energies

Article

Adaptive Hosting Capacity Forecasting in Distribution Networks with Distributed Energy Resources

Md Tariqul Islam †, M. Jahangir Hossain *,†, Md. Ahasan Habib and Muhammad Ahsan Zamee *

School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia; mdtariqul.islam@student.uts.edu.au (M.T.I.); mdahasan.habib@student.uts.edu.au (M.A.H.)
* Correspondence: jahangir.hossain@uts.edu.au (M.J.H.); muhammad.zamee@uts.edu.au (M.A.Z.)
† These authors contributed equally to this work.

Abstract: The sustainable integration of distributed energy resources (DER) into distribution networks requires accurate forecasting of hosting capacity. The network and DER variables alone do not capture the full range of external influences on DER integration. Traditional models often overlook the dynamic impacts of these exogenous factors, leading to suboptimal predictions. This study introduces a Sensitivity-Enhanced Recurrent Neural Network (SERNN) model, featuring a sensitivity gate within the neural network's memory cell architecture to enhance responsiveness to time-varying variables. The sensitivity gate dynamically adjusts the model's response based on external conditions, allowing for improved capture of input variability and temporal characteristics of the distribution network and DER. Additionally, a feedback mechanism within the model provides inputs from previous cell states into the forget gate, allowing for refined control over input selection and enhancing forecasting precision. Through case studies, the model demonstrates superior accuracy in hosting capacity predictions compared to baseline models like LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM, and GRU. Study shows that the SERNN achieves a mean absolute error (MAE) of 0.2030, a root mean square error (RMSE) of 0.3884 and an R-squared value of 0.9854, outperforming the best baseline model by 48 per cent in MAE and 71 per cent in RMSE. Additionally, Feature engineering enhances the model's performance, improving the R-squared value from 0.9145 to 0.9854. The sensitivity gate also impacts the model's performance, lowering MAE to 0.2030 compared to 0.2283 without the sensitivity gate, and increasing the R-squared value from 0.9152 to 0.9854. Incorporating exogenous factors such as the time of day as a sensitivity gate input, further improves responsiveness, making the model more adaptable to real-world conditions. This advanced SERNN model offers a reliable framework for distribution network operators, supporting intelligent planning and proactive DER management. Ultimately, it provides a significant step forward in hosting capacity analysis, enabling more efficient and sustainable DER integration within next-generation distribution networks.

Keywords: artificial intelligence; hosting capacity; distributed energy resources; PV; EVs; battery energy storage systems; sensitivity gate; contextual input; sensitivity-enhanced recurrent neural network (SERNN)

check for updates

Academic Editors: Anastasios Dounis, Inga Zicmane and Svetlana Beryozkina

Received: 15 November 2024
Revised: 23 December 2024
Accepted: 5 January 2025
Published: 9 January 2025

Citation: Islam, M.T.; Hossain, M.J.; Habib, M.A.; Zamee, M.A. Adaptive Hosting Capacity Forecasting in Distribution Networks with Distributed Energy Resources. *Energies* **2025**, *18*, 263. https://doi.org/10.3390/en18020263

1. Introduction

The increasing share of distributed energy resources (DER) like photovoltaic (PV), electric vehicles (EVs) and battery energy storage systems (BESS) in the power supply mix presents a great potential to reduce reliance on fossil fuels [1]. To ensure sustainable DER integration and reliable network operation, accurate forecasting of network capabilities is essential [2]. Hosting capacity analysis and forecasting can provide accurate system information and operational data for informed operational planning and sustainable DER integration [3]. The intermittent nature of the DER [4] output complicates grid management and integration planning [5]. Forecasting hosting capacity, the maximum DER integration capability of the network that maintains system reliability and stability [6], is essential for grid planning, operational decisions, and ensuring a reliable energy supply. Advanced hosting capacity forecasting models are required to capture the dynamic characteristics of the DER and their impact on distribution networks. Both conventional and data-driven hosting capacity analysis approaches [6] depend on network analysis. Therefore, these approaches are unsuccessful in capturing the inherent complexities of the dynamic network condition. The dynamic and time-sensitive nature of DERs and network variables, coupled with the limitations of existing methods to adapt to these factors, necessitates the development of an advanced model capable of capturing temporal dependencies and non-linear relationships. This study introduces the Sensitivity-Enhanced Recurrent Neural Network (SERNN) model that can dynamically adapt to time-sensitive and contextual variables while maintaining high predictive accuracy across diverse network conditions.

