ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.1648

A Model-Free Approach for Assessing Renewable Generation Penetration in Power Systems

Oizis Aksunamun Melgar Rodríguez[1], Isaac Fernando Hernández Urbina[2], Jonathan Muñoz Tabora[3], Ozy D. Melgar Dominguez[4]
[1]Universidad Nacional Autónoma de Honduras, *oizis.melgar@unah.hn*
[2]Universidad Nacional Autónoma de Honduras, *isaac_hernandez@unah.hn*
[3]Universidad Nacional Autónoma de Honduras, *Jonathan.munoz@unah.edu.hn*
[4]Universidad Nacional Autónoma de Honduras, *ozzydamedo@gmail.com*

***Abstract*—Integrating variable renewable energy (VRE) generation into power systems presents significant challenges due to its inherent variability and uncertainty. Traditionally, the impact of this generation is assessed using power flow analysis tools. However, this method requires modeling the characteristics of the electrical grid, and its solution can entail excessive computational effort for real-time applications. In this regard, methods based on artificial intelligence have emerged as an alternative for operating electrical power systems. Therefore, this paper proposes a strategy based on a multilayer perceptron (MLP) to assess the impact of VRE generation, such as wind and solar, through a purely data-driven approach. To evaluate the proposed strategy, numerous operating scenarios were simulated on the IEEE 30-bus test system using a power flow tool, and the resulting solutions were used to train the MLP. Once the MLP was trained, the system voltage profiles could be estimated with a mean squared error (MSE) of approximately 1.47 $\times 10^{-5}$ and a coefficient of determination (R-squared) of 0.987 on the test set. These findings provide a scalable and adaptable tool for system operators to effectively manage the increasing penetration of VRE generation.**

***Index Terms*—Artificial intelligence, multilayer perceptrons, power system analysis, variable renewable energy generation.**

I. Introduction

In recent years, the planning and operation of electric power systems have been challenged by the increasing penetration of variable renewable energy (VRE) generation, such as solar and wind power. The variability and uncertainty of power production from these sources pose significant challenges to develop more robust analytical methods [1]–[3]. Techniques for solving electrical system planning and operation problems are based on detailed models of electrical grids. Solutions can be found using exact or iterative methods, such as the Newton-Raphson or Gauss-Seidel methods. One disadvantage of methods that require an exact model of the electrical grid is computational effort, which can be impacted by the size of the electrical grid and by the number of scenarios with high penetration of VRE generation [4]

Unlike prior neural-network-based power flow approximations that require retraining for topology changes, our contribution is a purely data-driven MLP framework that does not assume any grid model and whose training data can be sourced entirely from smart meters in real-world deployment. This key distinction enables model-free operation without knowledge of line parameters or breaker statuses.

In this context, approaches based on artificial intelligence (AI) have emerged as a promising alternative. Among these, artificial neural networks (ANNs) are of particular interest due to their ability to learn and predict the operating state of a system directly from data, without depending on a detailed model of the electrical grid [5]. Several studies have examined the use of AI to predict the real-time power system state estimation [6] and the management of the variability of the VRE generation [7]. The authors in [8] explore the robustness of a deep learning algorithm to solve the optimal power flow problem in hybrid renewable-based energy systems. On the other hand, [9] propose using AI-based tools to forecast wind power generation. However, the accuracy of the results is limited by the lack of training data. Similarly, in Bassi et al. [10], a model-free methodology is presented for carrying out voltage calculations using neural networks with real smart meter data. Another application of AI-based tools, the authors in [11] present a deep learning model to predict power flow solutions for three-phase unbalanced power distribution grids. However, the accuracy of these approaches is limited by the lack of training data.

Given the need for developing innovative tools, this work explores the application of a multilayer perceptron (MLP) to assess the penetration of VRE generation into power systems. The MLP has been developed to estimate the voltage profiles of an electric power system under different scenarios of generation and electric demand. The proposed strategy is presented as a flexible and scalable tool that could be useful to electric system operators in managing the integration of large penetration of VRE generation.

