Forecasting hybrid renewable power generation in Luxembourg: a comparative study of convolutional neural network's application

Vahid Arabzadeh
Interdisciplinary Centre for Security, Reliability and Trust (SnT)
University of Luxembourg
Luxembourg, Luxembourg
vahid.arabzadeh@uni.lu

Raphaël Frank
Interdisciplinary Centre for Security, Reliability and Trust (SnT)
University of Luxembourg
Luxembourg, Luxembourg
raphael.frank@uni.lu

***Abstract*— Access to reliable renewable power generation forecasting tools is crucial for optimizing grid operations and advancing the integration of renewable energy, which in turn leads to the sustainability of energy systems. This study develops a forecasting model utilizing Convolutional Neural Networks (CNN) for precise prediction of hybrid solar and wind power generation in Luxembourg. Through a comprehensive comparative analysis exploring various combinations of critical hyperparameters, we demonstrate the significant capability of the CNN approach to serve as an effective prediction tool for hybrid renewable energy production. Our findings underscore the reliable potential of CNNs to enhance the accuracy of renewable energy forecasts, thereby enabling a more seamless and efficient integration of renewable energy sources into Luxembourg's power grid. Our CNN model demonstrated exceptional performance, with an R-squared (R^2) exceeding 90%, particularly for forecasting horizons of 5, and 24 hours ahead.**

Keywords— Hybrid renewable power; Convolutional Neural Network; Multi-hour ahead forecasting.

I. Introduction

During the rapid evolution of society and economy, driven by increasing energy demands and concerns over fossil fuels' environmental impact, many nations are transitioning towards renewable energy sources such as solar and wind power [1]. Luxembourg stands at the forefront of this movement, setting ambitious goals for 2021-2030 aimed at reducing pollution and boosting renewable energy usage [2, 3]. With its commitment to sustainability, Luxembourg is implementing policies and initiatives to accelerate the adoption of solar and wind energy technologies, leading the way towards a cleaner, greener future [3]. Therefore, a reliable forecasting is crucial for seamlessly integrating renewables into existing energy systems, yet accurate predictions face challenges from intricate weather variables [4]. Deep learning, particularly Convolutional Neural Networks (CNNs), has emerged as a crucial tool in overcoming these challenges mainly in combination with other techniques and algorithms [4]. In wind power prediction using an optimized support vector regression (SVR) model, an improved Jellyfish Search (IJS) algorithm optimizes SVR parameters, boosting predictive accuracy [5]. Hourly stepwise forecasting for solar irradiance proposes a hybrid model integrating CNN, long short-term memory networks (LSTM), multi-layer perceptron (MLP), and variational mode decomposition (VMD) for precise hourly solar irradiance forecasting, surpassing traditional methods [6]. Short-term wind power forecasting with LSTM and attention mechanism (AM-LSTM) model dynamically weighs physical attribute data using an attention mechanism with CNN and LSTM networks, improving short-term wind forecasting accuracy [7]. Very short-term forecasting of wind power generation using a hybrid deep learning model employs convolutional and gated recurrent unit (GRU) layers for superior very short-term wind power generation forecasting accuracy [8]. Wind power forecasting with the optimized deep learning techniques synergizes WPD with CNN and LSTM networks for offshore wind power forecasting, achieving superior accuracy [9]. The GWO-Nested CEEMDAN-CNN-BiLSTM model for wind speed forecasting integrates grey wolf optimization (GWO), complete ensemble empirical node decomposition with adaptive noise (CEEMDAN), CNN, and Bidirectional LSTM, achieving superior wind speed forecasting accuracy [10]. Short-term wind power forecasting based on attention mechanism and CNN-LSTM networks dynamically weighs input data importance using an attention mechanism with CNN and LSTM networks, improving short-term wind power forecasting accuracy [11]. CNN with LSTM units, enhanced by the Coati Optimization Algorithm (COA), to improve the accuracy of hybrid renewable energy forecasting [12]. Building upon the advancements highlighted in recent literature on renewable energy forecasting, our study aims to make a unique contribution by focusing solely on the power of CNN for renewable energy forecasting. This investigation is driven by the following hypothesis:

- A well-tuned CNN, free from the complexities of additional algorithms or techniques, can offer a solution with low computational demand while still delivering high-quality forecasting results.

