energies

MDPI

Article

Combined MIMO Deep Learning Method for ACOPF with High Wind Power Integration

Altan Unlu [1,*] and Malaquias Peña [2]

[1] Department of Electrical & Computer Engineering, University of Connecticut, Storrs, CT 06268, USA
[2] Department of Civil & Environmental Engineering, University of Connecticut, Storrs, CT 06268, USA; mpena@uconn.edu
* Correspondence: altan.unlu@uconn.edu

Abstract: The higher penetration of renewable energy sources in current and future power grids requires effective optimization models to solve economic dispatch (ED) and optimal power flow (OPF) problems. Data-driven optimization models have shown promising results compared to classical algorithms because they can address complex and computationally demanding problems and obtain the most cost-effective solution for dispatching generators. This study compares the forecast performance of selected data-driven models using the modified IEEE 39 benchmark system with high penetration of wind power generation. The active and reactive power load data of each bus are generated using Monte Carlo simulations, and synthetic wind power data are generated by utilizing a physical wind turbine model and wind speed samples withdrawn from a Weibull distribution. The objective is to design and evaluate an enhanced deep learning approach for the nonlinear, nonconvex alternating current optimal power flow (ACOPF) problem. The study attempts to establish relationships between loads, generators, and bus outcomes, utilizing a multiple-input, multiple-output (MIMO) workflow. Specifically, the study compares the forecast error reduction of convolutional neural networks (CNNs), deep feed-forward neural networks (DFFNNs), combined/hybrid CNN-DFFNN models, and the transfer learning (TL) approach. The results indicate that the proposed combined model outperforms the CNN, hybrid CNN-DFFNN, and TL models by a small margin and the DFFNN by a large margin.

Keywords: data-driven optimal power flow (OPF); renewable energy integration; deep learning; combined/hybrid deep learning; convolution neural network (CNN); deep feed-forward neural network (DFFNN); transfer learning (TL); multiple-input, multiple-output workflow (MIMO)

check for updates

Citation: Unlu, A.; Peña, M. Combined MIMO Deep Learning Method for ACOPF with High Wind Power Integration. *Energies* **2024**, *17*, 796. https://doi.org/10.3390/en17040796

Received: 22 December 2023
Revised: 13 January 2024
Accepted: 25 January 2024
Published: 7 February 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In a modern power system, utilities and transmission systems operators face large-scale uncertainty due to the high penetration and intermittent nature of renewable energy systems, such as solar and wind energy resources, and their integration into the electricity grid. The available generation capacity from conventional energy resources, such as coal-fired, oil-fired, gas-fired, nuclear power plants, and renewable resources, should be more significant than the peak load and the required reserve margin [1]. An optimization model is needed for economic dispatch (ED) and optimal power flow (OPF) problems to reduce fuel costs and obtain the most economical solution for dispatching generators. The objective function of the ED problem is to minimize total generation costs. To obtain a better ED solution, consider an accurate power network. This network model should represent a set of power flow equations mathematically. Applying load flow equations as a whole network model while considering ED problem constraints will enable an OPF model to be found [1,2].

Recently, utility companies and systems operators have collected more valuable electricity data from distributed renewable energy (RE) resources, generators, and loads. Histor-

ical power system operation data give insights into the electricity grid's current conditions and can be used to make modern power systems more stable, efficient, reliable, and secure.

Generally, the OPF problem can be classified into the direct current optimal power flow (DCOPF) and alternating current optimal power flow (ACOPF) subtypes. Indeed, the security constraint optimal power flow (SCOPF) problem is an expansion of the standard ACOPF/DCOPF problem, where binary variables are included to incorporate the switching of transmission lines, which enables more efficient and secure solutions [3]. Due to the ACOPF problem being computationally inefficient, using the Jacobian gradient for Newton–Raphson or Hessian matrices for the interior point method requires significant time to solve the problem. The system's bus lines' flows, voltages, and angles, which can represent grid state variables, are unknown in the OPF problem. Researchers have applied convex relaxation techniques to linearize and simplify the non-linear ACOPF problem to the DCOPF problem [4–7]. However, the solution to the DCOPF problem made a few assumptions to converge faster to the solution; for example, it was assumed that, for each bus, the voltage magnitude is one per unit, and the phase angle difference between buses is trivial. Therefore, the solution to the DCOPF problem did not capture all the current state conditions of the grid. On the other hand, the ACOPF is a more accurate representation of power system operations, although it is demanding and complex. Recently, numerous data-driven OPF solution techniques have been applied to solve this optimization problem more efficiently and effectively. Some techniques have been used to search machine learning (ML) and deep learning (DL) in OPF problems, namely the direct mapping of OPF variables, predicting active or chance constraints, using physics-informed neural network (PINN) models, and predicting warm start points.

The direct mapping technique aims to match the system's state variables to its control variables without repeatedly solving optimization problems [8–11]. Reference [10] utilized a neural network (NN) to map system loads to optimal outputs without actually solving the OPF to enable efficient solutions to be found by avoiding complex approximations, distributed techniques, or computationally intensive platforms, thus ensuring the ability to find feasible solutions through a rapid iterative process. The authors of [12] explored a two-step approach to training a deep neural network (DNN) for security-constrained DCOPF, initially focusing on load and generation mapping, followed by phase angle reconstruction, effectively streamlining the process and reducing the DNN's size and training requirements. The authors of [11] utilized a DNN approach based on a preventive framework, calibrating generation and transmission line limits in DNN training to anticipate approximation errors and ensure the feasibility of the predicted solutions. Generally, in the direct mapping method, the input parameters of the OPF problem are generated from simulations or collected from historical data. Then, these input parameters are introduced to the OPF problem, and the solution is captured at each sample point. Lastly, the model is trained on data using input and output pairs, learning the mapping from input parameters to the optimal OPF solution. Then, a new batch of data feeds into the previously trained model. This approach also allows for sequential learning, where the model can be updated as new batches of data become available, enhancing its accuracy and adaptability over time. This model approximates OPF solutions in real-time operation conditions; no further optimization is necessary for this model. The direct mapping method provides more rapid solutions than the traditional iterative solution-based OPF approach, particularly for large-scale power systems.

Predicting active and chance sets, rather than determining scenario-based heavy computation deterministic conditions for the ACOPF problem, utilizes approximate bounds on the joint chance constraints [13–19]. These studies have explored how to handle uncertainty for chance-constrained OPF problems. Reference [18] forecasted errors as random variables in an SCOPF problem and used chance constraints to limit technical violations, thereby requiring reformulation. Another study [19] proposed a generalized framework for chance-constrained OPF problems under Gaussian uncertainties, emphasizing the effectiveness of affine feedback policies across various distribution types. The authors of reference [15]

proposed joint chance constraints for the OPF problem, addressing multiple constraints simultaneously with a specified probability and creating a less conservative set of single chance constraints for more reliable solutions. In reference [13], the authors presented a streaming process design to learn active sets from training samples, including input parameters and optimal solutions across diverse problem types and structures, unrestricted by the probability distribution of the inputs. When given the system's current or future conditions, predicting active and chance sets relies on determining which constraints are binding in the OPF problem. Predicting active constraints helps to evaluate active remaining constraints throughout the OPF process, significantly supporting the reduction in the solution time. The OPF model solver depends on previously predicted active constraints. On the other hand, chance constraints determine and account for uncertainties while also satisfying particular probability conditions. Classification models are suitable for predicting the probability of active constraints for the system's state. The process of predicting active and chance constraints is similar to the direct mapping technique, except that it includes the active constraints. In this model, the accuracy of the predicted active and chance constraints is crucial because the OPF solvers rely on predicted constraints to speed up the process.

Another data-driven OPF approach is the PINN, which combines the NN model and incorporates physical constraints for the system to satisfy those constraints. PINN ensures that predicted power flow solutions are consistent with network topology and power system operation limits by introducing additional loss functions in the NN [20–23]. This survey [23] provided a comprehensive overview of PINN in the power systems domain, summarizing various PINN paradigms such as PI loss functions, PI initialization, PI architectural design, and hybrid physics–deep learning models. The authors in this study [24] introduced a framework for PINN in power systems and optimized training parameters using physical laws. Their approach simplified the network's structures and improved the accuracy of PINN, specifically in rapidly determining dynamic states and uncertain parameters such as inertia and damping for single-machine infinite bus systems. This study [20] presented an approach using the PINN scheme. It utilized stacked extreme learning machine (SELM) frameworks. Their approach simplified the OPF model by decomposing it into three stages to reduce complexity and bias. The method included a classification strategy for better feature extraction and is adaptable to various systems due to SELM minimal hyperparameter tuning. The approach in [21] integrated the AC power flow equations in NN during the training phase, thus reducing worst-case constraint violations and maintaining prediction optimality. The PINN is particularly important when the RE outputs are variable and difficult to predict. Therefore, PINN can ensure that predicted power flow solutions are feasible and within operational conditions. PINN accuracy tends to be higher than the previously introduced data-driven OPF models. However, implementing PINN requires developing the set of physical constraints into NN, a challenging task. While there have been successful implementations of PINNs in certain network topologies [20], generalizing this approach across a wide range of topologies remains a significant challenge. This complexity arises during the training phase, where researchers might encounter issues due to limited, accurate datasets representing system states and conditions.

Predicting warm start points for ACOPF, instead of solving the OPF from scratch, the NN predicts the optimal decision for continuous and binary variables. Power flow solvers utilize those predicted variables later [15,25–27]. The author [25] demonstrated an NN model to emulate ACOPF solvers without enforcing specific training set restrictions. The model was trained using data from previous ACOPF runs. The model operated iteratively, making small steps toward the optimum instead of directly predicting the ACOPF solution. This work [28] introduced an ML framework called Smart-PGSim that accelerated power system simulations using multitask-learning neural networks to predict initial values. Their work enhanced problem-solving efficiency and maintained solution accuracy. In this study, ref. [29] proposed a method for rapidly learning warm start points using a multiple-target binary decision tree with post-pruning to speed up solving processes. Due to the nonlinear

and non-convex problem, the ACOPF problem can result in falling into local optimal points throughout the process. A good starting point (warm start) dramatically influences the quality of the result, reducing convergence speed and avoiding local minima optima points. Therefore, the warm start method enhances and reduces computational complexity and efficiently gives near-real optimal solutions for the OPF problem.

While solving data-driven OPF models, each related methodology offers a unique strength. The best approach depends on specific application and system challenges. At the same time, the methods mentioned earlier are generally primary for solving the data-driven OPF approach. Several other methodologies, which include machine learning (ML) and DL, are notably applied to execute OPF problems. It is essential to recognize that advanced DL algorithms, known for their higher accuracy and lower error rates, enhance the capability to tackle complex problems in various fields. This survey [30] explored ML and DL methodologies in energy systems. There are several methods applied for ML and DL, such as DNN [12], convolution neural network (CNN) [27,31,32], reinforcement learning (RF) [33], Gaussian process (GP) [34], graph neural network (GNN) [31,35,36], Lagrangian-based approaches [37], compact learning and principal component analysis (PCA) [38], meta-learning [39], and the learning-aided OPF approach [40]. A hybrid or combined model reunites physics-based and statistical methods or two or more individual methods [41–43]. The combined model utilizes the individual models' strengths and produces more accurate results for the data-driven problem. The combined approach leverages an effective relationship of various features and structures of the data-driven models and optimization framework, which is not commonly utilized by individual methods [37]. This integration increases the depth and scope of information extraction and efficient exploration while adding robustness against the potential risks of relying on a single methodology. To this end, the significant contributions of this study are as follows:

- A two-stage combined convolution and deep feed-forward neural network (CNN-DFFNN) DL technique is proposed for the data-driven ACOPF problem. This methodology enhances prediction accuracy compared to individual CNN, deep feed-forward neural network (DFFNN), hybrid CNN-DFFNN, and transfer learning (TL)-based methods. The approach leverages the spatial feature extraction capabilities of CNN for analyzing grid topologies and subsequently utilizes the regression capabilities of DFFNN.
- The proposed multiple-input, multiple-output (MIMO) methodology aims to utilize the traditional OPF solver mechanism. The grid parameters take inputs for loads and produce all the generator's set points and voltages/angles of each bus in a single attempt. The applied approach is designed for a comprehensive computation in one execution, streamlining the process of optimization outputs by providing a complete set of control and state variables.
- Integration of high-penetration-level wind power (WP) that is non-parametric in nature into the IEEE 39 bus system was utilized from the physical WP model. This approach was chosen to ensure that the variability and intermittency inherent in wind energy were accurately captured, reflecting a realistic simulation environment.

The organization of the rest of this paper is as follows. The OPF formulation is presented in Section 2. Data and procedures are covered in Section 3. The DL methodology for the ACOPF problem is demonstrated in Section 4. Results and discussion are presented in Section 5. The final section, Section 6, provides information about the conclusion.