II. Problem description

Unlike dispatchable power plants, VRE generation is characterized by its intermittent nature and high variability, which leads to the need to implement strategies to deal with uncertainties [12]. High penetration of VRE generation introduces significant operational challenges due to their inherent variability and forecasting uncertainty. These characteristics can compromise grid stability by causing fluctuations in voltage

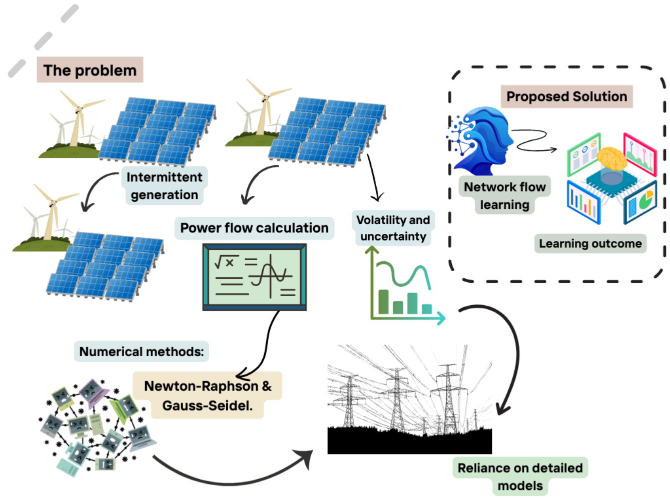

Fig. 1. Framework for the analysis of power systems and proposed structure

profiles, reducing system inertia, and complicating real-time balancing between generation and load demand.

To assess the power system operation with high penetration levels of VRE sources, system operators commonly implement classical techniques based on numerical methods for power flow analysis, such as Newton-Raphson, Gauss-Seidel, among others. These methods are mathematically robust and effective ways to obtain the solution of the problem. However, their computational effort and convergence may be affected by the size of the system and the number of elements connected to the electrical grid under study. In this regard, with the rise of artificial intelligence, new opportunities arise to automate power systems, providing innovative tools applied for decision-making processes [5]. In order to illustrate the problem to be addressed in this work, Fig. 1 is presented. From this figure we can observed that the steady-state operation of the power system is calculated using numerical methods. These numerical methods require a detailed network model. In this context, the proposal in this work to estimate the impact of connecting high levels of VRE generation consists of a model-free approach to the electrical grid. Therefore, generation, demand, and electrical system profile data are required to train the MLP neural network.

III. Methodology

This work proposes an approach based on MLP to address the challenges of estimating the state of an electrical grid with high penetrations of VRE generation. Following a comparative analysis of the literature, this architecture was selected because it was identified as offering a favorable balance between predictive capacity and computational efficiency for medium-scale power systems. Therefore, the proposed methodology focuses on generating training scenarios, architecting and configuring the model, and evaluating its performance.

A. Generation of stochastic data

A large number of works have been developed to model and capture the uncertainties associated with VRE sources [13]–[15]. In this regard, the main objective of this phase is to generate an extensive and diverse dataset that represents a wide range of operating scenarios, both for electricity demand and renewable power production. To model variability, a probability density function (PDF) is used for the system's electricity demand. This PDF include a uniform distribution ranging from 45% to 115% of the nominal value. To simulate the intermittent nature of wind and solar renewable generation, a beta PDF was modeled with parameters $\alpha = 2$ and $\beta = 5$.

B. Neural network-based emulator

The proposed strategy consists of developing a nonparametric MLP regression model to emulate the behavior of the system, when high penetration levels of VRE generation are incorporated Iniatilly, the MLP model requires sufficient information to be train itself and estimate the system's behavior. This training process follows a supervised learning approach, utilizing as input data the active and reactive power associated with power generation and system demand (P_g, Q_g, P_l, Q_l). Meanwhile, the output variables are defined as the voltage magnitude levels of each bus in the power system (V_j).

