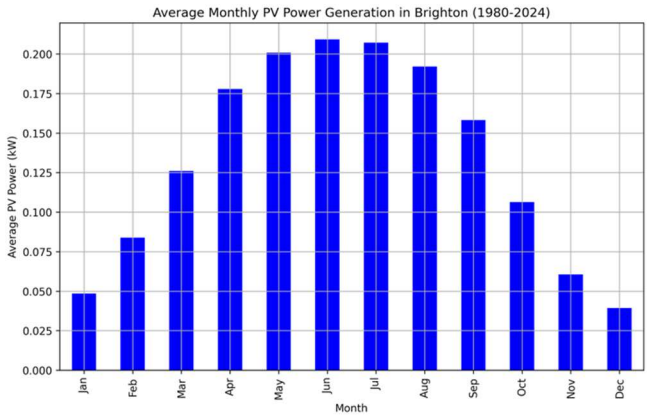


Fig. 2. Monthly Average PV generation in Brighton from January 1980 to December 2024.

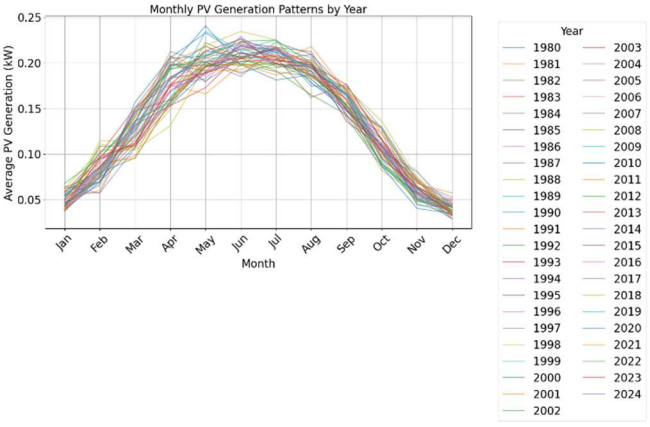


Fig. 3. Monthly PV generation for Brighton for each year from 1980 to 2024.

D. Model Architecture

Previous studies [16,24,25,26] have based their strategy for different TSF problems, such as solar power, photovoltaic power generation (PVPG), and load forecasting using hybrid models. This study follows the similar methodology proposed by [26] to forecast solar power, by using CNNs and LSTM network (CNN+LSTM) to forecast PV power for Brighton.

Convolutional neural networks (CNNs), first introduced for digit classification [6], and subsequently proven effective across different image classification applications [27, 28], excel at detecting local patterns that occur at adjacent input elements (i.e., pixels). Since time series data exhibits local dependencies, where adjacent points in time are often corelated, CNNs can be effectively used to classify one dimensional sequential data. On the other hand, long-short-term memory (LSTM) networks [29] are a type of recurrent neural network (RNN) that have feedback connections which allows them to exploit temporal dependencies across sequential data. Furthermore, LSTMs were specifically designed to mitigate the vanishing/exploding gradient problem in vanilla RNNs during training on sequential data, and that makes them effective for tasks that involve sequential data, such as TSF.

Fig. 4 presents the architecture of the CNN+LSTM model used for forecasting PV power for Brighton. The model accepts ten input features: historical electricity values as the primary predictive variables, engineered cyclical features encoding hourly, weekly, and monthly patterns through sine and cosine transformations, direct irradiance, diffuse irradiance, and temperature. Input sequences are structured as three dimensional tensors with the following dimensions:

$$(\text{batch size}, \text{length}, \text{features}) = (\text{None}, 168, 10) \quad (2)$$

where None denotes variable batch size, 168 represents the sequence length (temporal window), and 10 corresponds to the number of input features.

The causal convolutional layer (Conv1D) first processes these inputs to extract local temporal patterns. The resulting features from the Conv1D layer are then feed into stacked bidirectional LSTM layers. These layers model both forward and backward dependencies within the input window, and capture complex time-varying patterns. The output from stacked bidirectional LSTM layers is then passed through two dense layers, where the final linear output layer produces the forecast value.