2. Optimal Power Flow

General OPF Problem Formulation

This section describes the formulation of the ACOPF. The ACOPF model establishes the power system's operational states, considering the relationship between the fuel price cost for generators and the supplied power of each generator as the current system operational constraint. The objective function of ACOPF is to minimize the sum of the generation costs of all generating units, and this objective function can be expressed as follows:

$$\text{Minimize} \sum_{i \in S_G} \left(a_i P_{Gi}^2 + b_i P_{Gi} + c_i \right) \tag{1}$$

P_G represents the active power generation of the i-th generating units within the set of all generators S_G, and ai, bi, ci are the coefficients of the cost function for the generators. Furthermore, the objective function is optimized subject to power flow balance constraints, as detailed in Equations (2) and (3). Additionally, Equations (4) and (5) define the branch flow constraints.

The power balance constraint equations are as follows:

$$\sum_{g \in i} (P_{Gi} - P_{Di}) = \sum_{j=1}^{N} |V_i||V_j|(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \tag{2}$$

$$\sum_{g \in i} (Q_{Gi} - Q_{Di}) = \sum_{j=1}^{N} |V_i||V_j|(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \tag{3}$$

Q_G represents the reactive power generation, while P_D and Q_D denote the active and reactive power demand at bus i, respectively. The voltage magnitudes at buses i and j are denoted by $|V_i|$ and $|V_j|$. The real and imaginary parts of the admittance matrix are represented by G_{ij} and B_{ij}, respectively. The phase angle difference between buses i and j is described by $\theta_{ij} = \theta_i - \theta_j$. Lastly, N indicates the total number of buses in the system.

The branch flow constraints are as follows:

$$P_{ij} = G_{ij}(V_i^2 - |V_i||V_j| \cos \delta_{ij}) - B_{ij}|V_i||V_j| \sin \delta_{ij} \tag{4}$$

$$Q_{ij} = -B_{ij}(V_i^2 - |V_i||V_j| \cos \delta_{ij}) - G_{ij}|V_i||V_j| \sin \delta_{ij} \tag{5}$$

Meanwhile, the system operation constraints are presented in Equations (6)–(8).

The operation limit constraints are as follows:

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max} \tag{6}$$

$$Q_{Gi}^{\min} \leq Q_{Gi} \leq Q_{Gi}^{\max} \tag{7}$$

$$V_i^{\min} \leq V_i \leq V_i^{\max} \tag{8}$$

$P_{Gi}^{\min}$ and $P_{Gi}^{\max}$ denote the minimum and maximum allowable active power limits for the ith generators, respectively. Similarly, $Q_{Gi}^{\min}$ and $Q_{Gi}^{\max}$ represent the reactive power limits. Lastly, $V_i^{\min}$ and $V_i^{\max}$ demonstrate the voltage limits at bus i.

The transmission line constraints are as follows:

$$\sqrt{P_{ij}^2 + Q_{ij}^2} \leq S_{ij,\max} \tag{9}$$

S_{ij} represents the apparent power constraint between nodes i and j. The power transmitted along a line must be within its designated transmission capability limit.

3. Wind Energy and Load Data for Simulation

3.1. Modified IEEE 39 Bus System

The IEEE ISO New England (ISONE) 39-bus test system, as described in [44,45], consists of 39 buses, 21 loads, 10 generators, and 46 branches. The test system has a maximum loading capacity of 6254 MW and 6626 MVA. Additionally, the system operates with a base power of 100 MVA and a base voltage of 345 kV. In this study, the test system was modified by integrating WP plants into six specific buses.

Figure 1 represents the modified test system, including the locations of renewable WP plant integration at buses 4, 8, 16, 20, 28, and 29.

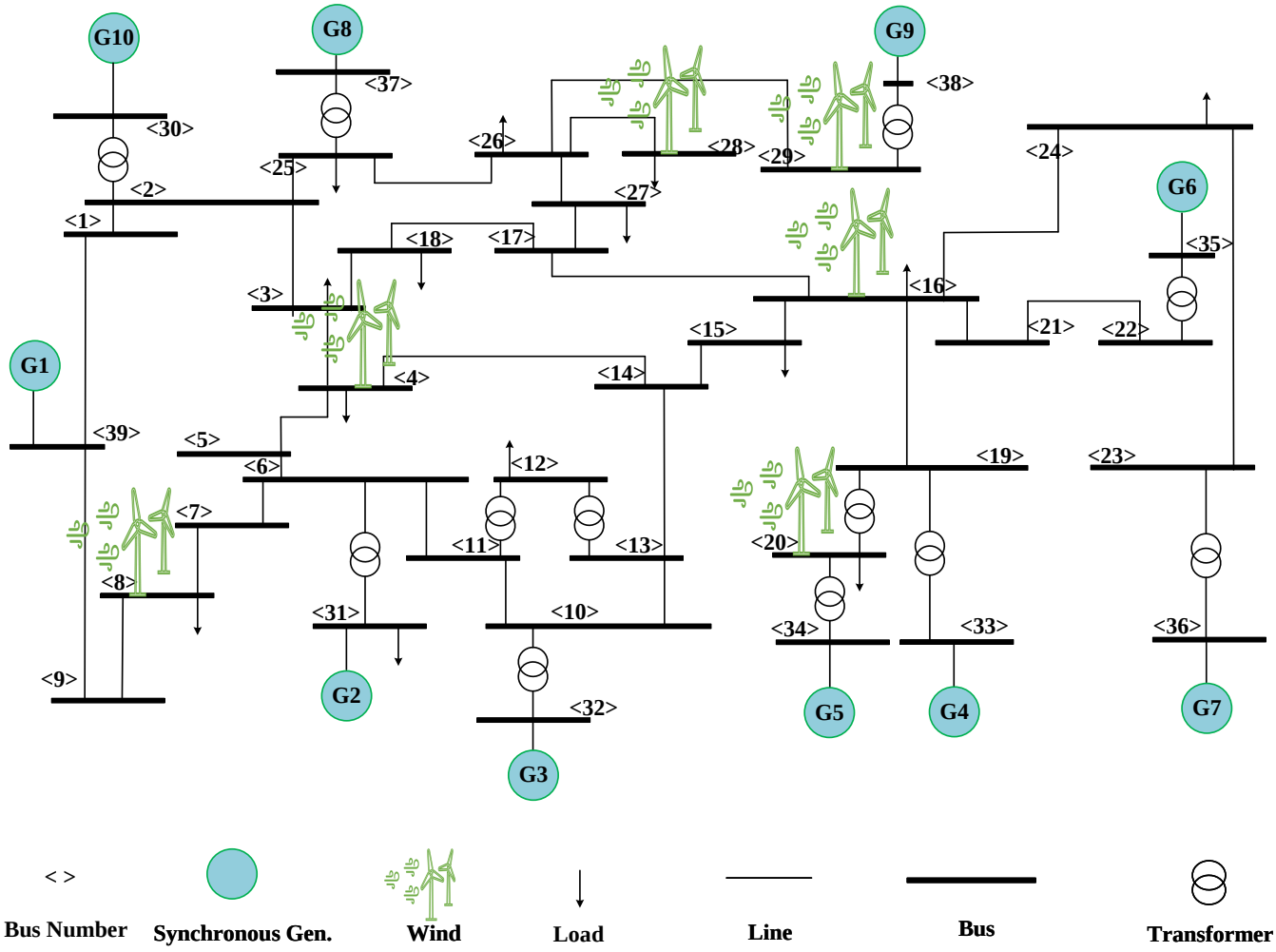


Figure 1. The modified IEEE 39-bus test system.

Furthermore, Table 1 describes the total number of wind turbines (WTs) integrated into the system and provides the maximum generation capacity in MW for each location. For example, bus 4 has integrated 70 WTs, with a maximum energy generation capacity of 370 MW for the system. The rest of the buses follow with their respective capacities and number of WTs, as listed in the table. The system's renewable WP plant integration accounts for 30% of the total system loads. The reactive power has been adjusted to maintain the same power factor for all buses to mimic the initial conditions.

Table 1. Renewable wind power integration for IEEE 39-bus system.

Bus Number	Max. MW Amount	Number of Wind Turbines
4	370	70
8	390	72
16	240	45
20	520	94
28	160	28
29	210	40

3.2. Wind Data and Power Production of Wind Power Plants

This subsection provides information on several aspects regarding WP. It covers wind data generation, statistical modeling, electricity generation from WP plants, and integrating

RE resources into different buses. A Weibull distribution represents the variability of wind speed (WS) and subsequently derives electrical power production [46]. Through Weibull analysis, the data can be obtained for suitable planning of a WP plant for a specific location and evaluating the power characteristic of the wind turbine (WT). Two critical parameters are required to determine the Weibull probability density function (PDF): the scale factor 'c' and the shape factor 'k'. The equation corresponding to the PDF of the Weibull distribution is presented in Equation (10):

$$v(k,c) = \begin{cases} (k/c)\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^{k}} & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \quad (10)$$

The Weibull probability density function (PDF), denoted as f, represents the conditional probability of WS (v). The Weibull shape and scale factor are denoted as k and c, respectively. The unit of v is m/s, whereas k and c are unitless. To represent the variability of WS, c and k are prescribed as 12.2 and 2.1, respectively. With these parameters, a simulation of 1000 WS observations was conducted to generate outputs for a 5.6 MW WP turbine. The data from the Vestas V162-5.6 MW were used for this study [47].

Table 2 provides details of the WT's operating and rotor data. The WT's essential characteristic speeds are its cut-in, rated, and cut-out WS. The turbine rotates and generates electrical power at the cut-in WS region. The WT continues to generate electrical power based on the power curve equations until the rated speed region. At the nominal WS region, the power output from the WT reaches the maximum electrical generation capacity limit. In the cut-out region, the WT is adjusted to force it to stop the turbine blades at the cut-out WS for safety purposes to prevent potential damage. The WT equations are presented as follows in Equation (11):

$$P(v_s) = \begin{cases} 0 & \text{if } v_s < v_{\text{cut-in}} \\ \frac{1}{2} A \rho c_f v_s^3 & \text{if } v_{\text{cut-in}} \leq v_s \leq v_{\text{rated}} \\ P_{\text{rated}} & \text{if } v_{\text{rated}} \leq v_s \leq v_{\text{cut-out}} \\ 0 & \text{if } v_s > v_{\text{cut-out}} \end{cases} \quad (11)$$

- P = Power output of wind turbine (W);
- v_s = Wind speed (m/s);
- A = Swept area of the rotor (m^2);
- ρ = Air density (kg/m^3);
- c_f = Capacity factor.

Table 2. Vestas V162-5 wind turbine data.

Parameter	Value	Unit
Generators' Rated Power	5600	kW
Turbine Cut-in Wind Speed	3	m/s
Turbine Cut-out Wind Speed	25	m/s
Turbine Rotor Diameter	162	m
Turbine Swept Area	20,612	m^2

The physical characteristics equations of the wind turbine were utilized to capture the WT power curve, along with the cut-in, cut-out, and rated WS regions. Subsequently, power curves from multiple WP plants were aggregated to represent the integration of large turbines across various bus locations. Figure 2 illustrates the aggregated power curves of six WTs from different bus locations and their respective number of WTs.

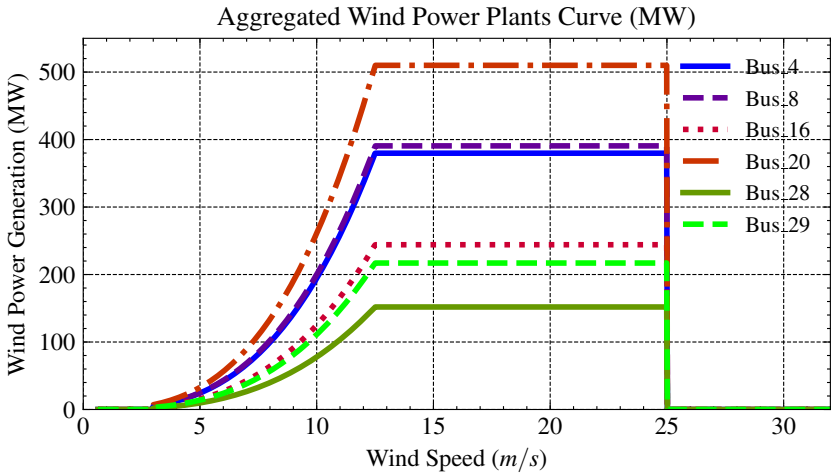


Figure 2. Wind power curve for different buses and capacities.

3.3. Load Data Variations for the Implementation

This subsection illustrates the use of Monte Carlo simulation to represent the variability and fluctuations of system loads. Load variations for active and reactive power at all 39 buses were randomly sampled. This sampling followed a normal distribution with a standard deviation of 10% of the mean. A total of 1000 simulations were executed, resulting in 39,000 data points for active and reactive power across the load buses.