C. Model architecture and training

As illustrated in Fig. 2, the model is configured with three hidden layers, each containing 128 neurons. This specific depth and width were determined through a trial-and-error process, ensuring a balance between predictive capacity and computational efficiency. The Rectified Linear Unit (ReLU) activation function was used to model the complex nonlinearities of the electrical system. ReLU was specifically selected to mitigate the vanishing gradient problem, which is common in deep architectures, thereby accelerating training and improving the model's ability to map the non-linear relationship between power injections and voltage magnitudes. The model is defined to be trained under a supervised learning approach with the active and reactive powers of the generation and system demand as input variables. This training process allows the model to predict the voltage profiles in the system's buses.

To ensure optimal performance, the data information should be preprocessed using a StandardScaler, and the model weights iteratively adjusted to minimize the loss function in terms of the mean squared error (MSE) and coefficient of determination (R^2). It is important to mention that, this kind of architecture allows to obtain suitable estimations without the need for modelling a detailed network topology [16].

IV. Data information and validation

This section provides the key data required to validate the proposed strategy, including a detailed description of the tested power system, the data-generating process used for the training, and the data validation to estimate the correlation between input and output variables.

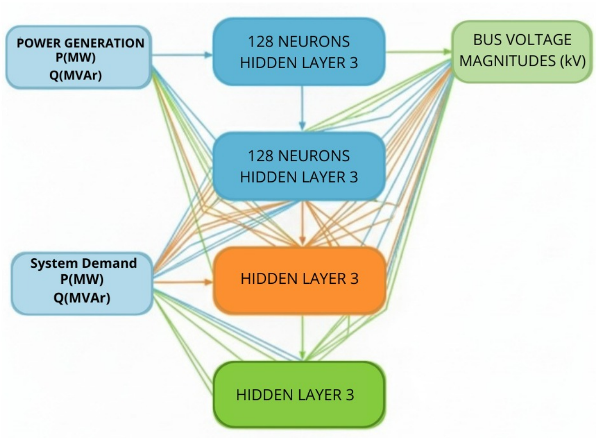

Fig. 2. Representation of the implemented MLP architecture

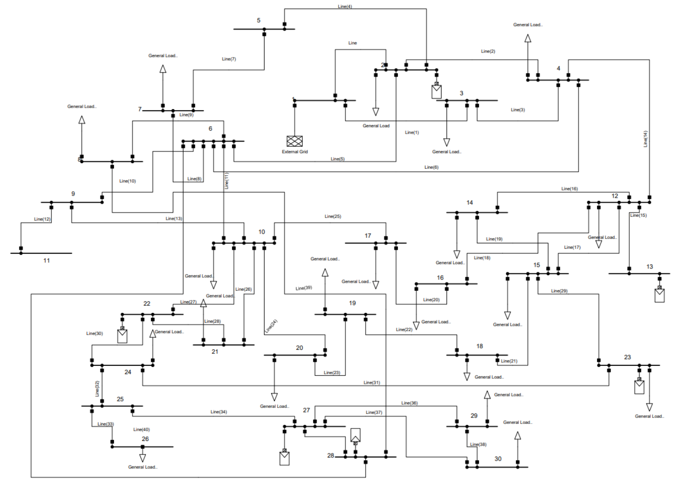

Fig. 3. Diagram of the IEEE 30 bus-system

A. Description of the power system

To validate the developed strategy, the IEEE 30-bus system is studied, shown in Fig. 3. This system consists of 30 nodes interconnected by 41 branches distributed across 37 transmission lines and four strategically located transformers. In addition, six generating units located at buses 1 (slack), 2, 5, 8, 11, and 13. These units supply active and reactive power to the system. The total active power demand is 283.4 MW and the reactive power demand is 126.2 MVar. The system buses are classified into three categories: a slack bus (Bus 1) for angular reference; five PV buses (Buses 2, 5, 8, 11, and 13), which control voltage while injecting active power; and 24 PQ buses, which represent passive loads. Finally, the system incorporates shunt elements in ten critical buses for voltage control.