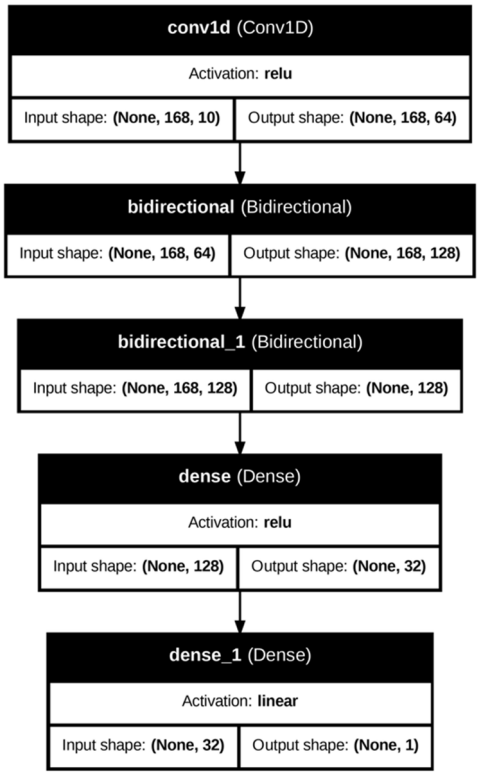


Fig. 4. Architecture of the hybrid model for PV forecasting (CNN+LSTM).

E. Evaluation Metrics

To evaluate the performance of the CNN+LSTM model, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) evaluation metrics were used. MAE was calculated as the average of absolute errors between the ground truth value and the forecasted value:

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|\hat{y}_n - y_n| \quad (3)$$

where $\hat{y}_n$ and y_n are the forecasted and ground truth electricity value for the given timestep, N is the number of data points in the validation set.

RMSE was calculated as the average of the root mean of squared error between the ground truth value and the forecasted value. RMSE was calculated using (4), while MAPE was computed using (5).

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\hat{y}_n - y_n)^2} \quad (4)$$

$$MAPE = \frac{100}{N}\sum_{n=1}^{N}\left|\frac{\hat{y}_n - y_n}{y_n}\right| \quad (5)$$

F. Implementation Details

For all experiments, Adam optimiser with Huber loss was employed. The batch size was set to 64. To determine the optimal learning rate for the neural network, a systematic approach was adopted. An exponentially increasing learning rate schedule was implemented, starting at 10^{-6} and growing by a factor of $10^{0.05}$ per epoch over 30 epochs. The model was trained using this schedule, tracking loss values across learning rates. After training, the loss per epoch was plotted against the learning rate per epoch, and the optimal learning rate was set to 10^{-4}. After identifying the optimal learning rate, the model was trained for 100 epochs with the optimal learning rate and using early stopping regularization technique to ensure robust generalization and avoid overfitting.

G. Experiment Settings

All experiments were conducted on Google Collaboratory (Colab), a cloud-based Jupyter notebook platform, using TensorFlow and Python's scientific computing libraries (e.g., NumPy, Pandas, Scikit-learn). To accelerate the training of neural network, a high-performance accelerator such as NVIDIA A100 GPU was used.

III. Experimental Results

A. Training and Validation Loss

Fig. 5 shows the training and validation loss curves using Huber loss across 100 epochs. Training curve shows rapid decrease for the first 20 epochs. This indicates that for the first 20 epochs, the model is effectively learning dominant temporal features. Then for the rest of epochs both training and validation loss curves decrease in parallel without divergence. This suggests that there is no overfitting. Finally, for the last epochs (e.g., from 60 to 100) the loss curves plateau, which implies model convergence.

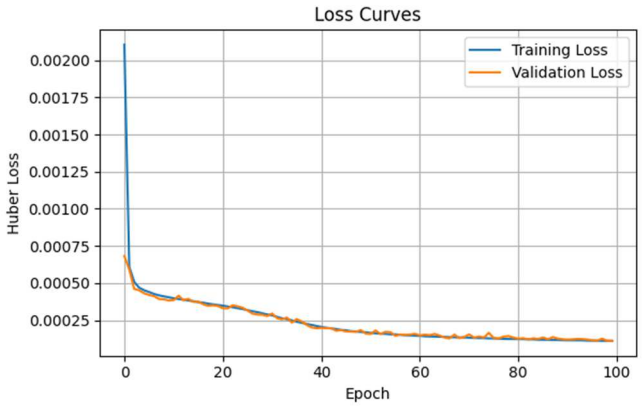


Fig. 5. Loss of training and validation dataset in the training phase.

B. Case Study

To analyse the forecasting performance of the CNN + LSTM hybrid model, a multi-layer perceptron (MLP) network was used as a baseline approach, i.e., benchmark to forecast the PV generation for Brighton. Fig. 6 shows the comparison between the actual PV generation and the PV generation forecasts produced using the MLP network and the CNN + LSTM hybrid model for Brighton for the week of 1-7 December 2024.