This approach offers several advantages, including decreased complexity compared to hybrid models, resulting in simpler implementation and interpretation of outcomes. Additionally, CNN models typically exhibit faster computation times, making them suitable for real-time forecasting applications. Moreover, the architecture of CNN allows for scalability, enabling the incorporation of additional data sources or features without significant overhead. By leveraging these benefits, we aim to demonstrate the efficiency of CNN in renewable energy forecasting, paving the way for streamlined and effective integration of sustainable energy sources into power systems. To achieve this, we focus on three main novelties:

- Investigating how adjusting CNN hyperparameters impacts the accuracy of forecasting hybrid renewable energy production in our studied case, Luxembourg.
- Introducing a streamlined CNN model optimized with the best hyperparameter settings, aimed at delivering accurate energy forecasts at 5, and 24-hour intervals.
- Enhancing forecasting models by integrating a variety of datasets, including historical energy output across different years, to empower accuracy and demonstrate consistent reliability.

Despite the array of sophisticated methods detailed in recent literature for renewable energy forecasting-ranging from hybrid deep learning models to advanced data preprocessing techniques and optimization algorithms-the current study seeks to evaluate the standalone performance of a CNN in forecasting hybrid renewable power generation. Indeed, our study uses a simplified deep learning framework to predict major contributions to hybrid renewable energy forecasting. In addition to contributing to the advancement of scientific discourse, this study offers a reliable tool for practical energy strategies and applications. The paper is organized methodically: Section II provides an overview of methodology, Section III presents the findings, and Section IV concludes.

II. METHOD

A. General overview

Our methodology aims to develop a CNN model for forecasting hybrid wind and solar power generation data by training and evaluating the CNN model using different combination of major hyperparameters. We analyze the results to determine the optimal model's hyperparameters configuration and explore the effect of individual hyperparameters.

B. Introduction of applied data

Here, we investigate into the dataset that underpins our methodology. This data originates from the ENTSO-E Transparency Platform, an extensive European resource providing real-time and transparent data on electricity transmission and market operations throughout the continent [13]. To evaluate the predictive performance of our method, we examine solar and onshore wind power generation data from Luxembourg spanning the years 2015 to 2021. Over these years, there's a noticeable trend of growth across all metrics. Starting in 2015, the mean power generation was recorded at 18.45 MW, with the maximum power output reaching 97 MW, and a standard deviation of 16.15 MW. This trend of increasing power generation continued steadily, with the mean output rising to 23.13 MW in 2016, and then more significantly to 65.90 MW by 2021. The maximum hybrid power output also showed a remarkable increase from 97 MW in 2015 to 243 MW in 2021, alongside a steady rise in the standard deviation from 16.15 MW to 43.73 MW over the same period. This data demonstrates a clear upward trajectory in Luxembourg's hybrid renewable power generation capabilities, highlighting improvements in efficiency and capacity. Our dataset comprises 8,760 samples annually. We plan to develop the model using data from two years (2015-2016). Specifically, 2015 data will be split for training (80%) and validation (20%). The entire 2016 dataset, unseen by the model during training, will be used for testing. The model's testing performance will serve as the benchmark for selecting the optimal model from the simulation pool.