The network operates at three nominal voltage levels (132 kV, 33 kV, and 11 kV) and is interconnected by transformers in branches 4-12, 6-9, 6-10, and 27-28. Topologically, the network is organized into three functional areas: Area 1, defined by buses 1-8, houses the primary generation; Area 2, with buses 9–14 and 30, defines the transmission core; and, Area 3, with buses 15–29, houses distributed loads.

B. Data generation for training

To implement the developed strategy, a total of 10,000 operating scenarios were generated using the PDFs discussed in Section III-A. These scenarios represents the stochastic data behavior to model uncertainties in electric demand and VRE generation. To obtain the information required to train the MLP, with the 10,000 generating scenarios, the Pandapower library [17] is used to AC power flow simulations. The solution of each simulation provides data such as voltage magnitude profiles. In order to ensure an impartial evaluation and mitigate overfitting, the dataset was divided into a training subset (90%) and a validation subset (10%). Therefore, all variables were normalized using a standardization transformation, which is a pre-processing step that promotes efficient convergence in the training process.

As a complement to data generation, it is verified that the training data is consistent with the defined stochastic approach. Fig. 4 and 5 shows the histograms of demand factors and VRE generation. Fig. 4 shows that electric demand exhibits a uniform probability distribution, which guarantees exploration of demand conditions. On the other hand, Fig. 5 illustrates the PDF adopted to captures uncertainties in VRE generation and represents a beta PDF with parameter $\alpha = 2$ and $\beta = 5$.

C. Physical-statistical correlation

Fig. 6 illustrates the relationship between total active power demand and power generation under 1000 simulated scenarios. It can be observed that the scatter plot shows a strong linear correlation between these two variables. This validation ensures the convergence of the power flow simulations, which in turn guarantees the proper training of the MLP model Therefore, it can be verified that the system's power balance is satisfied according to (1). From this equation, $P^g_{n,s}$ represents the active power generated by the power plant n. In addition, $P^d_{i,s}$ and $P^{loss}_{ij,s}$ represent the active power demand at bus i and the active power losses at circuit ij, respectively. It is worth

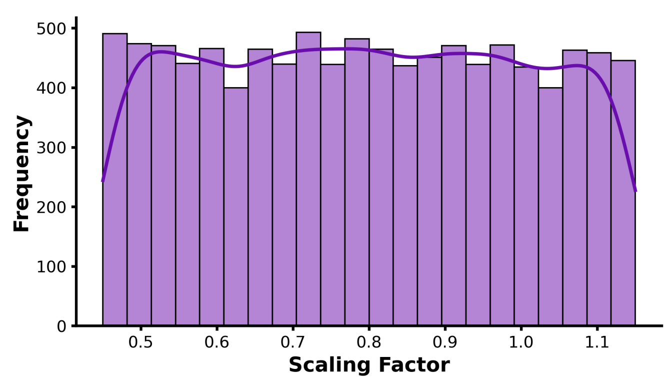

Fig. 4. Distribution of the demand scaling factor (FDt), showing the frequency of load variations and their associated probability density

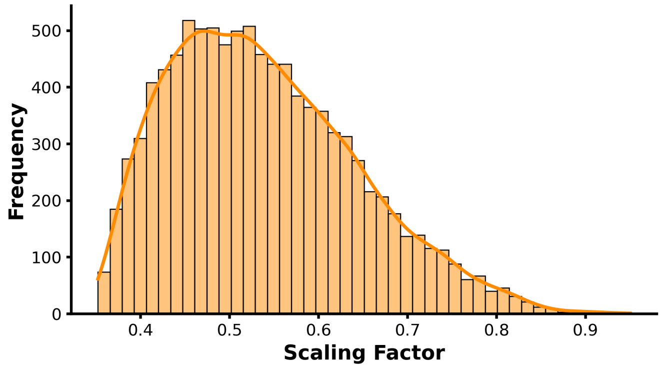

Fig. 5. Distribution of the generation scaling factor (FGt), illustrating the frequency of generation variations and their corresponding probability density

noting that this equation must be satisfied for each scenario s.

$$\sum_{n\in G} P^{g}_{n,s} = \sum_{i\in B} P^{d}_{i,s} + \sum_{ij\in L} P^{loss}_{ij,s}, \qquad \forall s \in S. \quad (1)$$

The vertical deviation of the data points from a line of unit slope ($y = x$) quantitatively represents the active power losses (P_{loss}). This is inherent to the transmission system in each scenario. The close clustering and low variance of the points validate the physical consistency of each power flow solution and confirm that the power balance has been met.