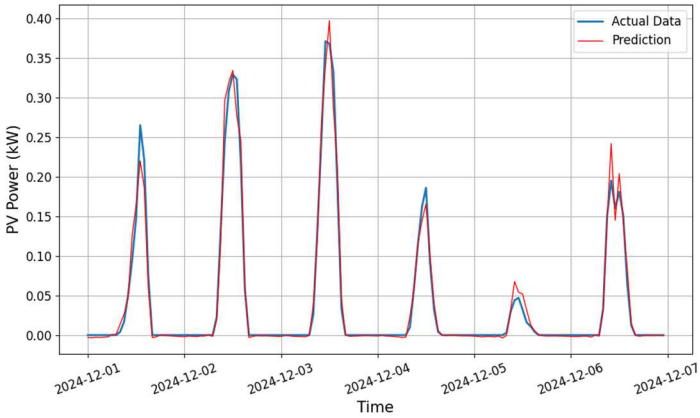
Comparing the results shown in Fig. 6. it can be seen that the CNN+LSTM hybrid model performs better than the MLP in forecasting PV generation. More specifically, the MLP struggles to accurately capture the off-peak hours, as the predictions for this period show irregular fluctuations. However, these irregular fluctuations are not present in the forecasts made by CNN+LSTM. This is because the hybrid model integrates CNNs for feature extraction and LSTM network for sequential modelling and can better capture the underlying patterns in the data. On the other hand, traditional neural networks such as MLP cannot accurately model important sequential dependencies and spatial relationships that are crucial for accurate PV generation forecasting.

Fig. 7 shows the actual PV generation and the PV generation forecasts made by both the CNN + LSTM hybrid model and the MLP network for the months of January and November 2024.

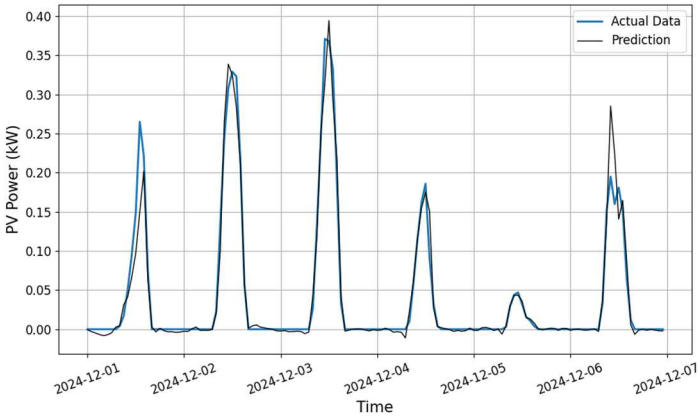
The MAE, RMSE, and MAPE results for the MLP and CNN+LSTM hybrid model are summarised in Table I. The MAE and RMSE of the CNN+LSTM hybrid model on the test set are 0.0061 and 0.0121, respectively, which are significantly lower than the metrics obtained by the MLP. Furthermore, the MAPE of the CNN + LSTM hybrid model (4.86%) is less than the MAPE obtained by the MLP network.

TABLE I. EVALUATION METRICS (MAE, RMSE, AND MAPE) FOR THE MLP AND CNN+ LSTM IN FORECASTING PV GENERATION.

Algorithm	MAE	RMSE	MAPE
MLP	0.0143	0.0264	10.12%
CNN+LSTM	**0.0061**	**0.0121**	**4.86%**