C. Convolutional Neural Network

The applied CNN architecture (see Fig 1) is designed for sequential input data, focusing on hybrid renewable power generation sequences. It starts with a Sequence Input layer and follows with multiple 1-D convolutional blocks (from 1 to N in Fig 1), each containing a convolution layer, batch normalization, and rectified linear unit (ReLU) activation to extract features and introduce non-linearity [12]. The convolutional layers use filters to capture spatial and temporal patterns, enhancing the model's ability to learn complex data relationships [14]. Batch normalization stabilizes training by normalizing layer activations, while ReLU activation promotes sparsity and efficient learning by setting negative inputs to zero [14]. Following the convolutional stages, the network includes fully connected layers for feature integration, with dropout regularization to prevent overfitting [15]. The architecture culminates with fully connected layers, the nodes of which correspond to the anticipated scale of forecasted values, followed by a regression output layer designed for predicting continuous variables. It employs a mean-squared error loss function to quantify the accuracy of predictions. This streamlined design enables the model to learn hierarchical time series representations for accurate forecasting in hybrid power generation contexts.

D. Applied method

The developed MATLAB code presents an approach to forecasting hybrid power generation using CNNs. The code aims to train a CNN model that can accurately predict future hybrid power generation based on historical data. We applied a method to prevent overfitting called early stopping which is implemented using a custom callback function in our training process. This function monitors the validation loss during training and halts the training process if the validation loss fails to improve for a specified number of epochs. To prevent overfitting other applied techniques are also employed. Firstly, dropout layers are integrated into the network architecture, randomly deactivating neurons during training to enhance the network's resilience and reduce its reliance on specific features. Additionally, batch normalization layers are utilized to normalize layer activations, stabilizing the training process and acting as a form of regularization. We simulate a total of 2496 cases, incorporating various combinations of hyperparameters for two distinct forecasting horizons (5, and 24 hours ahead).

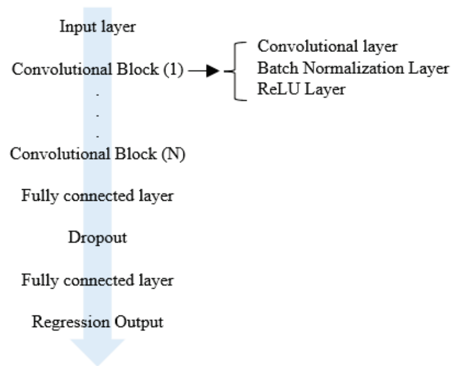

Fig 1: Applied CNN architecture in this study. The model features filter sizes of 3, 5, 7, and 9, with counts from 16 to 96, 1-6 convolutional blocks, dropout rates of 30%-70%.

Selecting 5-hour and 24-hour forecasting horizons for model performance evaluation strategically addresses both immediate and daily energy management needs [16]. The 5-hour horizon is essential for intra-day operations, including energy trading and adjusting to demand-supply changes, enabling efficient resource dispatching. Conversely, the 24-hour forecast is vital for daily planning, such as maintenance, energy procurement, and reserve management, ensuring grid stability and economic operations [17]. Together, these horizons offer a comprehensive assessment of the model's utility in supporting renewable energy integration for both short-term adjustments and strategic decision-making. The selection of hyperparameters is based on prior studies that have demonstrated their effectiveness in enhancing model performance [6, 7, 12, 15, 18-20]. The hyperparameters in this study are crucial for configuring the CNN model used in forecasting. “Window size” defines the temporal extent of the input data, set at intervals of 1h, 8h, 24h, and 48h in this study, to accommodate various periods of analysis. “Filter size” refers to the dimensions of the kernels employed in the convolutional layers for feature extraction, with sizes of 3, 5, 7, and 9 explored. “Filter number” indicates the quantity of distinct filters in each convolutional layer, enabling the detection of a diverse set of features, with values set at 16, 32, 48, 64, and 96 in our study. “Dropout” is implemented as a regularization strategy, randomly omitting a proportion of the input units during training to mitigate overfitting, with rates of 0.1, 0.3, 0.5, and 0.7. Lastly, the “CNN layer number” describes the overall depth of the CNN architecture, critical for the model’s ability to capture complex patterns, with the depth varying from 1 to 6 layers. The MATLAB code developed for this study is available upon request.