Different studies have highlighted challenges for DER integration through hosting capacity analysis of distribution networks. The researchers have considered different network and DER constraints like over-voltage, under-voltage, thermal violations, equipment power rating limitations, reverse power flow, inverter control capabilities, DER technology types, and spatial impacts on power quality [7]. Voltage profile was considered as the limiting factor for hosting capacity analysis in [8,9]. Besides voltage profile, researchers considered constraints like line current and transformer overloading in [10,11]. Voltage profile and thermal violation were examined in [12]. In [13], authors considered voltage profile, thermal violation and transformer overloading as constraints for hosting capacity analysis. Active and reactive power losses and voltage deviation were studied in [14]. The authors in [15] highlighted network stability issues. PV was considered in many studies including [16,17]. Due to the increased share of EVs in the transportation fleet [18], different studies considered various aspects of EV hosting capacity. Authors in [19,20] considered the impact of EVs on the distribution network in their study. The authors in [21] investigated the coordinated active and reactive power control approach of DERs and found better overvoltage and overcurrent management in distribution networks. In recent days, hosting capacity analysis using the historical network and DER data has gained much attention among researchers.

Artificial intelligence (AI) can significantly enhance the integration of DER by providing real-time monitoring and forecasting capabilities. Different AI algorithms like Spatial-Temporal LSTM (ST-LSTM) [22], Support Vector Machine (SVM) [23] and whale optimization algorithm [24] were investigated for hosting capacity analysis of distribution networks with DER. In [22], the authors proposed a sensitivity gate and dual forget gate to capture the temporal and spatial impact. The authors in [23] studied the support vector machine (SVM) highlighting the grid-based features. In [25], the authors investigated the regression and classification-based machine learning algorithms. The study found classification-based algorithms like support vector machine (SVM) more promising for PV HC analysis. Hybrid PSO-GD Algorithm in [26] and the deep reinforcement learning-based approach applying the coordinated voltage control scheme in [27] were investigated to evaluate the enhanced hosting capacity of the network. In [24], the authors considered the thermal violation, voltage deviation, voltage fluctuation and short-circuit current for estimating network hosting capacity using the whale optimization algorithm. The multiple linear regression model was investigated using the operational data in [28]. Different machine learning models were investigated in [29] for the distribution network's hosting capacity analysis and forecasting.

The model-free approach using the smart meter data for calculating network capability to accommodate additional DER has gained attention to replace the model-based approaches. This approach estimates the maximum amount of PV systems to integrate with the distribution network without violating voltage limits. The voltage constraint was considered for calculating PV hosting capacity in different studies. The voltage profile was prepared using the smart meter data to calculate network hosting capacity by the univariate linear regression algorithm in [30] and locational hosting capacity analysis algorithm in [31]. Real-time customer voltage data and solar irradiance data were utilized in [32] deploying the deep reinforcement learning-based algorithm. The neural network deployed in [33] for hosting capacity determination from the smart meter data considering the over-voltage constraint. The study in [34] derived the voltage sensitivity coefficient from the smart meter data and estimated the locational voltage-constrained hosting capacity by deploying the linear algebra and optimization technique. The authors in [35] estimate the hosting capacity using the probabilistic evolution method by using the time variability of load and PV generation derived from the smart meter data.

Table 1 compares the relevant studies on hosting capacity analysis and forecasting of distribution networks with DER. This comparison highlights the methodologies used, focus areas of the research, constraints and factors considered for analysis, results or findings achieved and limitations encountered, establishing the foundation for the proposed hosting capacity analysis and forecasting model.

Table 1. Comparison of Hosting Capacity Analysis Methods.

Ref.	Methodology	Focus Area	Factors	Findings	Limitations
[8]	Spatio-temporal Analysis	Hosting capacity optimization	Voltage profile	Estimated 33% and 39% HC for IEEE-37 and IEEE-123 test network, respectively	Static and Scenario-based analysis, Limited to PV systems
[9]	Reconfiguration Algorithm	Low voltage feeder capacity	Voltage, thermal constraints	Increased hosting capacity due to configuring PV in low-voltage feeders	Static and Scenario-based analysis, Limited to PV systems
[10]	Hybrid Optimization	DER placement optimization	Cost, power losses and voltage deviation	Increased both single and multiobjective optimization	Limited generalizability
[11]	Stochastic Probabilistic Analysis	Hosting capacity under uncertainty	Voltage, thermal limits	Improved accuracy under uncertain scenarios	Scenario-based static analysis
[12]	EV Hosting Capacity Management	Low voltage grid management	Voltage, Line congestion, transformer capacity	Identified under voltage as the most restrictive EV impact	Limited to EV charging scenarios
[16]	Stochastic analysis	Planning risks	Voltage, consumption and PV output	Less strict planning risks requires more investment	Static analysis, limited to PV systems
[22]	Spatio-Temporal LSTM	Temporal and spatial impacts on hosting capacity	Voltage profile and DER variability	Decrease the percentage error by 7.9 per cent	Computationally intensive, limited adaptability to dynamic data