Notably, 21 of the simulation cases did not converge. Therefore, unsolved simulation results were eliminated. As a result, the active power and reactive power of 38,181 data points were established to represent system loads. Figure 3 illustrates active and reactive power load variability for bus 4. While Monte Carlo simulations generated active and reactive power load data for all 39 buses, only bus 4 is presented here for the sake of simplicity and brief illustration.

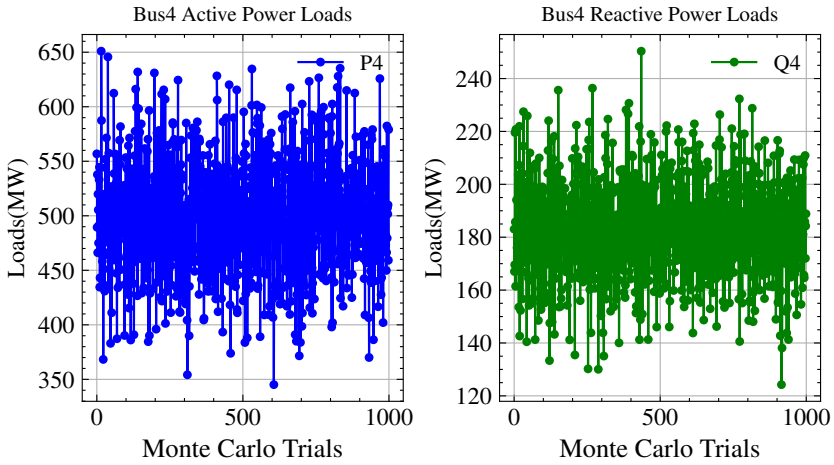


Figure 3. Monte Carlo samples for loads.

3.4. Simulation Results to Prepare Deep Learning Model

This section provides data preprocessing, training, validation, test split, feature scaling, and an example of data visualization. Utilizing Newton–Raphson as the solver in MATPOWER [45], the data of the ACOPF simulation results were simulated and accumulated. Table 3 provides a concise summary of data points collected from the simulations, detail descriptions such as active and reactive power loads, voltages and angles, WS and corresponding WT power outputs, branch current, conventional generator power outcomes, and electricity production cost.

Table 3. Summary of data points from simulation.

Description	Number of Data Points
Active and reactive power loads	38,181
Voltage magnitude and angles	38,181
Wind speed and wind turbine active and reactive power	979
Generator active and reactive power	9790
Branch current active and reactive power	45,034
Generators' objective cost functions	979

Later, the simulation dataset was reorganized to separate information for each bus load and generator. The refined data contain 979 data points for each bus's active and reactive power for loads as well as voltage and angles. In addition, 979 data points were restructured for generators, capturing both active and reactive power. The generated dataset will be uploaded and serve as the primary data source for our deep learning model training [48]. The structured dataset from the simulations feeds into the deep learning model. The dataset was divided into 65% for training, 15% for validation, and 20% for the test set. After the dataset split, the RobustScaler was utilized for feature scaling. RobustScaler is effective in identifying outliers and scales to features that might be robust against outliers. Although the Min-Max scaler was also experimented with, the RobustScaler provided superior results compared to the Min-Max scaler.

Figure 4 presents a 3D plot that showcases 24 simulation results derived from variable loads using Monte Carlo methods and the maximum integration of WP plants into the test network, corresponding to the voltage profile results of each bus.

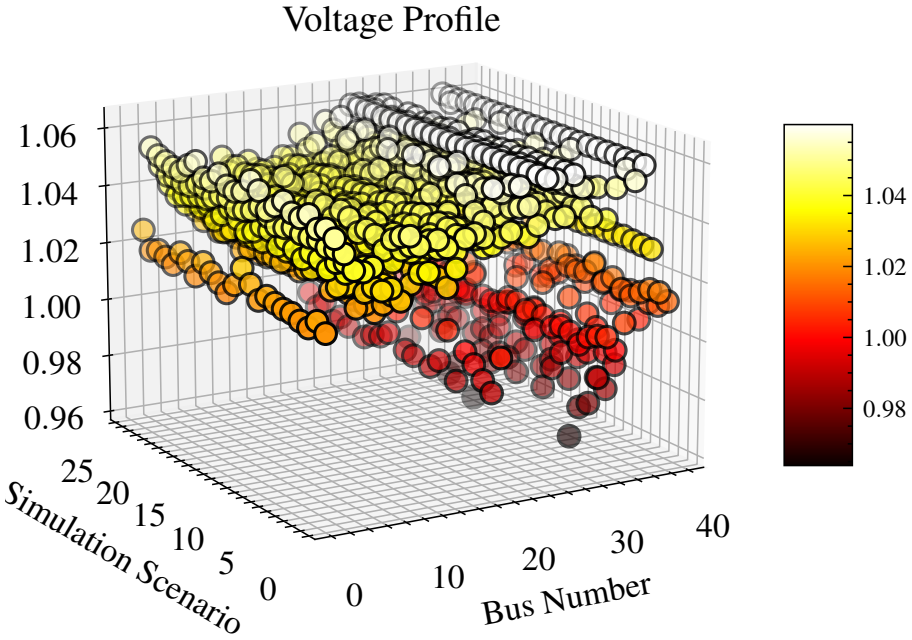


Figure 4. Simulation results for integration of wind power.

4. Deep Learning Methodology for ACOPF

4.1. Convolution Neural Network (CNN)

The CNN DL model commonly uses convolutional, pooling, and fully connected dense layers. In the context of OPF, convolutional layers are particularly effective in identifying the spatial relationships among grid variables. Pooling layers contribute to this process by systematically decreasing the dimension of data, which mitigates computational complexity and the risk of overfitting. Following the feature extraction and pooling layers, the fully connected layers are crucial in recognizing patterns of OPF variables, which predicts the optimization variables. For OPF, these output variables typically include the generation

dispatch for active and reactive powers, voltage magnitudes, and phase angles of the grid. For a detailed analysis of CNN as applied in this research, refer to [49], and additional background information about CNN can be found in [50].

4.2. Deep Feed-Forward Neural Network (DFFNN)

The DFFNN comprises an input layer, several hidden layers, and an output layer. The hidden layers of the DFFNN are key in capturing and modeling the non-linear relationships between the input grid variables and the output optimization variables. DFFNN is primarily used to solve supervised learning and classification problems by taking a set of inputs and performing complex transformations on output layers via hidden layers. Specifically for the OPF problem, DFFNN approximates the generation level, including the corresponding generators' active and reactive power as well as voltage levels and phase angles for system buses mapping from the given system load and power network configurations.

4.3. Transfer Learning (TL)

The primary goal of TL is to allow solving a problem and leveraging the previously learned knowledge to apply related problems without starting the learning of the knowledge from scratch. CNN and DFFNN models require a significant number of high-quality labeled data, and it is expensive and often complex to obtain those labeled datasets. The model accuracy tends to be low due to limited data. On the other hand, TL enables a warm start while utilizing previously learned patterns and model training weights, and it can increase the converging speed and often provides robust accuracy. TL model accuracy relies on several factors, such as the similarity of the target task, fine-tuning and feature extraction models, and the model complexity. If the pre-training model is used for feature extraction without fine-tuning, the TL model might not perform optimally or generalize effectively to capture the target task.

4.4. Problem Methodology of Combination of CNN-DFFNNs

This study concentrated on the ACOPF problem, thus reflecting regression tasks to provide grid control parameters such as load active and reactive power (P_{Di}, P_{Qi}), and predict the active and reactive power (P_{Gi}, Q_{Gi}) for the all the generators in the system. Additionally, predictions are generated for all bus voltage and angles (V_i, δ) from the load's active and reactive power. The proposed DL architecture is shown in Figure 5. The proposed methodology aims to predict all the generator's outputs in a single shot after training the model. The studied combined model architectures utilize five stacked CNNs and five stacked DFFNNs.

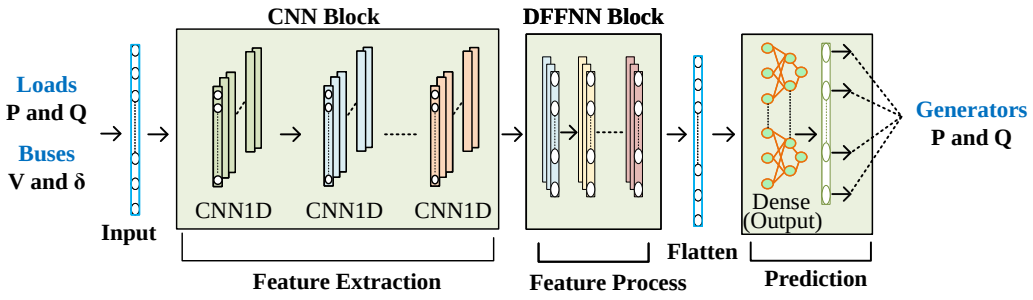


Figure 5. Proposed combined CNN-DFFNN architecture for generators (a).

Table 4 provides a detailed configuration of the combined CNN and DFFNN architecture. It highlights the filter, kernel sizes, and activation functions employed in the CNN and DFFNN layers. The model hyperparameters were kept the same for all six case studies to prevent bias from one model to another. Training parameters such as epochs, batch size, loss function, and learning rate used for model optimization are also demonstrated.

Additionally, it presents the output layers designated for three separate case studies for Sections 5.1–5.3.

Table 4. Combined CNN and DFFNN architecture.

CNN Filter Sizes	512, 256, 128, 64, 32, 16
CNN Kernel Size	2
Activation Function	ReLU
DFFNN Layer Sizes	512, 256, 128, 64, 32, 16
Activation Function	ReLU
Flatten Layer	Applied
Dense (Output) Layers (Section 5.1)	78
Dense (Output) Layers (Section 5.2)	20
Dense (Output) Layers (Section 5.3)	20
Epochs	500
Batch Size	32
Loss Function	MSE
Learning Rate	0.0001
$\beta_1, \beta_2, \epsilon$	$0.9, 0.999, 1 \times 10^{-7}$

5. Results and Discussion

This section summarizes the findings of six studied deep learning methods for the ACOPF problem. The initial two models in our analysis are traditional deep learning methods DFFNN and CNN. These models serve as a foundation of performance baseline for hybrid models. Following the models, we introduce two hybrid models that combine the CNN and DFFNN architectures, with each configuration as depicted in Figures 5 and 6.

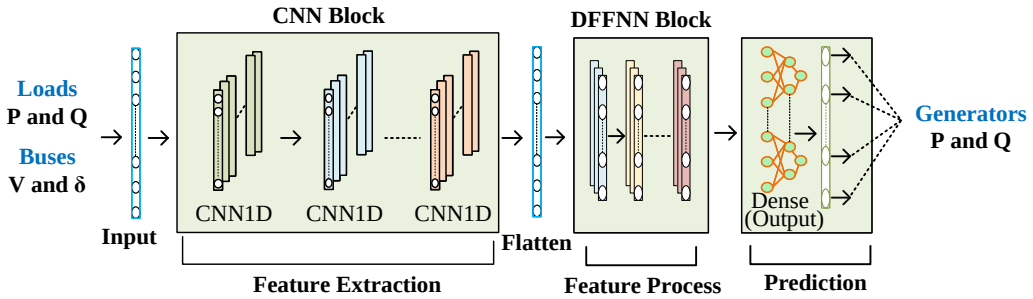


Figure 6. Hybrid CNN-DFFNN architecture for generators (b).

The first hybrid model, labeled as CNN-DFFNN (a), represents the proposed combined CNN-DFFNN architecture for generators (a), as depicted in Figure 5. The second hybrid model, labeled as CNN-DFFNN (b), represents the hybrid CNN-DFFNN architecture for generators (b), shown in Figure 6. These hybrid models demonstrate variations in performance evaluation due to their distinct architectural approaches.

The model structures for the TL approaches, TL (a) and TL (b), are built upon the architectures previously detailed in Figures 5 and 6 of the paper. The critical distinction in the TL models lies in their foundation on the pre-trained CNN models. This approach demonstrates the effectiveness of using pre-trained models without extensive fine-tuning.

The study concentrated on MIMO DL techniques to predict the generator's active and reactive power and the voltage and phase angle for each bus result. The proposed model utilizes all 39 buses' load active and reactive powers, voltage, and phase angles as an input and predicts all generators' active and reactive power as an output. This section is divided into three case studies for each DL model accuracy and efficient, and results are presented in Tables 5–10.

5.1. MIMO Forecasting for Bus Voltages and Angles

The first case study focuses on utilizing the active and reactive power loads as input to predict the voltages and angles for each bus.

$$X_1^T = [P_1^L, Q_1^L, \ldots, P_n^L, Q_n^L] \tag{12}$$

X represents the input load feature vector in the given dataset starting from bus 1 to bus n. In this network model, n is equal to 39. The letter L represents load, and n represents the bus numbers.

$$Y_1^T = [V_1^B, \delta_1^B, \ldots, P_{gn}^B, Q_{gn}^B] \tag{13}$$

Y demonstrates each bus's associated voltages and angles. The primary objective is to accurately match the input feature vectors to the respective output vectors. This mapping ensures that voltage and angles are satisfied while adhering to the system constraints. The goal is to minimize the loss function, between the predicted (V_{pred}, δ_{pred}) and the actual values (V_{actual}, δ_{actual}).