III. RESULTS

Our results originate from the domain of time series forecasting, which specifically evaluates the performance of various configurations of a CNN model across different forecasting horizons, which varies among 5, and 24 hours ahead predictions. The dataset records the impact of different hyperparameters such as window size, filter size, number of filters, dropout rate, and the number of CNN layer on the model's forecasting accuracy.

The objective of our analysis is to dissect how these configurations influence the model's effectiveness, gauged through metrics like Mean Squared Error (MSE), R-squared (R^2). By understanding these dynamics, we aim to uncover insights that could guide the optimization of time series forecasting models for improved accuracy and reliability. The MSE offers insight into the average magnitude of the model's errors, providing a clear measure of predictive accuracy, while R^2 offers an indication of the model's explanatory power.

The dataset created from simulated cases, structured as a table (with 2496 rows), facilitates a thorough analysis of how different model configurations affect time series forecasting accuracy. Each column in this dataset represents a critical aspect of the model's behavior, documenting the configuration, hyperparameters, and performance metrics for each simulation. These performance metrics (MSE and R^2) quantitatively assess the model's accuracy and effectiveness in identifying the underlying patterns of the forecasted time series data.

From looking deeply into the dataset, we aim to extract actionable insights, such as identifying the optimal range of model parameters for balancing model complexity with forecasting precision, thereby laying a foundation for more informed and targeted model tuning and validation efforts. In general, the MSE demonstrates a considerable range, with a minimal value of 0.361 MW, suggesting instances of near-perfect forecasting accuracy, and extending to a maximum of 440.55 MW, where predictions substantially deviate from actual observations. The R^2 values, indicating the proportion of variance in the dependent variable predictable from the model, vary dramatically from low values, almost devoid of explanatory power in some cases, to an exemplary high of 0.999, denoting nearly flawless predictability. These indicators collectively furnish a nuanced picture of model efficacy, ranging from highly precise forecasts to scenarios marked by notable prediction inaccuracies, highlighting the pivotal role of model configuration in optimizing forecasting performance. This variation underscores the critical importance of model configuration, as different settings can dramatically influence the model's ability to capture and predict the underlying patterns in the data.

Examining the scatter plots for the forecasting horizons of 5, and 24 hours reveals nuanced insights into the impact of various individual hyperparameters on the MSE (See Fig 2). For forecasting horizon = 5, the MSE remains relatively unaffected by different window sizes, hinting that short-term forecasts may not be sensitive to the amount of input data. Smaller filter sizes tend to result in lower MSE, suggesting they might be optimal for capturing the necessary features at this forecast range. The number of filters shows that there may be an optimal quantity that minimizes MSE, beyond which the performance does not improve significantly. Dropout rates do not present a clear trend, indicating that an optimal value likely exists that must be fine-tuned to balance model complexity and prevent overfitting. As for CNN layers, an initial decrease in MSE is observed with an increase in layers, after which the benefit plateaus, suggesting that a moderate depth is most beneficial for short-term forecasting.

At the 24-hour horizon, a greater spread in MSE values suggests that a larger window size doesn't necessarily correlate with improved long-term forecasts, and smaller filter sizes might again be more effective. The pattern for the number of filters does not show a consistent decrease in MSE, pointing to an optimal count before complexity leads to diminishing returns. The dropout rate's impact on MSE implies that there might be an optimal rate that mitigates overfitting while still allowing sufficient learning. The MSE for varying numbers of CNN layers is relatively high overall, which could imply that long-term forecasts may benefit less from deeper architectures or that limitations in the data prevent the model from effectively leveraging increased depth.

In our analysis, we tailored the hyperparameters for each forecasting horizon to pinpoint the most efficient combination for predictive accuracy. We evaluate the model's performance during the testing phase by comparing outcomes across all simulated scenarios. Our objective is to identify the configuration that yields the lowest MSE while maintaining the simplest model structure. This involves selecting the model with the fewest convolutional layers, smallest window size, minimal filter size and count, and lowest dropout rate, thereby ensuring optimal performance with minimal complexity.