Table 1. *Cont.*

Ref.	Methodology	Focus Area	Factors	Findings	Limitations
[23]	Support Vector Machine (SVM)	Grid-based feature analysis	Voltage, Equipment overloading	Achieve almost 90 per cent accuracy	Reliance on expert observations, limited adaptability to real data
[24]	Whale Optimization Algorithm	Hosting capacity assessment	thermal violation, voltage profile and short-circuit current	Error reduction by 12.3%	High computational cost, low convergence and static analysis
[28]	Multiple Linear Regression	Renewable energy hosting capacity	Voltage constraints	Achieved R-squared value 0.9218	Limited to linear relationships
[30]	Smart Meter Data	Model-free PV hosting capacity estimation	Voltage profile	Accurate PV hosting capacity estimation	Narrow scope
[32]	Deep Reinforcement Learning	Real-time PV hosting capacity assessment	Voltage, irradiance	High adaptability to real-time conditions	Extensive training data required

Despite significant advancements, the traditional models face critical challenges in capturing the influence of external environmental factors and managing long-term dependencies on hosting capacity analysis and forecasting. Conventional approaches such as voltage-constraint analysis focus on static network parameters overlooking contextual and time-sensitive variables, thus failing to dynamically adapt to contextual variables or manage long-term dependencies in sequential data effectively. Model-free approaches leveraging smart meter data, while promising, focus narrowly on voltage constraints and overlook dynamic temporal impacts of external factors critical for hosting capacity forecasting. Furthermore, conventional AI-based models like LSTM, GRU, and ConvLSTM, though capable of sequential learning, struggle to retain long-term dependencies, leading to the loss of critical historical information and resulting in suboptimal forecasting accuracy over extended horizons. Additionally, most models lack mechanisms to dynamically incorporate contextual variables, limiting their effectiveness in capturing the dynamic variability of DER and network conditions. For instance, LSTMs excel at retaining long-term dependencies but do not account for real-time contextual variability, which is crucial for hosting capacity forecasting. GRUs streamline computations by reducing gate complexity but do not effectively incorporate exogenous influences, thereby limiting their adaptability in dynamic responses. ConvLSTM improves spatial-temporal learning but remains insensitive to external temporal variables and struggles with long-term dependencies. Addressing these limitations is essential for accurate and responsive hosting capacity forecasting in next-generation distribution networks. These limitations underscore the need for a novel approach that integrates contextual sensitivity and dynamic adjustment mechanisms to address the complexities of hosting capacity forecasting in modern distribution networks. The Sensitivity-Enhanced Recurrent Neural Network (SERNN) model addresses these gaps by introducing a sensitivity gate that dynamically modulates responses to time-sensitive and contextual inputs critical for hosting capacity. By seamlessly integrating these inputs into the memory cell architecture, the model makes it better suited for dynamic scenarios in distribution networks. The model enhances its adaptability and effectiveness in handling the dynamic scenarios of distribution networks with DER.

This study explores the impact of external factors on the dynamic characteristics of networks and DER integration for hosting capacity analysis and forecasting by introducing the contextual sensitivity gate into the neural network memory cell architecture. The sensitivity gate output, combined with the previous cell state, modulates the input gate output to capture the impact of contextual factors and long-term dependencies on network and DER variables. The modified forget output controls information flow, reflecting the impact of contextual factors on hosting capacity analysis and forecasting dynamics. With modified neural network architecture, the proposed model enhances predictive capabilities and flexibility in handling complex time-series datasets. The sensitivity gate improves forecasting performance by addressing the distinct influence of contextual variables, creating a more robust and adaptable hosting capacity analysis and forecasting model. The contributions of the study can be summarized as follows:

(1) Improved Accuracy in Forecasting Network Hosting Capacity: The sensitivity gate of the forecasting model helps to adjust the influence of the network and DER variables dynamically to achieve higher accuracy in network hosting capacity forecasting. The model captures complex dependencies and non-linear characteristics of the data through improved responsiveness to temporal patterns and external factors that impact the network hosting capacity.
(2) Enhancing Network Hosting Capacity Forecasting through Effective Long-Term Dependency Management on Network and DER Data: Incorporating a sophisticated cell state update mechanism with forget and input gates, alongside sensitivity adjustments, the model effectively manages long-term dependencies on Network and DER data. By optimizing the retention of relevant historical information while discarding irrelevant data, the model improves hosting capacity forecasting performance.
(3) Improved Adaptability to Contextual Variables for Hosting Capacity Forecasting: The model supports extensive fine-tuning of hyperparameters, layers, and configurations to optimize the model's performance in varied DER deployment scenarios. Its flexibility in integrating contextual variables like time of the day enables effective adaptation to diverse datasets and scenarios, accommodating the intermittent nature of the DER output.