To investigate the efficacy of combined DL approaches, two primary performance metrics, root mean squared error (RMSE) and mean absolute error (MAE), are considered for the study. These performance evaluation measures are defined by Equations (14) and (15).

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{n=1}^{M}(y_n - \hat{y}_n)^2} \tag{14}$$

$$\text{MAE} = \frac{1}{M}\sum_{n=1}^{N}|y_n - \hat{y}_n| \tag{15}$$

y_n and $\hat{y}_n$ indicate the observed and forecasted generators' active and reactive power results, respectively. The symbol n represents individual simulation samples and M denotes the total number of simulation datasets.

Predictions were made for all bus voltages and phase angles in a single attempt. Figures 7–9 represent the DFFNN, the CNN, and the proposed model visualization results in terms of voltage magnitude and phase angle results, respectively.

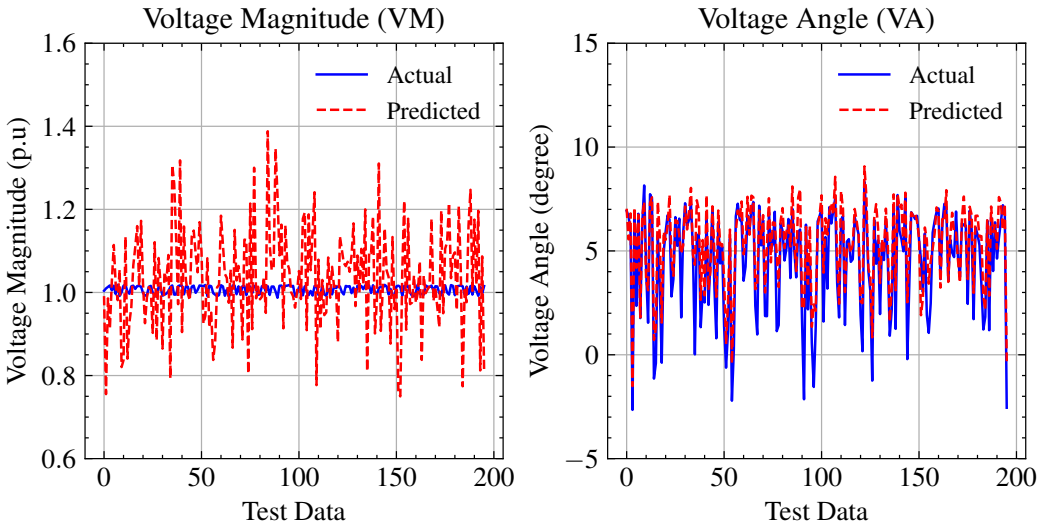


Figure 7. DFFNN-based prediction and actual results.

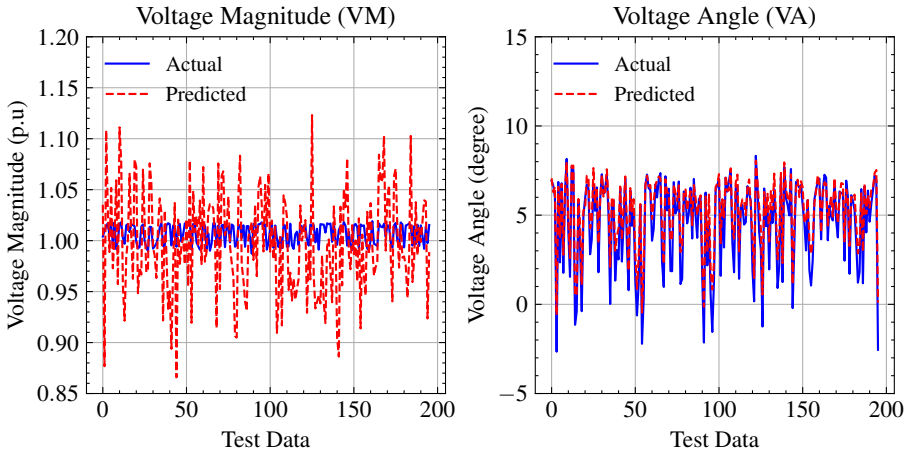


Figure 8. CNN-based prediction and actual results.

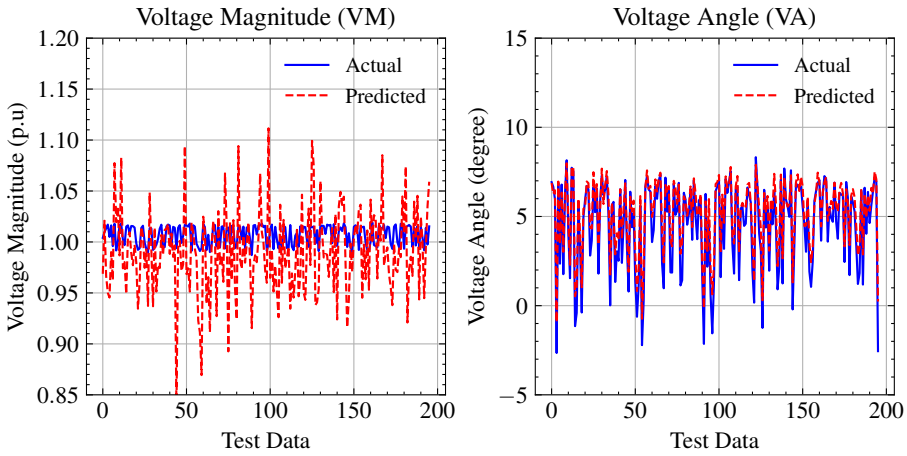


Figure 9. CNN-DFFNN (a)-based prediction and actual results.

The worst-performing model was the DFFNN, as shown in Table 5, and it was not able to capture the voltage magnitudes because the predicted voltage magnitudes reached as high as 1.4 p.u and dropped as low as 0.77 p.u. Significant predicted voltage deviations could lead to model inefficiency for the data-driven OPF problem.

On the other hand, the proposed method, as depicted in Figure 9, demonstrated its capability to predict more accurate results regarding voltage magnitude within the range of 1.11 and 0.85 p.u and phase angles. Although the proposed model was not the best-performing model, the CNN model was the best-performing model.

Table 5 shows the results of RMSE and MAE errors while considering each bus voltage and angles as outputs when taking each load active and reactive power to serve as inputs.

Table 5. Training and validation dataset results.

DL Model	Train RMSE	Train MAE	Val. RMSE	Val. MAE
DFFNN	0.230	0.135	0.614	0.356
CNN	0.121	0.079	0.191	0.118
CNN-DFFNN (a)	0.135	0.081	0.233	0.135
CNN-DFFNN (b)	0.256	0.162	0.313	0.197
TL (a)	0.319	0.172	0.389	0.207
TL (b)	0.406	0.225	0.458	0.255

The results in Table 5 demonstrate that the CNN was the best-performing model, although results were indistinguishable in terms of the lowest RMSE and MAE values on

both training and validation datasets when comparing the proposed model. On the other hand, DFFNN was the least effective model.

When comparing the hybrid and the proposed models, the combined model outperformed the traditional one. Similarly, when considering the TL model with different configurations, the TL model with CNN-DFFNN and flatten configuration performed better than the CNN, flatten, and DFFNN models.

Table 6 provides a comparative analysis of various models, focusing on their size, training duration, and speed of prediction to assess their overall performance efficiency and the associated computational expenses.

Table 6. DL model efficiency performance.

DL Model	Trainable Param.	Non-Trainable Param.	Training Time	Prediction Time
DFFNN	216,862	0	1 min 21 s	135 ms
CNN	441,150	0	1 min 40 s	171 ms
CNN-DFFNN (a)	624,942	0	2 min 37 s	365 ms
CNN-DFFNN (b)	1,117,966	0	1 min 54 s	213 ms
TL (a)	273,726	351,216	2 min 17 s	337 ms
TL (b)	766,750	351,216	1 min 37 s	230 ms

The DFFNN model had the shortest training and prediction times, indicating high efficiency. The CNN model required more training time than the DFFNN and had a slightly longer prediction time. The CNN-DFFNN (a) took the longest to train and had the longest prediction time, suggesting a trade-off for its complexity.

The CNN-DFFNN (b) and TL (b) models were more efficient in training and prediction times than their counterparts. However, they had a more significant number of parameters. The comparative additional results of the model can be found in the Appendix A, as referenced in Appendix A.1.

5.1.1. Section Key Findings

- The DFFNN model was the least effective, especially in capturing voltage magnitudes.
- The proposed method demonstrated better accuracy in predicting voltage magnitudes and phase angles.
- The CNN model appeared to be the best performer.
- The proposed model demonstrated superior performance over the traditional hybrid model.
- The DFFNN model had the highest efficiency with the shortest training and prediction times.
- The CNN-DFFNN (a) model had the longest training duration and prediction time.

5.2. MIMO Forecasting for Generators' Active and Reactive Powers from Loads

In the second case study, the focus is on utilizing the active and reactive power loads as input parameters and predicting the generators' active and reactive power outputs.

$$X_2^T = [P_1^L, Q_1^L, \ldots, P_n^L, Q_n^L] \tag{16}$$

In this second case, the input feature vector remains consistent with that of case one. However, the distinction lies in the output feature vector for the generator's optimal dispatch setting. The primary objective in this scenario is for the model to predict all generators' output values accurately.

$$Y_2^T = [P_{g1}^G, Q_{g1}^G, \ldots, P_{gn}^G, Q_{gn}^G] \tag{17}$$

Figures 10–12 demonstrate predicted and actual results for one of the generators, the poorest, the intermediate, and the proposed model results in the system.

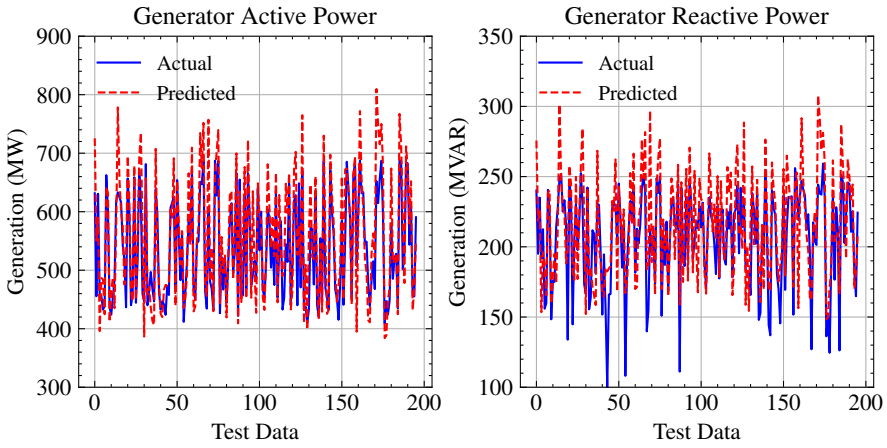


Figure 10. DFFNN-based prediction and actual results.

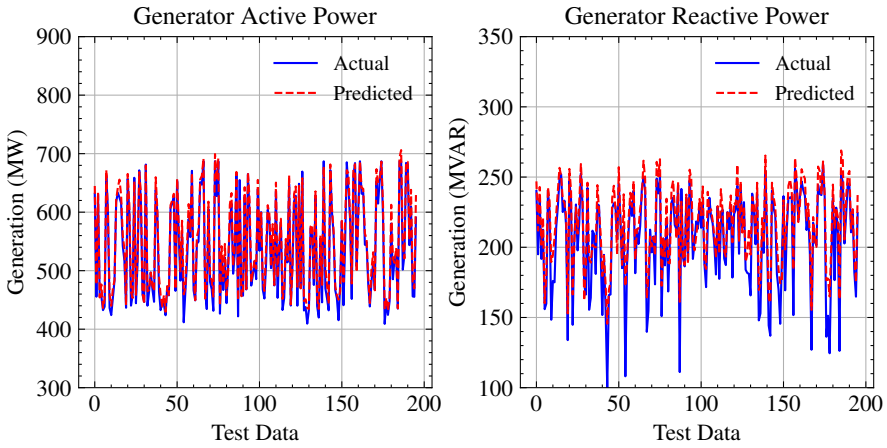


Figure 11. CNN-based prediction and actual results.