For a short-term, 5-hour horizon, the best results came from a simple yet effective configuration: a window size of 1, filter size of 7, 48 filters, dropout of 0.1, and 2 CNN layers, achieving an the lowest MSE. This reinforces the idea that less can be more, with a minimalistic model yielding high accuracy for short-term forecasts. For the long-term, 24-hour forecasts, we observed a different dynamic: the lowest MSE was achieved with a window size of 24, filter size of 1, 96 filters, a dropout rate of 0.3, and 6 CNN layers. Despite the larger window and increased dropout hinting at a rising MSE trend, this combination suggests that capturing more extended patterns is key, but a moderate approach to complexity is essential.

The optimized models were deployed to forecast data for the years 2019 to 2021, demonstrating consistent performance across this period (see Fig 3). Initially, it achieved a MSE of 4.443 MW and maintained a high R^2 value above 0.990 for forecasts made 5 hours ahead. Although the MSE incrementally rose to 17.839 MW by 2021, the model's ability to accurately reflect key data features underlines its robustness in adapting to annual variations. The steady increase in MSE might suggest changing data dynamics or a growing complexity in the underlying patterns. Nonetheless, the consistently high R^2 values across years highlight the model's robustness and its capability to capture essential trends with a relatively simple setup. The 24-hour forecast model with its optimized configuration demonstrates an initial strong performance with an MSE of 53.670 MW in 2019 and an R^2 of 0.914, suggesting effective capture of long-range patterns with a broad window size and substantial filter depth. Over time, the MSE increases, peaking at 211.902MW in 2021 with an R^2 of 0.915, which may reflect the model's struggle against increasingly complex or volatile data. Yet, the model maintains relatively almost constant and high R^2 values throughout the example years, affirming its capability to account for a significant portion of the variance within the data.

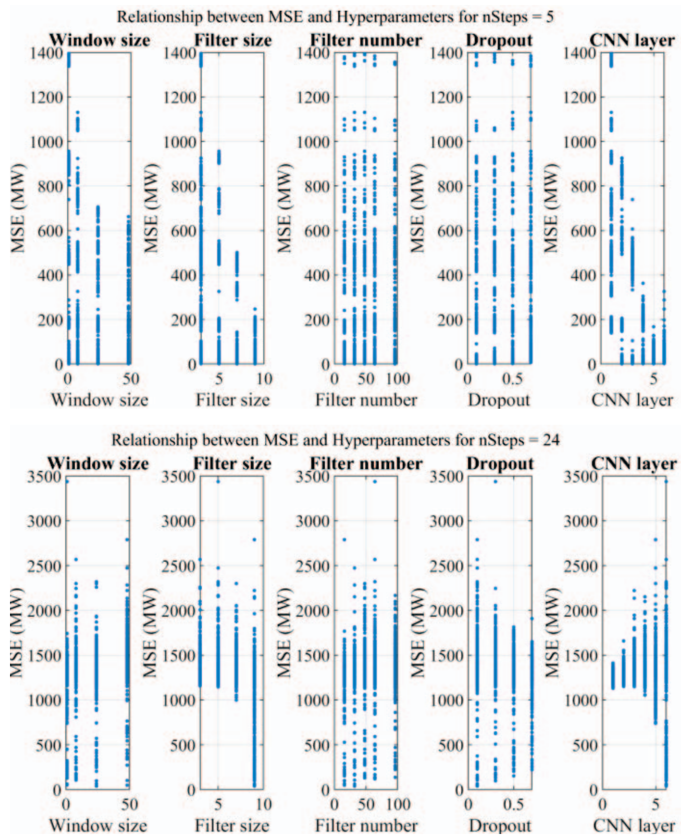

Fig 2*: Comparative analysis of CNN's hyperparameters on forecasting accuracy across 5, and 24 hours ahead*