V. Numerical Results

In the previous section, the information necessary to train the MLP was presented and validated. Thereby, this section shows numerical results obtained from simulate different operating conditions in the 30-bus power system by varying the number of scenarios from 10, 100, 1000 and 10000, which are the result of running the power flow. These conditions aim to model the variability and uncertainty of electricity demand and VRE generation, such as solar photovoltaics and wind.

The mean squared error (MSE) was used to evaluate the predictive accuracy of the proposed MLP approach. The results are summarized in Table I. The data reveals a consistent improvement in model precision as the training set expands; specifically, the MSE is reduced by two orders of magnitude when moving from 10 to 10,000 scenarios. This trend highlights the high sensitivity of the MLP to the volume of data, transitioning from a coarse estimation to a high-fidelity mapping of the system's states. Achieving such a low error rate in the most extensive case confirms that the network has effectively captured the underlying non-linear physics of the 30-bus system, reaching a high level of variance explanation without the need for traditional power flow equations

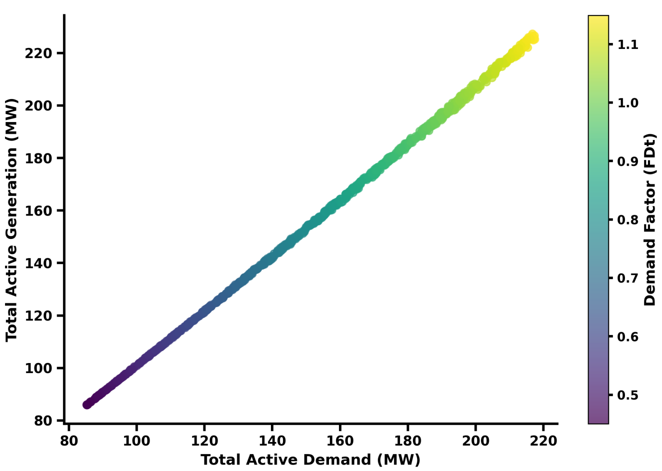

Fig. 6. Verification of compliance with the linear relationship and power balance

TABLE I
MSE as a Function of the Number of Training Scenarios

Number of scenarios	Mean Squared Error (MSE)
10	1.37×10^{-4}
100	4.46×10^{-5}
1,000	1.27×10^{-5}
10,000	2.06×10^{-6}

It is observed that with 10 scenarios, lacking sufficient samples to learn the non-linear relationships of the electrical system, the model shows a low performance, obtaining a higher mean squared error (MSE) of the analysis, of 1.37×10^{-4}. This result is visually validated in Fig. 7, which illustrates the dispersion of the predicted voltage profiles versus the actual values for this limited dataset.

On the other hand, increasing the training dataset to 100 scenarios results in an MSE of 4.46×10^{-5}, representing a reduction of approximately 70% compared to the MSE obtained when the model was trained with only 10 scenarios. By analyzing the results obtained from that increase, it is possible to observe that the model begins to capture characteristic patterns of the electrical system. This improvement is depicted in Fig. 8, where the alignment of the voltage magnitude profiles begins to tighten around the identity line, reflecting a more consistent prediction across the 100 power generation and demand scenarios. When the number of training data is increased to 1,000 scenarios, the MLP has a diverse dataset

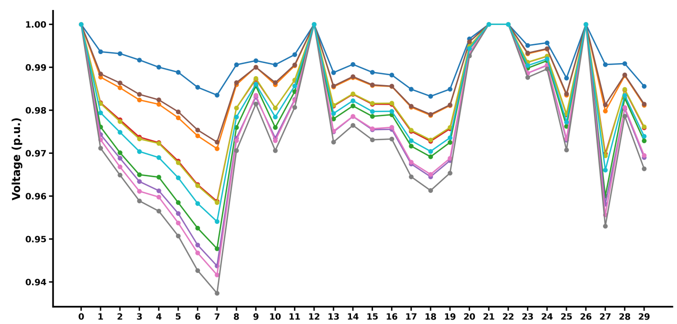

Fig. 7. Predicted voltage profiles for the trained MLP with 10 scenarios.