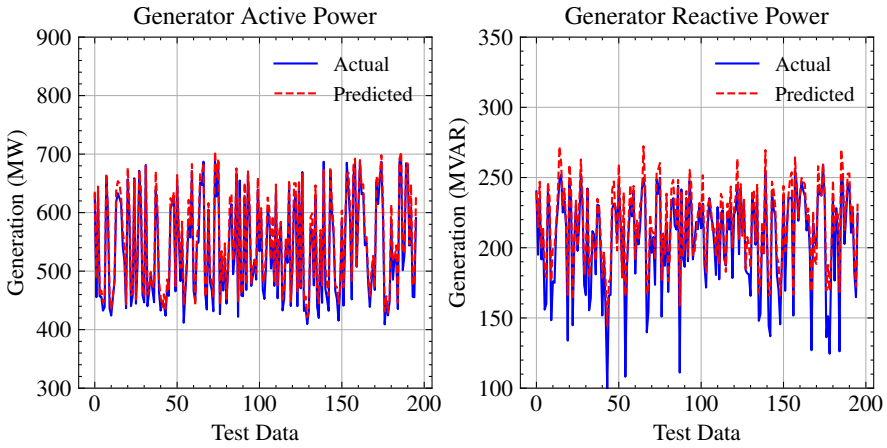


Figure 12. CNN-DFFNN (a)-based prediction and actual results.

The predicted MW and MVAR outputs of generators have a more significant margin than the actual generator outputs for the DFFNN model, and the CNN model visualization results are indistinguishable from the proposed model. Table 7 denotes the performance metric errors for the training and validation datasets.

The results indicate that the proposed combination of CNN-DFFNN outperformed the validation dataset RMSE and MAE by 17.89 and 13.13, respectively, compared to other DL models. The results in Table 7 represent that the CNN model remained the second best model, and the training model performance was superior to the other models. However,

on the validation datasets, the proposed model outperformed the rest. The DFFNN model is still the poorest model for this case study.

Table 7. Training and validation dataset results.

DL Model	Train RMSE	Train MAE	Val. RMSE	Val. MAE
DFFNN	18.122	12.629	43.417	31.027
CNN	11.310	8.325	18.074	13.371
CNN-DFFNN (a)	13.864	9.837	17.898	13.131
CNN-DFFNN (b)	16.878	11.355	21.012	15.728
TL (a)	20.937	14.747	23.315	16.516
TL (b)	27.911	19.978	32.048	24.401

The hybrid model, integrating both CNN and DFFNN models, performed close to the proposed model and standalone CNN model. The TL model with CNN-DFFNN and flatten configuration continued to outperform the CNN, flatten, and DFFNN models.

Table 8 presents a comparison of different methods, detailing their model size, training time, and prediction speed to evaluate performance efficiency and computational cost.

Table 8. DL model efficiency performance.

DL Model	Trainable Param.	Non-Trainable Param.	Training Time	Prediction Time
DFFNN	215,876	0	1 min 22 s	175 ms
CNN	374,276	0	1 min 38 s	367 ms
CNN-DFFNN (a)	558,068	0	2 min 38 s	336 ms
CNN-DFFNN (b)	1,116,980	0	1 min 57 s	219 ms
TL (a)	206,852	351,216	2 min 15 s	353 ms
TL (b)	765,764	351,216	1 min 31 s	235 ms

The DFFNN model stands out for its training speed and swift prediction capability. On the other hand, the slowest model was the CNN in this case study. When comparing combined and hybrid models, CNN-DFFNN (a) and (b), the hybrid model offered a more efficient model than CNN-DFFNN (a). For the TL models, the TL (b) hybrid model configurations were more outstanding than the combined model in terms of model efficiency. The comparative additional results of the model can be found in the Appendix A, as referenced in Appendix A.2.

5.2.1. Section Key Findings

- The proposed CNN-DFFNN (a) combination outperformed other DL models.
- The CNN models were identified as the second-best performers.
- The hybrid showed performance close to both the proposed and the standalone CNN.
- The DFFNN still outperformed the rest in training speed and prediction capabilities.
- The CNN model was the slowest in this case study.
- The hybrid configuration was outstanding and more efficient than CNN-DFFNN (a).

5.3. MIMO Forecasting for Generators' Active and Reactive Powers from Loads and Buses

In the final case study, the active and reactive powers of each bus load, along with the voltages and angles for each bus, serve as the input to predict the generators' active and reactive power outputs.

$$X_3^T = [P_1^L, Q_1^L, V_1^B, \delta_1^B, \ldots, P_n^L, Q_n^L, V_n^B, \delta_n^B] \tag{18}$$

X represents the input sample space vector containing both the load's active and reactive powers and bus voltages and angles.

$$Y_3^T = [P_{g1}^G, Q_{g1}^G, \ldots, P_{gn}^G, Q_{gn}^G] \tag{19}$$

Y denotes the predicting generator's optimal dispatch settings for active and reactive powers. Figures 13–15 illustrate the worst-performing model, the CNN, and the proposed model showcases both active and reactive power prediction and actual results of one of the generators in the system, respectively.

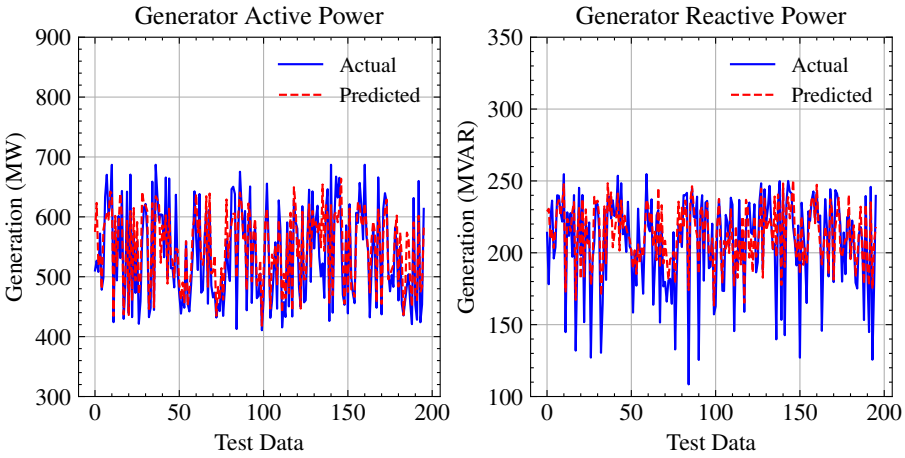


Figure 13. DFFNN-based prediction and actual results.

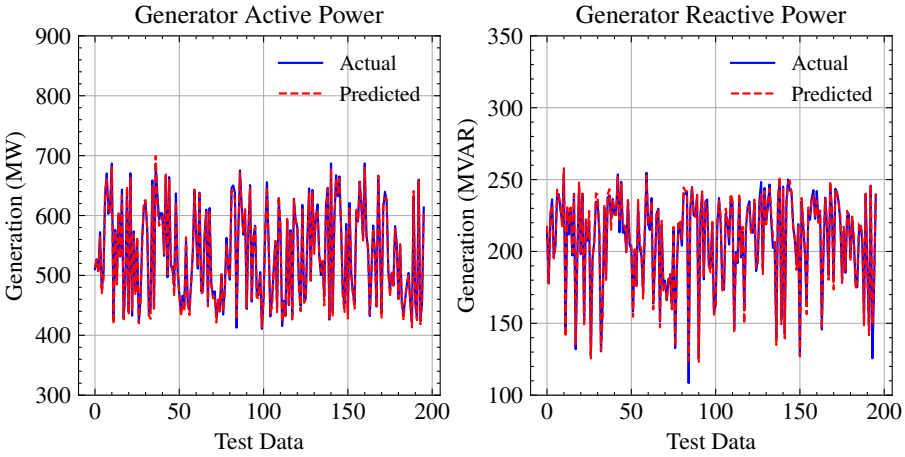


Figure 14. CNN-based prediction and actual results.

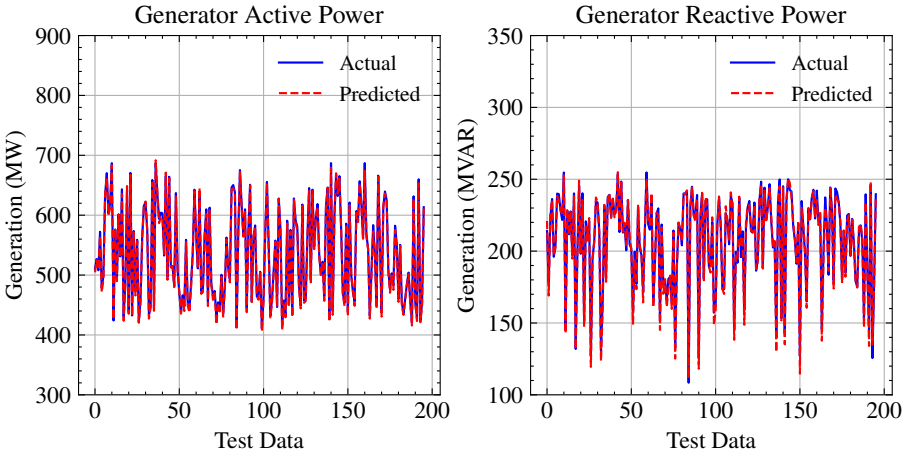


Figure 15. CNN-DFFNN (a)-based prediction and actual results.

Table 9 demonstrates that the CNN-DFFNN outperformed others on both the training and validation datasets. The rest of the deep learning models remained consistent with

the observed performance in case studies Sections 5.1 and 5.2. It can be seen that both case study Section 5.2 and case study Section 5.3 aim to predict generator outputs with additional inputs, taking into account voltage magnitude and phase angle features for case Section 5.3 when compared to case Section 5.2. The target outputs remained the same, and all of the generator's active and reactive power predictions were based on the proposed method. It can also observed that case Section 5.3's generator's output accuracy increased and enabled more accurate predictions when including additional input features. Table 10 provides a comparative analysis of various models' overall performance efficiency and the associated computational cost.

Table 9. Training and validation dataset results.

DL Model	Train RMSE	Train MAE	Val. RMSE	Val. MAE
DFFNN	15.107	10.089	39.290	24.648
CNN	5.831	3.791	8.952	6.223
CNN-DFFNN (a)	4.821	3.164	8.449	5.833
CNN-DFFNN (b)	14.273	10.378	15.213	11.124
TL (a)	14.781	10.331	16.481	11.395
TL (b)	20.876	14.451	23.311	16.682

Table 10. DL model efficiency performance.

DL Model	Trainable Param.	Non-Trainable Param.	Training Time	Prediction Time
DFFNN	255,812	0	1 min 21 s	136 ms
CNN	399,236	0	1 min 48 s	394 ms
CNN-DFFNN (a)	583,028	0	2 min 48 s	350 ms
CNN-DFFNN (b)	1,755,956	0	2 min 1 s	221 ms
TL (a)	231,812	351,216	2 min 17 s	371 ms
TL (b)	1,404,740	351,216	1 min 40 s	224 ms

In case study Section 5.3, model efficiency and computation costs remained consistent compared to the results from case study Section 5.2. The comparative analysis of RMSE and MAE performance metrics for active and reactive power outputs is illustrated in Figures 16–18.

The proposed model achieves significantly lower errors, with RMSE and MAE for active power peaking at 11 MW and 8 MW, respectively. In contrast, the MAE remained under 5 MVAR for most generators for reactive power. On the other hand, the DFFNN model exhibited higher errors, with RMSE increasing to 140 MW for active power and 125 MVAR for reactive power and the corresponding MAE reaching 110 MW and 102 MVAR. The CNN model peaked an RMSE of around 15 MW and 12 MVAR for reactive powers. These results indicate that the proposed model enhanced predictive accuracy and robustness, significantly outperforming the DFFNN model and slightly outperforming the CNN model.

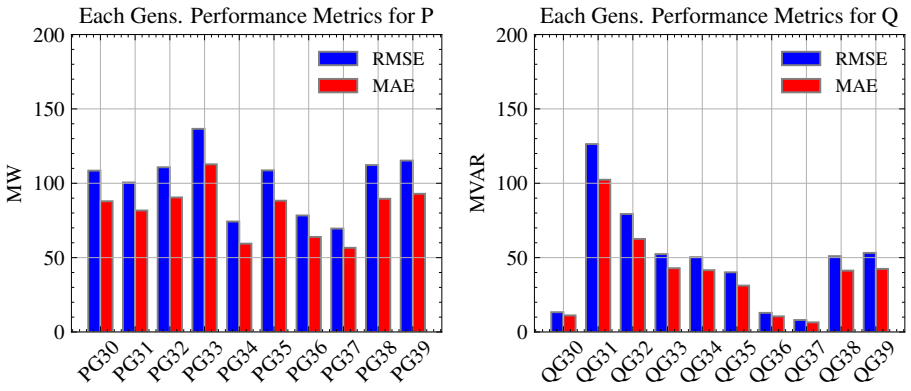


Figure 16. DFFNN-based RMSE and MAE error results for all generators.