IV. Conclusion

In this study, we explored the potential of CNNs and their hyperparameters in shaping forecasting outcomes for hybrid renewable power generation in Luxembourg. Our analysis unearthed the significant influence of hyperparameters on forecasting accuracy, revealing those variations in factors like window size, filter size, number of filters, dropout rate, and CNN layers distinctly impacted model performance. Notably, smaller filter sizes and moderate numbers of filters were found to yield lower MSE, indicating enhanced accuracy, particularly for short-term forecasts. Achieving a delicate balance between model complexity and regularization techniques, such as dropout rates, emerged as crucial for optimizing accuracy across different forecasting horizons. Tailoring model configurations to specific forecast horizons, we identified optimal hyperparameter combinations for the applied dataset. For short-term forecasts (5 hours ahead), a simplified model with a window size of 1, small filter size, moderate number of filters, low dropout rate, and 2 CNN layers showcased impressive accuracy. As the forecast horizon lengthened, slightly more complex configurations were required, underscoring the need to balance capturing longer patterns with avoiding overfitting. Therefore, we found for longer-term forecasts (24 hours ahead), the best configuration comprised a larger window size of 24, a filter size of 1, 96 filters, a dropout rate of 0.3, and 6 CNN layers. Our study underscores the pivotal role of the CNN layer and filter size in enhancing the performance of CNNs. Additionally, we demonstrate that a finely tuned CNN offers a cost-effective solution for forecasting hybrid renewable power generation in Luxembourg. Looking ahead, our research points towards promising avenues for further exploration, such as refining hyperparameters through optimization techniques, investigating ensemble methods for data preprocessing, and advancing techniques for dynamic model adaptation.