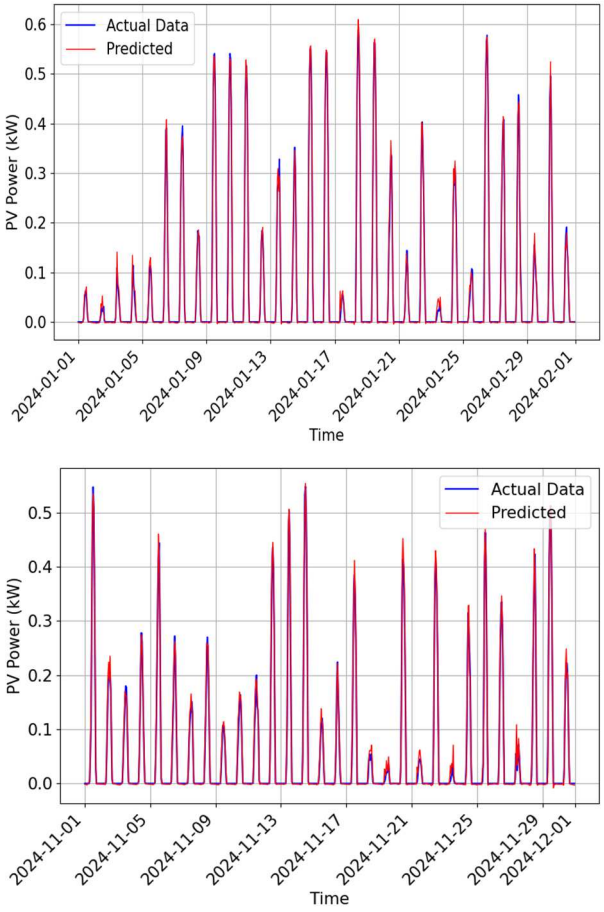


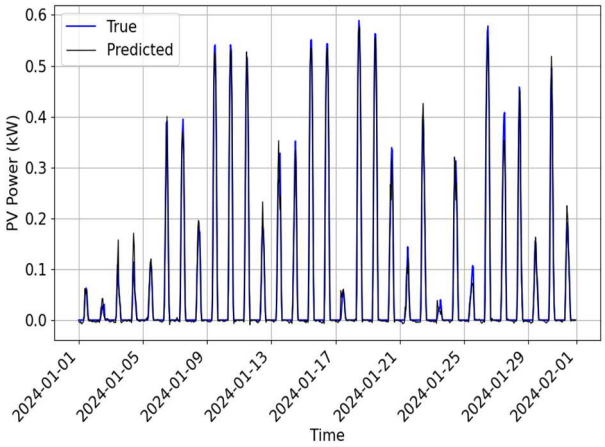
(a) True and forecasted PV generation using CNN + LSTM hybrid model.

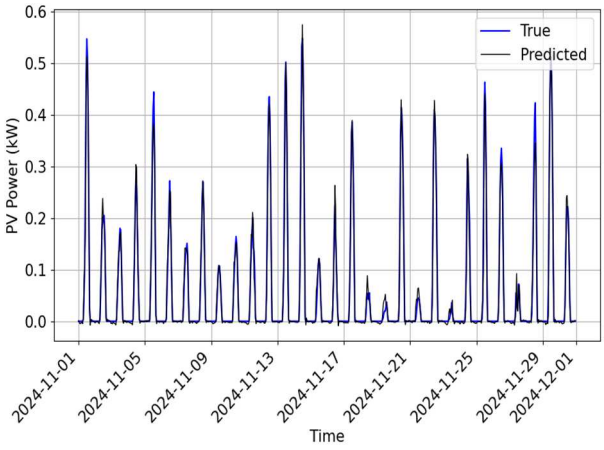


(b) True and forecasted PV generation using MLP network.

Fig. 6. True and forecasted PV generation for Brighton from December 1st to 7th 2024 using CNN+LSTM hybrid model and MLP network.







(a) CNN + LSTM hybrid model.

(b) MLP network.

Fig. 7. True and forecasted PV generation for Brighton for the months of January and November 2024 using CNN+LSTM hybrid model and MLP network.

IV. Conclusion and Future Work

This study has demonstrated an effective application of a hybrid CNN+ LSTM deep learning framework for forecasting PV generation in the UK. The experiments were conducted on the data downloaded from Renewables.ninja website that is based on MERRA-2 reanalysis and SARAH satellite radiation datasets. The experimental results have shown that the predictions made by CNN+LSTM hybrid model are close to the actual PV generation values. Furthermore, by analysing the MAE, RMSE, and MAPE metrics it was confirmed that the CNN+LSTM model has a consistent performance across the test set.

Future research will focus on forecasting wind output and cooling and heating demand across different UK regions with the emphasis on long-term (seasonal to annual) predictions.

References

[1] W. Li and K. L. E. Law, "Deep Learning Models for Time Series Forecasting: A Review," *IEEE Access*, vol. 12, pp. 92306–92327, 2024, doi: https://doi.org/10.1109/access.2024.3422528.

[2] Z. Yue *et al.*, "TS2Vec: Towards Universal Representation of Time Series," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 8, pp. 8980–8987, Jun. 2022, doi: https://doi.org/10.1609/aaai.v36i8.20881.

[3] Y. Li, H. Wang, J. Li, C. Liu, and J. Tan, "ACT: Adversarial Convolutional Transformer for Time Series Forecasting," *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Jul. 2022, doi: https://doi.org/10.1109/ijcnn55064.2022.9892791.

[4] H. Zhou *et al.*, "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, pp. 11106–11115, May 2021, doi: https://doi.org/10.1609/aaai.v35i12.17325.

[5] J. Kim, H. Kim, H. Kim, D. Lee, and S. Yoon, "A Comprehensive Survey of Time Series Forecasting: Architectural Diversity and Open Challenges," *arXiv.org*, May 01, 2025. https://arxiv.org/abs/2411.05793.