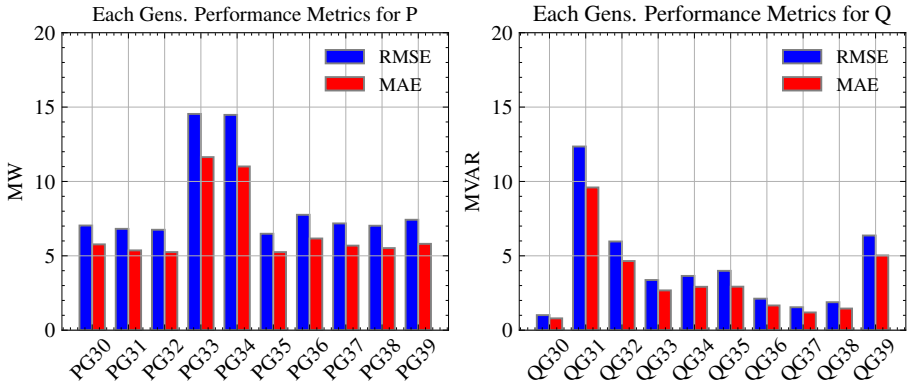


Figure 17. CNN-based RMSE and MAE error results for all generators.

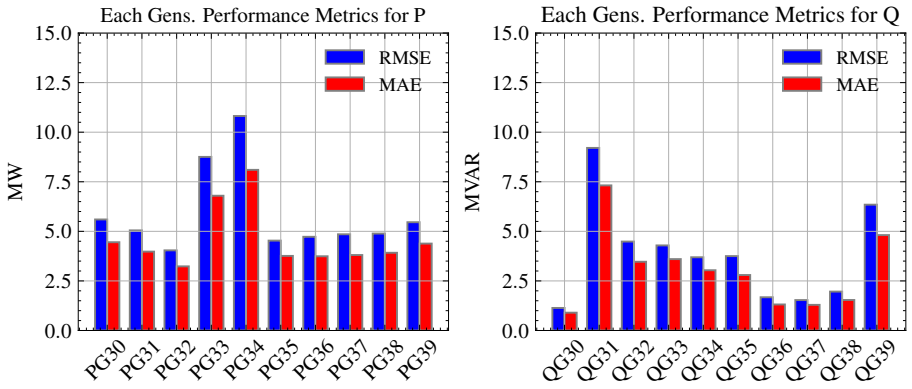


Figure 18. CNN-DFFNN (a)-based RMSE and MAE error results for all generators.

The error distribution for the comparison between the worst model, the second-best, and the best model is shown in Figures 19–21 for all ten generators' active and reactive power outputs within the system. The DFFNN model in Figure 19 operates within the error range of ±200 MW/MVAR, displaying a broader variability in predictive accuracy. Figure 20 illustrates the error variation of the CNN model across all generators. The results of the CNN model are similar to the proposed model, yet the proposed model demonstrates superior performance. On the other hand, the proposed model in Figure 21 significantly improves the prediction accuracy and operates in an error range of ±20 MW/MVAR. Despite the difference in error magnitudes, the median errors for the models are closely aligned with zero, indicating no significant systematic bias in either model. The median errors are highlighted in red, and circles demonstrate outliers in the boxplot figures.

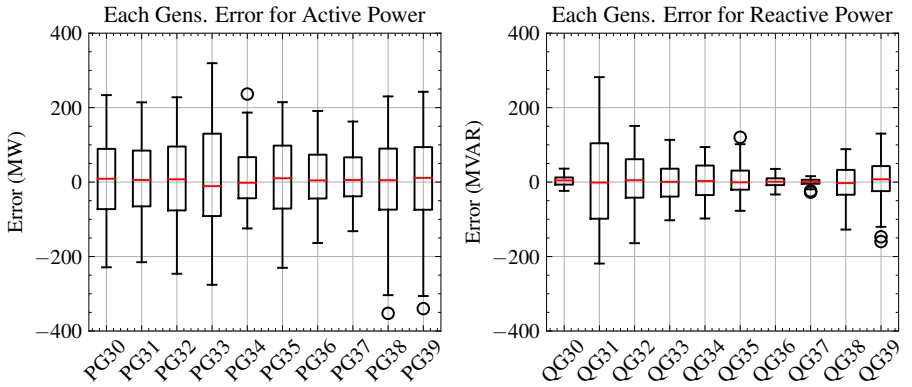


Figure 19. DFFNN-based boxplot error distribution for all generators.

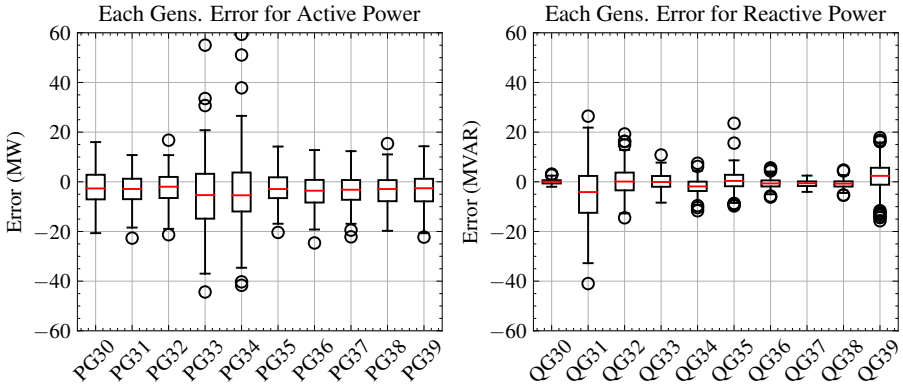


Figure 20. CNN-based boxplot error distribution for all generators.

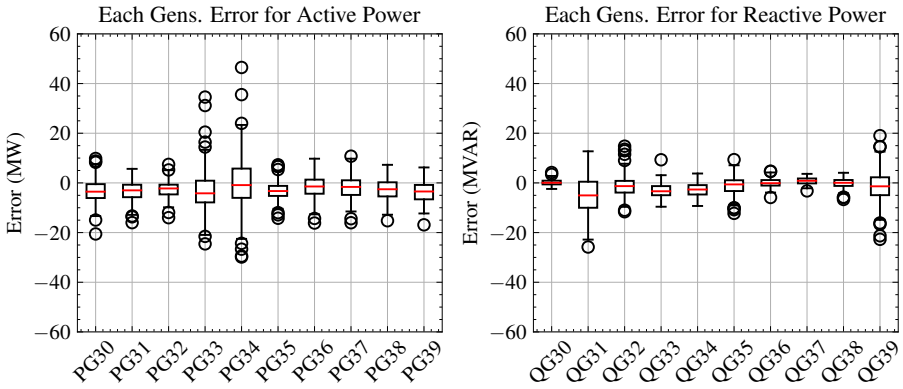


Figure 21. CNN-DFFNN (a)-based boxplot error distribution for all generators.

The scatter plots in Figures 22–24 illustrate the predicted and actual values for active and reactive power from the DFFNN model, the CNN model, and the proposed model. In Figure 22, the data points marked in orange, representing active power, are widely distributed from the line of perfect prediction along the blue line. The figure also shows that the model predictions for active power are less closely clustered around the actual values, which implies that the first model had lower accuracy. For reactive power in the figure, the points are closer to the line. However, there is still a noticeable spread from the blue line circles.

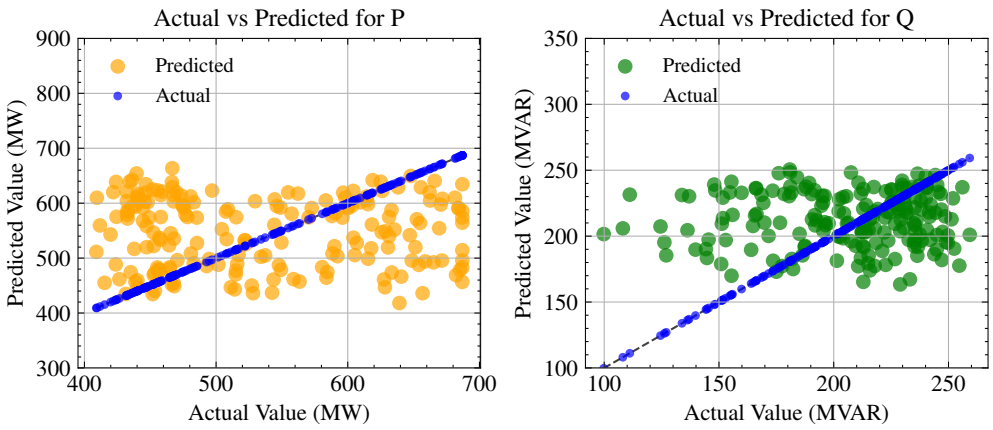


Figure 22. DFFNN-based prediction and actual results.

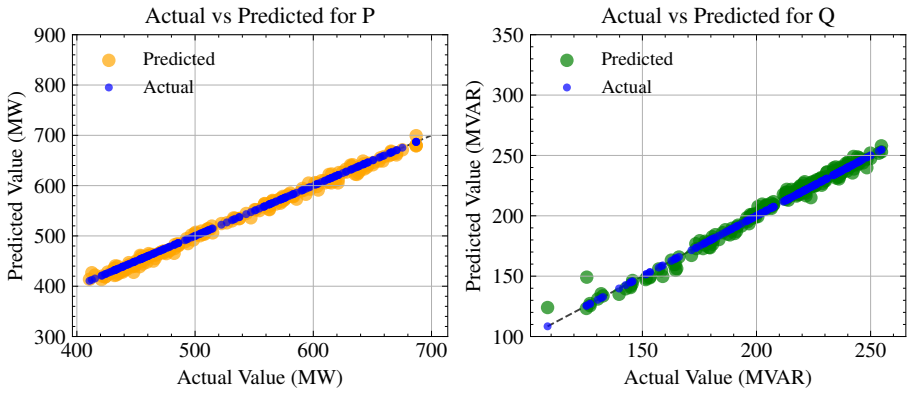


Figure 23. CNN-based prediction and actual results.

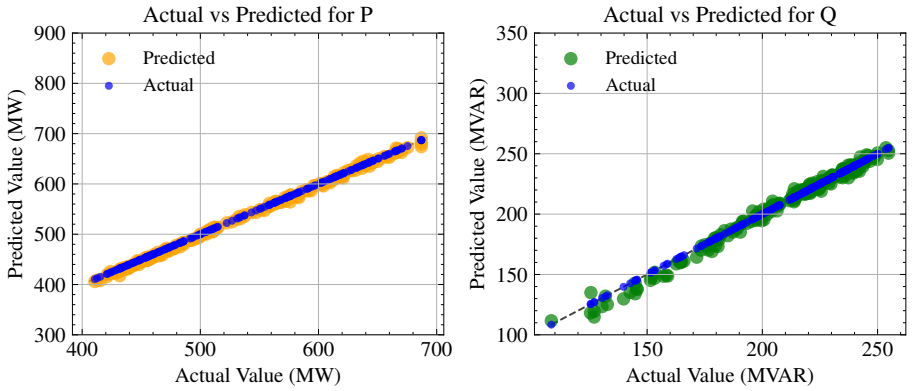


Figure 24. CNN-DFFNN (a)-based prediction and actual results.

In the second and third plots for Figures 23 and 24, the proposed model and the CNN model predictions for both active and reactive power are more closely aligned with the actual values. In addition, the range of predicted values for P and Q does not extend as far as in the first plot, demonstrating that the proposed model and the CNN are more consistent and reliable for prediction. Therefore, the third plot illustrates a superior model accuracy improvement between predicted and actual values, implying that the proposed model has a better fit and predictive performance.

Section Key Findings

- The proposed model still outperformed on validation datasets.
- Performance trends of other DL models remained consistent with those observed in case studies Sections 5.1.1 and 5.2.1.
- When additional input features were included, the model accuracy increased for generator outputs, leading to more precise predictions.
- Model efficiency and computation costs remained consistent when evaluating the case study results in Section 5.2.1.
- The proposed model enhanced predictive accuracy and robustness, significantly outperforming the DFFNN model and slightly outperforming the CNN model.

6. Conclusions

This study addresses the ACOPF problem using several data-driven OPF prediction approaches. The load data generation process relies on Monte Carlo simulations of load variations in all buses. Additionally, the study utilizes a Weibull distribution for wind speed, integrating the physical wind turbine model into six buses of the IEEE 39 test system to simulate corresponding wind power plants. The problem formulation utilizes MIMO learning workflow. We analyzed three different case studies to demonstrate the discussion and results.

The first case study relies on load active and reactive power to predict each bus's voltages and phase angles. The second case study utilizes load active and reactive power to predict the generator's active and reactive power. The last case study focused on load active and reactive power and each bus's voltages and phase angles to be considered as input and predict the generator's active and reactive power outputs.

Based on the results of this study, the studied combination of the CNN-DFFNN model outperformed the other individual CNN and DFFNN models, as well as hybrid and TL-based methods. The results showed that the CNN model outperformed the predicted bus voltage and phase angle case, and the remaining case studies demonstrated that the performance of the combination CNN-DFFNN model was more outstanding than other models on the validation datasets.