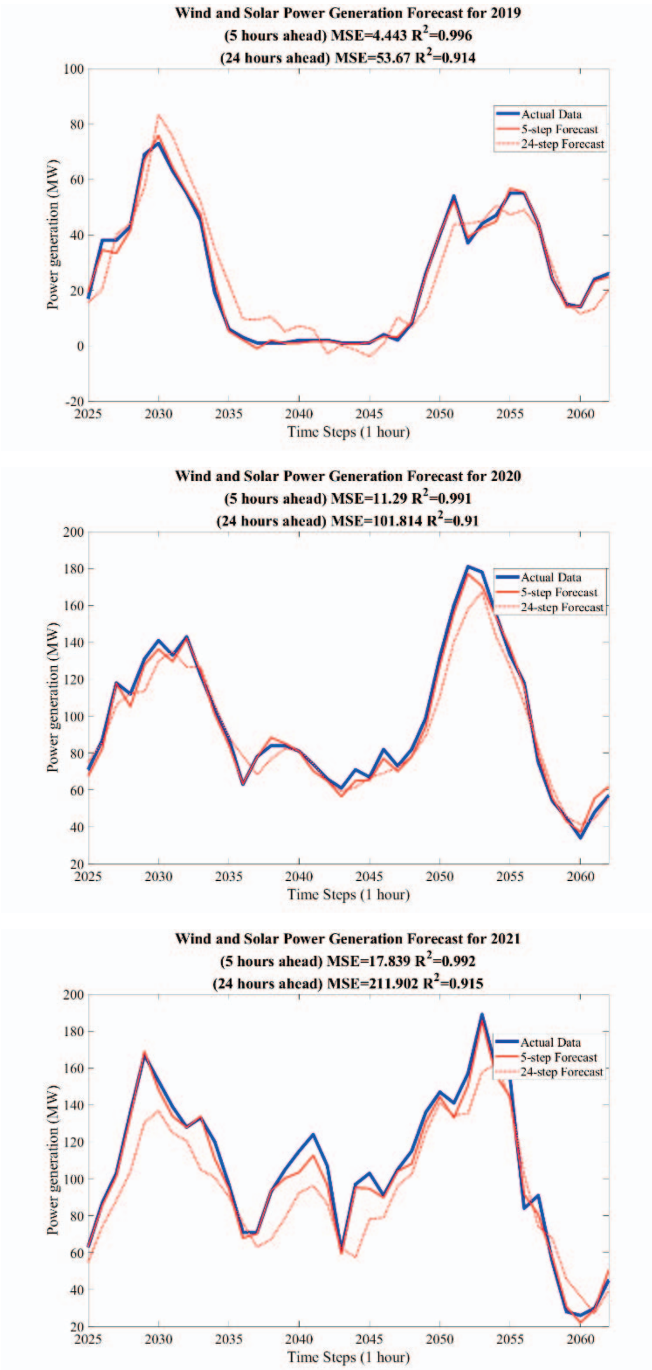

Fig 3: Performance of the optimized models for each forecasting horizon 5 hours, and 24 hours. The model has been tested for years 2019-2021 as example.

REFERENCES

[1] Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla,, W.M.-O. A. Pirani, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock,, and a.T.W.e. M. Tignor, An IPCC Special Report on the impacts of global warming of 1.5°C. 2019.

[2] Frank, R., Review and Growth Prospects of Renewable Energy in Luxembourg: Towards a Carbon-Neutral Future. 2023.

[3] luxembourgeois, G. Luxembourg's integrated national energy and climate plan for the period 2021-2030 (PNEC). 2023 [cited 2023 22/11/2023].

[4] Zendehboudi, A., M.A. Baseer, and R. Saidur, Application of support vector machine models for forecasting solar and wind energy resources: A review. Journal of Cleaner Production, 2018. 199: p. 272-285.

[5] Yuan, D.-D., et al. Wind Power Prediction Method: Support Vector Regression Optimized by Improved Jellyfish Search Algorithm. Energies, 2022. 15, DOI: 10.3390/en15176404.

[6] Liu, J., et al., Hourly stepwise forecasting for solar irradiance using integrated hybrid models CNN-LSTM-MLP combined with error correction and VMD. Energy Conversion and Management, 2023. 280.

[7] Wan, A., et al., Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism. Energy, 2023. 282.

[8] Rai, A., A. Shrivastava, and K.C. Jana, A robust auto encoder-gated recurrent unit (AE-GRU) based deep learning approach for short term solar power forecasting. Optik, 2022. 252: p. 168515.

[9] Hanifi, S., et al., Offshore wind power forecasting based on WPD and optimised deep learning methods. Renewable Energy, 2023. 218: p. 119241.

[10] Cui, X., X. Yu, and D. Niu, The ultra-short-term wind power point-interval forecasting model based on improved variational mode decomposition and bidirectional gated recurrent unit improved by improved sparrow search algorithm and attention mechanism. Energy, 2024. 288: p. 129714.

[11] Xiong, B., et al., Short-term wind power forecasting based on Attention Mechanism and Deep Learning. Electric Power Systems Research, 2022. 206: p. 107776.

[12] Abou Houran, M., et al., COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications. Applied Energy, 2023. 349.

[13] Committee, E.-E.M. ENTSO-E's Transparency Platform. 2024 12/1/2024.

[14] Shahid, F., et al., 1D Convolutional LSTM-based wind power prediction integrated with PkNN data imputation technique. Journal of King Saud University - Computer and Information Sciences, 2023. 35(10).

[15] Agga, A., et al., CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electric Power Systems Research, 2022. 208.

[16] Perera, M., et al., Day-ahead regional solar power forecasting with

[17] hierarchical temporal convolutional neural networks using historical power generation and weather data. Applied Energy, 2024. 361: p. 122971.

[18] Gupta, P. and R. Singh, Forecasting hourly day-ahead solar photovoltaic power generation by assembling a new adaptive multivariate data analysis with a long short-term memory network. Sustainable Energy, Grids and Networks, 2023. 35: p. 101133.

[19] Garg, S. and R. Krishnamurthi, A CNN encoder decoder LSTM model for sustainable wind power predictive analytics. Sustainable Computing: Informatics and Systems, 2023. 38.

[20] Alzubaidi, L., et al., Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 2021. 8(1): p. 53.

[21] Ghimire, S., et al., Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction. Measurement, 2022. 202.