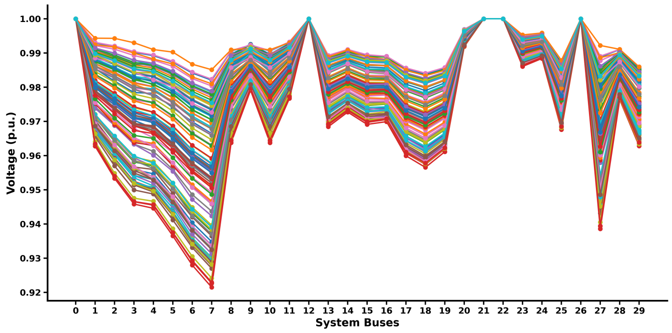

Fig. 8. Prediction results for the model trained with 100 scenarios.

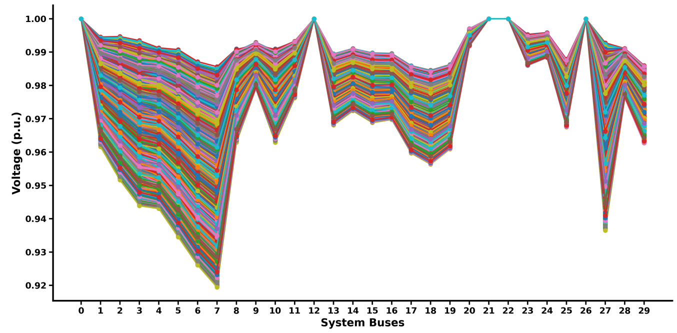

Fig. 10. Prediction result for the model trained with 10,000 scenarios.

information to robustly simulate the behavior of the power system.

In this case, the MSE is reduced to 1.27×10^{-5}, obtaining a high degree of accuracy. The robustness of the model at this stage is evidenced in Fig. 9, where the predicted voltage profiles closely follow the diagonal, indicating that the MLP has successfully learned the complex, non-linear behavior of the 30-bus system with 1,000 generation and demand scenarios.

By increasing the data tenfold, it becomes evident that the model significantly outperforms previous iterations Notably, when the MLP is trained under diverse operating conditions, it enhances its grasp of the electrical network, achieving an MSE of 2.06×10^{-6}, which represents the best performance among the evaluated cases. This high-fidelity estimation is visualized in Fig. 10, where the overlap between predicted and real values is nearly perfect. The results demonstrate that with a sufficiently large dataset, the MLP model becomes a highly reliable surrogate for traditional power flow analysis under extreme VRE variability.

To complement the analysis, a regression model was fitted to the results, comparing the number of electricity demand and generation scenarios used to train the MLP model and the MSE value. This analysis is illustrated, on a logarithmic scale, in Fig. 11. From this figure, it is possible to observe that the relationship between the number of scenarios and the MSE is approximately linear. An analysis of the results clearly indicates that the MLP's forecasting performance improves substantially as the volume of training data increases.

The results obtained demonstrate that the MLP-based approach effectively captures the non-linear sensitivities of the 30-bus system without requiring a physical model of the network. The significant reduction in MSE as training scenarios increased confirms that the network successfully learned the mapping between power injections and voltage magnitudes. Furthermore, the nearly linear relationship observed in the log-scale learning curve suggests that the model is robust and has not yet reached a saturation point. This implies that the methodology is highly scalable; while this study focused on a 30-bus system, the learning patterns suggest it could be adapted to larger, more complex grids with similar precision. This model-free capability represents a critical advantage for real-time monitoring, where speed and adaptability are paramount.