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: https://doi.org/10.1109/5.726791.

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Jun. 2009, doi: https://doi.org/10.1109/cvpr.2009.5206848.

[8] P. Rajpurkar, J. Zhang, K. Lopyrev, and P. Liang, "SQuAD: 100,000+ Questions for Machine Comprehension of Text," *arXiv (Cornell University)*, Jun. 2016, doi: https://doi.org/10.48550/arxiv.1606.05250.

[9] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, and S. R. Bowman, "GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding," *arXiv (Cornell University)*, Apr. 2018, doi: https://doi.org/10.48550/arxiv.1804.07461.

[10] J.-F. Chen, W.-M. Wang, and C.-M. Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Systems Research*, vol. 34, no. 3, pp. 187–196, Sep. 1995, doi: https://doi.org/10.1016/0378-7796(95)00977-1.

[11] W. Lulu, K. Zhou, S. Yang, and X. Lu, "Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting," *Energy*, vol. 171, pp. 1053–1065, Mar. 2019, doi: https://doi.org/10.1016/j.energy.2019.01.075.

[12] "Dataport | Login | Signup," *Pecanstreet.org*, 2024. https://dataport.pecanstreet.org (accessed May 25, 2025).

[13] "Home Page | DKA Solar Centre," *Dkasolarcentre.com.au*, 2023. https://dkasolarcentre.com.au (accessed May 25, 2025).

[14] K. Yan, W. Li, Z. Ji, M. Qi, and Y. Du, "A Hybrid LSTM Neural Network for Energy Consumption Forecasting of Individual Households," *IEEE Access*, vol. 7, pp. 157633–157642, 2019, doi: https://doi.org/10.1109/access.2019.2949065.

[15] S. Bai, K. J. Zico, and V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling," *arXiv.org*, 2018. http://arxiv.org/abs/1803.01271 (accessed May 26, 2025).

[16] H. Zheng, J. Yuan, and L. Chen, "Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation," *Energies*, vol. 10, no. 8, p. 1168, Aug. 2017, doi: https://doi.org/10.3390/en10081168.

[17] P.K. Baskar and K. Hariharan, "Multivariate Time Series Traffic Forecast with Long Short Term Memory based Deep Learning Model," *2020 International Conference on Power, Instrumentation, Control and Computing (PICC)*, Dec. 2020, doi: https://doi.org/10.1109/picc51425.2020.9362455.

[18] N. Jin, Y. Zeng, K. Yan, and Z. Ji, "Multivariate Air Quality Forecasting With Nested Long Short Term Memory Neural Network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8514–8522, Dec. 2021, doi: https://doi.org/10.1109/tii.2021.3065425.

[19] R. Gelaro *et al.*, "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)," *Journal of Climate*, vol. 30, no. 14, pp. 5419–5454, Jul. 2017, doi: https://doi.org/10.1175/jcli-d-16-0758.1.

[20] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, Nov. 2016, doi: https://doi.org/10.1016/j.energy.2016.08.060.

[21] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, Nov. 2016, doi: https://doi.org/10.1016/j.energy.2016.08.068.

[22] "Renewables.ninja," 2019. https://www.renewables.ninja/ (accessed May 26, 2025).

[23] R. Müller, U. Pfeifroth, C. Träger-Chatterjee, J. Trentmann, and R. Cremer, "Digging the METEOSAT Treasure—3 Decades of Solar Surface Radiation," *Remote Sensing*, vol. 7, no. 6, pp. 8067–8101, Jun. 2015, doi: https://doi.org/10.3390/rs70608067.

[24] X. Luo, D. Zhang, and X. Zhu, "Deep learning based forecasting of photovoltaic power generation by incorporating domain knowledge," *Energy*, vol. 225, p. 120240, Jun. 2021, doi: https://doi.org/10.1016/j.energy.2021.120240.

[25] P. Singla, M. Duhan, and S. Saroha, "A comprehensive review and analysis of solar forecasting techniques," *Frontiers in Energy*, Mar. 2021, doi: https://doi.org/10.1007/s11708-021-0722-7.

[26] W. Lee, K. Kim, J. Park, J. Kim, and Y. Kim, "Forecasting Solar Power Using Long-Short Term Memory and Convolutional Neural Networks," *IEEE Access*, vol. 6, pp. 73068–73080, 2018, doi: https://doi.org/10.1109/access.2018.2883330.

[27] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 2014, doi: https://doi.org/10.48550/arXiv.1409.1556.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016, doi: https://doi.org/10.1109/cvpr.2016.91.

[29] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: https://doi.org/10.1162/neco.1997.9.8.1735.