Additionally, the DFFNN model was the most efficient in terms of computational resources. However, the DFFNN model was the poorest model in the study regarding model accuracy. Despite the proposed model being the least efficient, it delivered more accurate and robust results, showcasing its effectiveness in prediction tasks. The CNN model found a middle ground, performing very closely to the proposed model regarding accuracy and robustness. At the same time, its computational efficiency was moderate, positioned between the DFFNN model and the proposed model. The results highlighted a trade-off between efficiency and predictive performance across the different models.

It is worth mentioning that while the study was conducted on and tackled the need for more accurate forecasts in DL-supported ACOPF, issues related to ensuring the feasibility of the resulting predictions, finding global optimality approaches, and varying topology for the end-to-end prediction remain for future study. Future studies will expand this research to larger electric network models and explore probabilistic and time-series DL approaches. We intend to expand our methodology by incorporating Bayesian ML and advanced data preprocessing approaches.

In addition, we plan to evaluate the capabilities of more sophisticated neural network architectures, such as Residual Networks (ResNet) and Visual Geometry Group Networks (VGG), for their potential to further enhance the accuracy and robustness of our models. In addition, further stages of the study will investigate and incorporate spatial and temporal DL methods while considering atmospheric weather conditions in specific locations in the Connecticut and New England regions.

Author Contributions: Conceptualization, A.U. and M.P.; methodology, A.U. and M.P.; software, A.U.; validation, A.U.; formal analysis, A.U.; investigation, A.U.; resources, A.U.; data curation, A.U.; writing—original draft preparation, A.U.; writing—review and editing, A.U. and M.P.; visualization, A.U.; supervision, M.P.; project administration, M.P.; funding acquisition, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Eversource Energy Center under the project AG 200208.

Data Availability Statement: Data are contained within the article.

Acknowledgments: A.U. and M.P. sincerely acknowledge and thank the support of Eversource, Inc. in carrying out this research.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

OPF	optimal power flow
ED	economic dispatch
DCOPF	direct currect optimal power flow
ACOPF	alternating currect optimal power flow
VM	voltage magnitude

VA	voltage angle
SCOPF	security constrained optimal power flow
DL	deep learning
NN	neural network
DNN	deep neural network
CNN	convolution neural network
DFFNN	deep feed-forward neural network
TL	transfer learning
CNN-DFFNN	combined convolution and deep feed-forward neural network
RMSE	root mean squared error
MAE	mean absolute error
MIMO	multiple-input multiple-output
RE	renewable energy
PDF	probability density function
WS	wind speed
WT	wind turbine
WP	wind power

Appendix A

Appendix A.1

Appendix A.1, Figure A1 through Figure A4, presents a comparative analysis of the predictive accuracy of DFFNN and the proposed model for VM and VA for Section 5.1 in the results and discussion section across all buses in the power system.

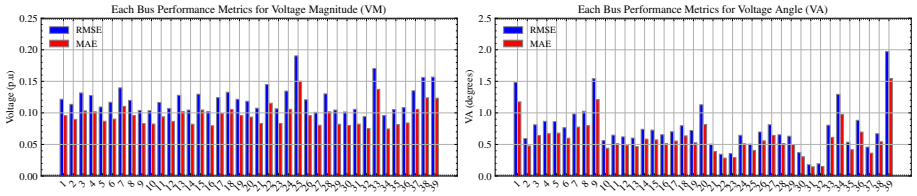


Figure A1. DFFNN-based RMSE and MAE error for all buses in the system.

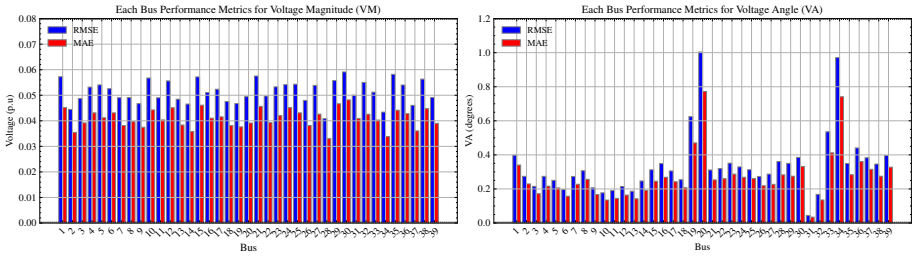


Figure A2. CNN-DFFNN (a)-based RMSE and MAE error for all buses in the system.

Figure A1 shows the DFFNN prediction of VM and VA errors with RMSE and MAE metrics on the left and right subplots, respectively. Figure A2 demonstrates that the proposed combined CNN-DFFNN model improves the prediction accuracy. Its lower RMSE and MAE in VM and VA predictions suggest that the proposed combined model enhanced predictive capabilities over the standalone DFFNN model.

Figures A3 and A4 represent error distribution boxplots for the DFFNN and CNN-DFFNN models, respectively. Figure A3 shows the spread of VM and VA prediction errors in the DFFNN model, with medians highlighted in red. Figure A4 shows that the proposed combined CNN-DFFNN model has diminished error ranges, indicating improved prediction consistency.

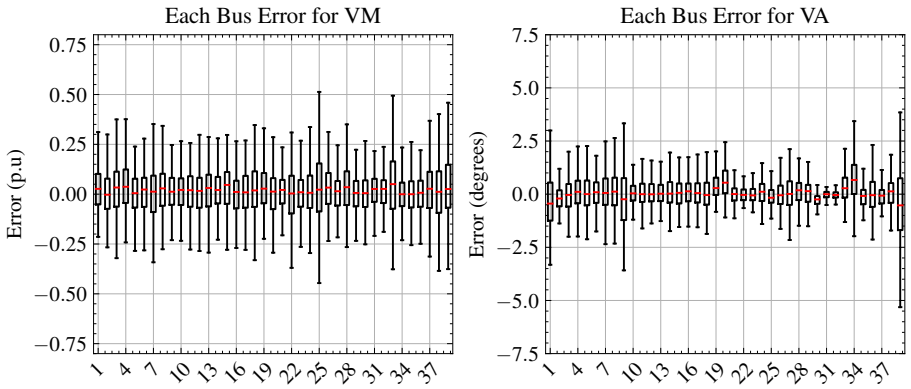


Figure A3. DFFNN-based boxplot error distributions for all buses in the system.

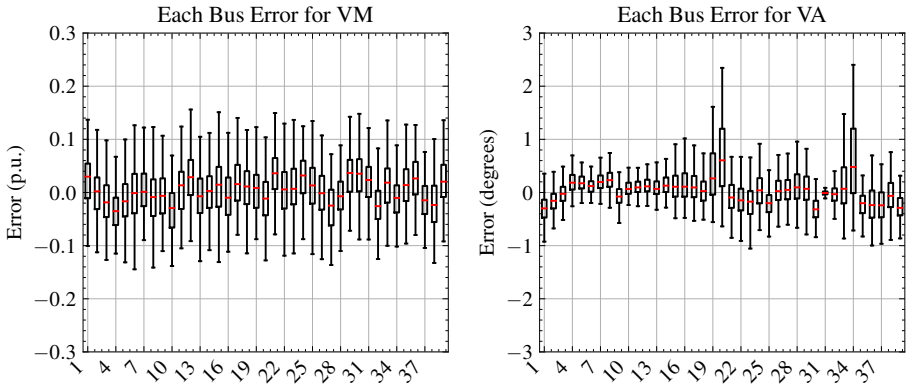


Figure A4. CNN-DFFNN (a)-based boxplot error distributions for all buses in the system.

Appendix A.2

Figure A5 through Figure A10 in Appendix A.2 present a comparative analysis of the predictive RMSE and MAE error of the DFFNN model and the proposed model for generators' active and reactive power for Section 5.2 in the results and discussion section across all generators in the power system. Figures A6, A8 and A10 in Appendix A.2 depict the performance of the proposed CNN-DFFNN model for prediction and actual results. The figures for the proposed model are consistent with earlier discussions, and these results further improve the proposed model, resulting in more accurate predictive ability compared to other DL methods throughout the study. The accuracy of the CNN-DFFNN model underscores its effectiveness in capturing the complexities of the dataset and providing reliable predictions.

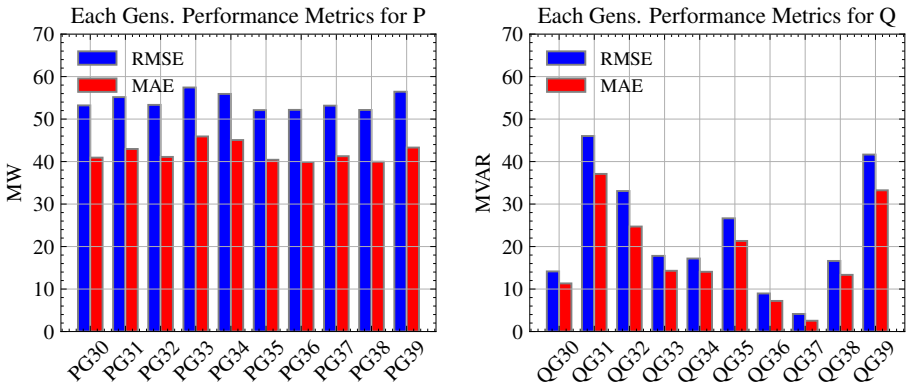


Figure A5. DFFNN-based RMSE and MAE error for all generators.

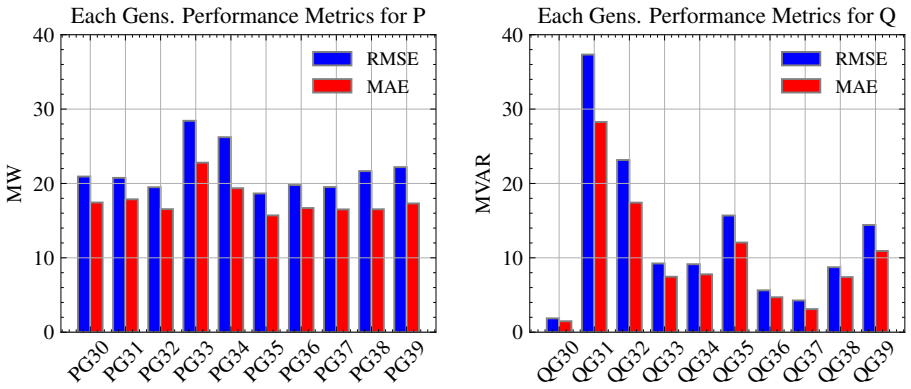


Figure A6. CNN-DFFNN (a)-based RMSE and MAE error for all generators.

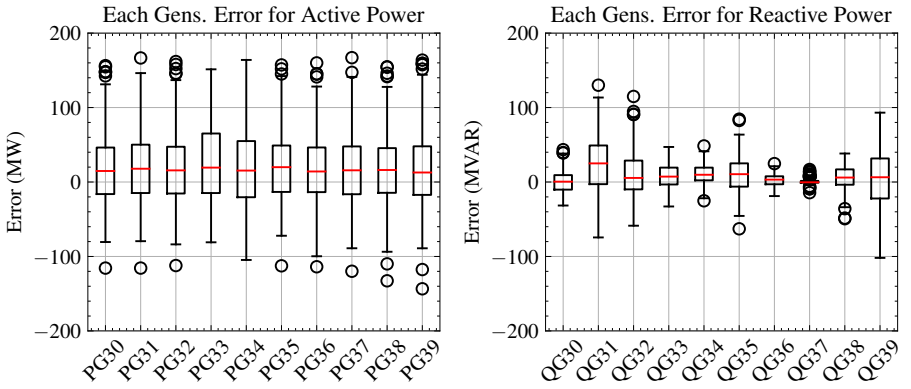


Figure A7. DFFNN-based boxplot error distribution for all generators.

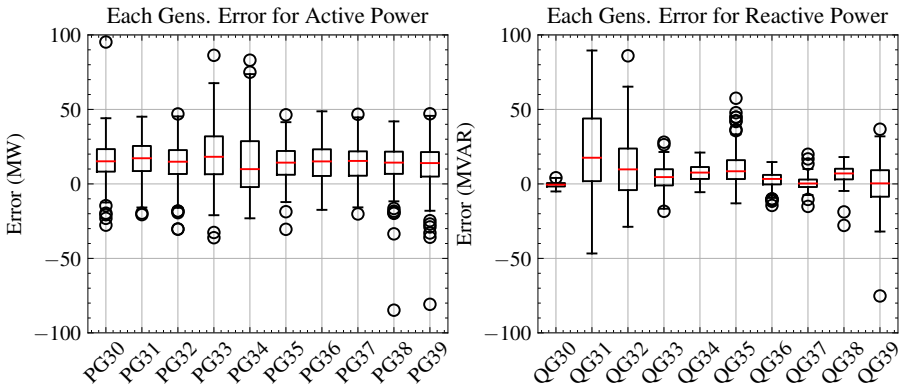


Figure A8. CNN-DFFNN (a)-based boxplot error distribution for all generators.