VI. Conclusions

The increased use of variable renewable energy (VRE) generation has created a need for robust tools that can analyze their impact on power systems. This challenge has led system operators to recast their computational tools for planning and operational studies. In this regard, this paper developed an artificial intelligence-based tool designed to predict the impact of VRE generation on power systems. The tool requires a training process that utilizes active power demand and generation data as input variables, with the voltage profile at

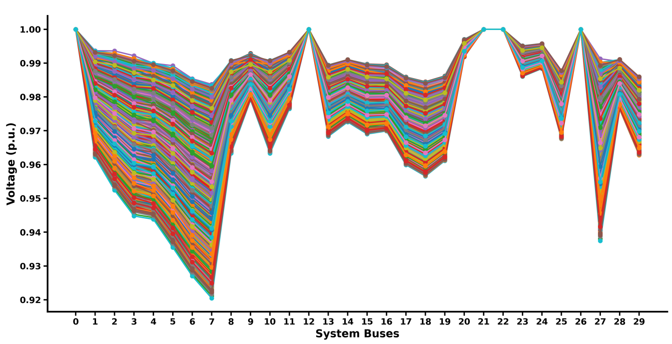

Fig. 9. Prediction result for the model trained with 1,000 scenarios

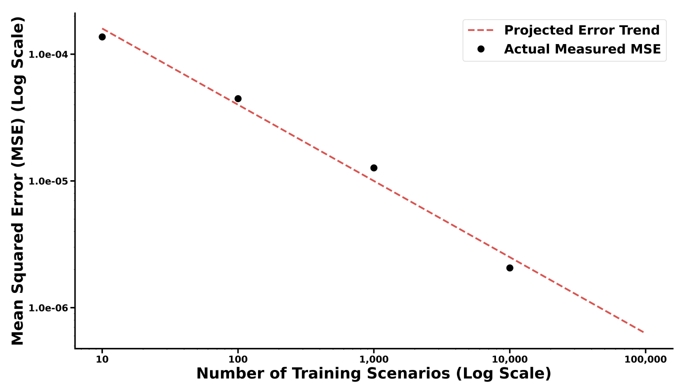

Fig. 11. The learning curve of the model is shown on a logarithmic scale.

each bus serving as the output. For the purposes of this study, the training was conducted using data derived from power flow simulations across various operating scenarios within a test system. However, for real-world applications, the proposed strategy can be trained using historical data collected directly from the system's metering infrastructure.

Although the present work focuses on demonstrating the feasibility of the MLP-based emulator on the IEEE 30-bus system, preliminary comparisons with a linear regression baseline (not shown for brevity) yielded an R^2 of 0.62 versus 0.987 for the proposed MLP, confirming the need for a nonlinear model. Regarding scalability, the purely data-driven nature of the approach implies that computational complexity scales with the number of neurons and training samples, not with system size—suggesting that larger grids would only require proportionally more training data, not a change in architecture. A systematic comparison with traditional power flow methods and validation on larger systems is planned as immediate future work.

The numerical results indicate that the accuracy of the Multilayer Perceptron (MLP) model is closely tied to the richness of the dataset used during training. The numerical results indicate that the accuracy of the Multilayer Perceptron (MLP) model is closely tied to the richness of the dataset. While a sample size of 100 scenarios already shows promising results, the analysis demonstrates that a dataset of 10^4 scenarios achieves a high-precision state with an MSE of 2.06×10^{-6}. This confirms that the proposed model-free strategy can effectively replace traditional power flow equations for steady-state analysis, significantly reducing computational burden. Nevertheless, even with a sample size of just 100 scenarios, the proposed strategy demonstrates a marked reduction in error, yielding promising results.

Future work will focus on validating this strategy using data from a real-world power system and expanding its formulation to support advanced applications such as power dispatch optimization, voltage regulation, and reactive power control.

References

[1] O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, and J. R. S. Mantovani, "Robust short-term electrical distribution network planning considering simultaneous allocation of renewable energy sources and energy storage systems," in *Robust Optimal Planning and Operation of Electrical Energy Systems*. Springer, 2019, pp. 145–175.

[2] O. D. Melgar-Dominguez, R. W. Salas, and J. R. S. Mantovani, "Short-term distribution system planning using a system reduction technique," *IEEE Access*, vol. 9, pp. 153 586–153 598, 2021.

[3] J. M. Tabora, Y. G. Velasquez, D. A. Rivera, D. A. Quijano, and O. D. Melgar-Dominguez, "A multiobjective strategy for generation expansion planning to enable increased penetration of variable renewable energy sources," *IEEE Access*, vol. 12, pp. 187 665–187 675, 2024.

[4] A. A. El-Fergany, "Reviews on load flow methods in electric distribution networks," *Archives of Computational Methods in Engineering*, October 2024. [Online]. Available: https://lens.org/148-457-420-609-961

[5] V. Paucar and M. J. Rider, "Artificial neural networks for solving the power flow problem in electric power systems," *Electric Power Systems Research*, vol. 62, no. 2, pp. 139–144, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779602000305

[6] L. Zhang, G. Wang, and G. B. Giannakis, "Real-time power system state estimation and forecasting via deep unrolled neural networks," *IEEE Transactions on Signal Processing*, vol. 67, no. 15, pp. 4069–4077, 2019.

[7] Y. Li, Y. Ding, S. He, F. Hu, J. Duan, G. Wen, H. Geng, Z. Wu, H. B. Gooi, Y. Zhao, C. Zhang, S. Mei, and Z. Zeng, "Artificial intelligence-based methods for renewable power system operation," *Nature Reviews Electrical Engineering*, vol. 1, no. 3, pp. 163–179, February 2024. [Online]. Available: https://lens.org/103-887-403-407-812

[8] G. Gurumoorthi, S. Senthilkumar, G. Karthikeyan, and F. Alsaif, "A hybrid deep learning approach to solve optimal power flow problem in hybrid renewable energy systems," *Scientific Reports*, vol. 14, no. 1, p. 19377, 2024, accessed 2025-04-04.

[9] P. V. Sireesha and S. Thotakura, "Wind power prediction using optimized mlp-nn machine learning forecasting model," *Electrical Engineering*, vol. 106, no. 6, pp. 7643–7666, May 2024. [Online]. Available: https://lens.org/173-846-400-423-25X

[10] V. Bassi and et al., "Demonstrating electrical model-free voltage calculations with real smart meter data," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, Dubrovnik, Croatia, 2024, pp. 1–4.

[11] D. Tiwari, M. J. Zideh, V. Talreja, V. Verma, S. K. Solanki, and J. Solanki, "Power flow analysis using deep neural networks in three-phase unbalanced smart distribution grids," *IEEE Access*, vol. 12, pp. 29 959–29 970, 2024.

[12] J. Tabora, U. Paixão Júnior, C. Rodrigues, U. Bezerra, M. Tostes, B. Albuquerque, E. Matos, and A. Nascimento, "Hybrid system assessment in on-grid and off-grid conditions: A technical and economical approach," *Energies*, vol. 14, no. 17, p. 5284, 2021.

[13] A. M. L. da Silva, L. A. F. M. Ferreira, and J. P. S. Catalão, "Uncertainty assessment in power systems through interval arithmetic and probabilistic analysis," *International Journal of Electrical Power Energy Systems*, vol. 64, pp. 913–919, Jan 2015.

[14] D. A. Quijano, J. Wang, M. R. Sarker, and A. Padilha-Feltrin, "Stochastic assessment of distributed generation hosting capacity and energy efficiency in active distribution networks," *IET Generation, Transmission & Distribution*, vol. 11, no. 18, pp. 4617–4625, 2017.

[15] O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, and J. R. S. Mantovani, "Adaptive robust short-term planning of electrical distribution systems considering siting and sizing of renewable energy based dg units," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 158–169, 2019.

[16] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer, 2009.

[17] S. Meinecke and M. Braun, "Pandapower—an open-source python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510–6521, 2018.