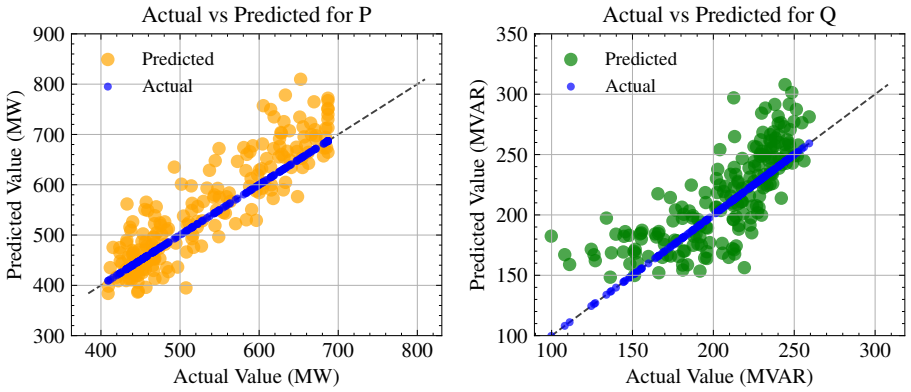


Figure A9. DFFNN-based prediction and actual results.

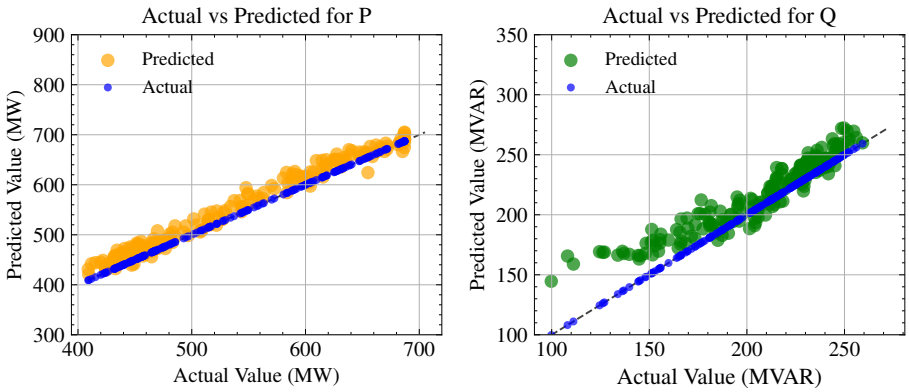


Figure A10. CNN-DFFNN (a)-based prediction and actual results.

References

1. Zhong, J. *Power System Economic and Market Operations*; CRC Press: Boca Raton, FL, USA, 2018.
2. Carpentier, J. Contribution to the economic dispatch problem. *Bull. Soc. Fr. Electr.* **1962**, *3*, 431–447.
3. Cain, M.B.; O'neill, R.P.; Castillo, A. History of optimal power flow and formulations. *Fed. Energy Regul. Comm.* **2012**, *1*, 1–36.
4. Dos Santos, T.N.; Diniz, A.L. A dynamic piecewise linear model for DC transmission losses in optimal scheduling problems. *IEEE Trans. Power Syst.* **2010**, *26*, 508–519. [CrossRef]
5. Low, S.H. Convex relaxation of optimal power flow—Part I: Formulations and equivalence. *IEEE Trans. Control Netw. Syst.* **2014**, *1*, 15–27. [CrossRef]
6. Manshadi, S.D.; Liu, G.; Khodayar, M.E.; Wang, J.; Dai, R. A convex relaxation approach for power flow problem. *J. Mod. Power Syst. Clean Energy* **2019**, *7*, 1399–1410. [CrossRef]
7. Overbye, T.J.; Cheng, X.; Sun, Y. A comparison of the AC and DC power flow models for LMP calculations. In Proceedings of the 37th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 5–8 January 2004; IEEE: Piscataway, NJ, USA, 2004; p. 9.
8. Sun, Y.; Fan, X.; Huang, Q.; Li, X.; Huang, R.; Yin, T.; Lin, G. Local feature sufficiency exploration for predicting security-constrained generation dispatch in multi-area power systems. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1283–1289.
9. Navidi, T.; Bhooshan, S.; Garg, A. *Predicting Solutions to the Optimal Power Flow Problem*; Project Report; Stanford University: Stanford, CA, USA, 2016. Available online: http://cs229.stanford.edu/proj2016/report/NavidiBhooshanGarg-PredictingSolutionstotheOptimalPowerFlowProblem-report.pdf (accessed on 24 January 2024).
10. Zamzam, A.S.; Baker, K. Learning optimal solutions for extremely fast AC optimal power flow. In Proceedings of the 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Tempe, AZ, USA, 11–13 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
11. Zhao, T.; Pan, X.; Chen, M.; Venzke, A.; Low, S.H. DeepOPF+: A deep neural network approach for DC optimal power flow for ensuring feasibility. In Proceedings of the 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Tempe, AZ, USA, 11–13 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
12. Pan, X.; Zhao, T.; Chen, M.; Zhang, S. Deepopf: A deep neural network approach for security-constrained dc optimal power flow. *IEEE Trans. Power Syst.* **2020**, *36*, 1725–1735. [CrossRef]
13. Misra, S.; Roald, L.; Ng, Y. Learning for constrained optimization: Identifying optimal active constraint sets. *INFORMS J. Comput.* **2022**, *34*, 463–480. [CrossRef]
14. Deka, D.; Misra, S. Learning for DC-OPF: Classifying active sets using neural nets. In Proceedings of the 2019 IEEE Milan PowerTech, Milan, Italy, 23–27 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
15. Baker, K.; Bernstein, A. Joint chance constraints in AC optimal power flow: Improving bounds through learning. *IEEE Trans. Smart Grid* **2019**, *10*, 6376–6385. [CrossRef]
16. Zhang, Y.; Cui, H.; Liu, J.; Qiu, F.; Hong, T.; Yao, R.; Li, F. Encoding frequency constraints in preventive unit commitment using deep learning with region-of-interest active sampling. *IEEE Trans. Power Syst.* **2021**, *37*, 1942–1955. [CrossRef]
17. Liu, J.; Yang, Z.; Zhao, J.; Yu, J.; Tan, B.; Li, W. Explicit data-driven small-signal stability constrained optimal power flow. *IEEE Trans. Power Syst.* **2021**, *37*, 3726–3737. [CrossRef]
18. Roald, L.; Oldewurtel, F.; Van Parys, B.; Andersson, G. Security constrained optimal power flow with distributionally robust chance constraints. *arXiv* **2015**, arXiv:1508.06061.
19. Mühlpfordt, T.; Faulwasser, T.; Hagenmeyer, V. A generalized framework for chance-constrained optimal power flow. *Sustain. Energy Grids Netw.* **2018**, *16*, 231–242. [CrossRef]

20. Lei, X.; Yang, Z.; Yu, J.; Zhao, J.; Gao, Q.; Yu, H. Data-driven optimal power flow: A physics-informed machine learning approach. *IEEE Trans. Power Syst.* **2020**, *36*, 346–354. [CrossRef]
21. Nellikkath, R.; Chatzivasileiadis, S. Physics-informed neural networks for minimising worst-case violations in dc optimal power flow. In Proceedings of the 2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Aachen, Germany, 25–28 October 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 419–424.
22. Gao, M.; Yu, J.; Yang, Z.; Zhao, J. A Physics-Guided Graph Convolution Neural Network for Optimal Power Flow. *IEEE Trans. Power Syst.* **2023**. [CrossRef]
23. Huang, B.; Wang, J. Applications of physics-informed neural networks in power systems-a review. *IEEE Trans. Power Syst.* **2022**, *38*, 572–588. [CrossRef]
24. Misyris, G.S.; Venzke, A.; Chatzivasileiadis, S. Physics-informed neural networks for power systems. In Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2–6 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
25. Baker, K. Emulating AC OPF Solvers With Neural Networks. *IEEE Trans. Power Syst.* **2022**, *37*, 4950–4953. [CrossRef]
26. Chatzos, M.; Mak, T.W.; Van Hentenryck, P. Spatial network decomposition for fast and scalable AC-OPF learning. *IEEE Trans. Power Syst.* **2021**, *37*, 2601–2612. [CrossRef]
27. Chen, L.; Tate, J.E. Hot-starting the AC power flow with convolutional neural networks. *arXiv* **2020**, arXiv:2004.09342.
28. Dong, W.; Xie, Z.; Kestor, G.; Li, D. Smart-PGSim: Using neural network to accelerate AC-OPF power grid simulation. In Proceedings of the SC20: International Conference for High Performance Computing, Networking, Storage and Analysis, Atlanta, GA, USA, 9–19 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–15.
29. Cao, Y.; Zhao, H.; Liang, G.; Zhao, J.; Liao, H.; Yang, C. Fast and explainable warm-start point learning for AC Optimal Power Flow using decision tree. *Int. J. Electr. Power Energy Syst.* **2023**, *153*, 109369. [CrossRef]
30. Forootan, M.M.; Larki, I.; Zahedi, R.; Ahmadi, A. Machine learning and deep learning in energy systems: A review. *Sustainability* **2022**, *14*, 4832. [CrossRef]
31. Falconer, T.; Mones, L. Leveraging power grid topology in machine learning assisted optimal power flow. *IEEE Trans. Power Syst.* **2022**, *38*, 2234–2246. [CrossRef]
32. Jia, Y.; Bai, X.; Zheng, L.; Weng, Z.; Li, Y. ConvOPF-DOP: A data-driven method for solving AC-OPF based on CNN considering different operation patterns. *IEEE Trans. Power Syst.* **2022**, *38*, 853–860. [CrossRef]
33. Rahman, J.; Feng, C.; Zhang, J. A learning-augmented approach for AC optimal power flow. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106908. [CrossRef]
34. Pareek, P.; Nguyen, H.D. Gaussian process learning-based probabilistic optimal power flow. *IEEE Trans. Power Syst.* **2020**, *36*, 541–544. [CrossRef]
35. Owerko, D.; Gama, F.; Ribeiro, A. Optimal power flow using graph neural networks. In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 5930–5934.
36. Liu, S.; Wu, C.; Zhu, H. Topology-aware graph neural networks for learning feasible and adaptive AC-OPF solutions. *IEEE Trans. Power Syst.* **2022**, *38*, 5660–5670. [CrossRef]
37. Velloso, A.; Van Hentenryck, P. Combining deep learning and optimization for preventive security-constrained DC optimal power flow. *IEEE Trans. Power Syst.* **2021**, *36*, 3618–3628. [CrossRef]
38. Park, S.; Chen, W.; Mak, T.W.; Van Hentenryck, P. Compact optimization learning for AC optimal power flow. *arXiv* **2023**, arXiv:2301.08840.
39. Chen, Y.; Lakshminarayana, S.; Maple, C.; Poor, H.V. A meta-learning approach to the optimal power flow problem under topology reconfigurations. *IEEE Open Access J. Power Energy* **2022**, *9*, 109–120. [CrossRef]
40. Liu, J.; Liu, Y.; Qiu, G.; Shao, X. Learning-aided optimal power flow based fast total transfer capability calculation. *Energies* **2022**, *15*, 1320. [CrossRef]
41. Aksan, F.; Li, Y.; Suresh, V.; Janik, P. CNN-LSTM vs. LSTM-CNN to Predict Power Flow Direction: A Case Study of the High-Voltage Subnet of Northeast Germany. *Sensors* **2023**, *23*, 901. [CrossRef]
42. Unlu, A.; Peña, M.; Wang, Z. Comparison of the Combined Deep Learning Methods for Load Forecasting. In Proceedings of the 2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 16–19 January 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
43. Yan, Z.; Xu, Y. A hybrid data-driven method for fast solution of security-constrained optimal power flow. *IEEE Trans. Power Syst.* **2022**, *37*, 4365–4374. [CrossRef]
44. Athay, T.; Podmore, R.; Virmani, S. A practical method for the direct analysis of transient stability. *IEEE Trans. Power Appar. Syst.* **1979**, *2*, 573–584. [CrossRef]
45. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* **2010**, *26*, 12–19. [CrossRef]
46. Bie, Z.; Li, G.; Liu, H.; Wang, X.; Wang, X. Studies on voltage fluctuation in the integration of wind power plants using probabilistic load flow. In Proceedings of the 2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, PA, USA, 20–24 July 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 1–7.
47. Available online: https://www.vestas.com/en/ (accessed on 24 January 2024).

48. Available online: https://ieee-dataport.org/documents/data-driven-optimal-power-flow-opf (accessed on 24 January 2024).
49. Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Deep Learning for Time Series Classification: A Review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [CrossRef]
50. Aggarwal, C.C. *Neural Networks and Deep Learning*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.