The rest of the paper is organized as follows: Section 2 describes the methodology and forecasting model, Section 3 illustrates the results and discussion, and Section 4 contains the conclusion of the study.

2. Methodology and Forecasting Model

This section outlines the methodology and model description for forecasting the future value of the distribution network's hosting capacity. Figure 1 shows the different stages of the proposed Sensitivity-Enhanced Recurrent Neural Network (SERNN) model. The process begins with a power flow analysis of the distribution network with DER like PV, EVs, and BESS. The power flow analysis provides the time series data containing the network hosting capacity and the network and DER variables. After cleaning and processing, the data fit into the hosting capacity analysis and forecasting model. The feature engineering step provides additional features that enhance the dynamic response of the model [36]. After training, validation, and testing, the hosting capacity analysis and forecasting model will provide the future value of the network hosting capacity for reliable network operation and sustainable DER integration.

Figure 1 illustrates the adaptive HC forecasting process in distribution networks with integrated DER. The figure provides a high-level overview of the methodology, showcasing time-series data preparation, input data processing, and hosting capacity forecasts through the SERNN forecasting model. The process begins with input data acquisition, such as the network parameters, DER data, and weather information. The power flow analysis provides the time-series input data required for the forecasting architecture. The forecasting architecture includes data processing and feature engineering techniques that enhance the time-series data. The processed data is then fed into the SERNN model, which incorporates a sensitivity gate for dynamically adjusting the response to time-sensitive and contextual inputs. Sequential processing is performed through LSTM and dense layers, which manage long-term dependencies and non-linear relationships in the data. The model then forecasts the hosting capacity which is evaluated using standard performance metrics like MAE, RMSE, and R-squared value. This comprehensive process ensures accurate and robust predictions of hosting capacity under dynamic network conditions.

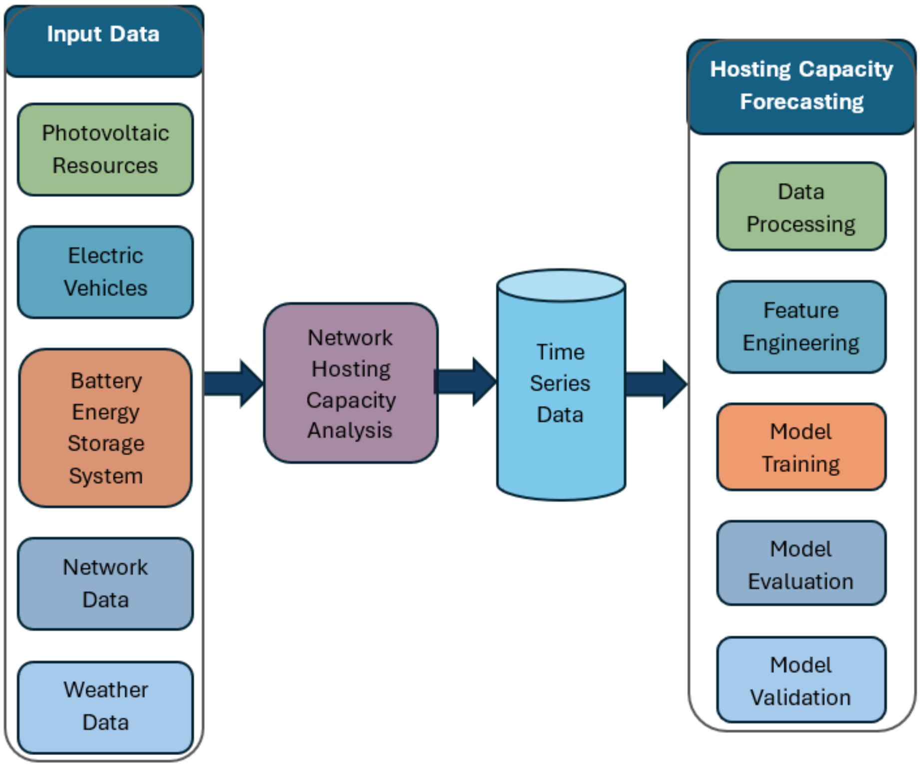


Figure 1. Process Flow Diagram of the SERNN Model.

The following sections outline the methodology and forecasting model development stages with necessary mathematical equations.

2.1. Hosting Capacity Analysis and Time-Series Data Preparation

To calculate the hosting capacity, network, and DER variables, the required time-series data gathered by power flow analysis of the IEEE-123 distribution network with DER like PV, EVs, and BESS considering over-voltage, under-voltage, thermal limits, and transformer overloading constraints. The power flow analysis was conducted by OpenDSS [37]. The hosting capacity analysis and forecasting model was derived from the distribution network with DER using Python. Data exchange between OpenDSS and Python was performed using the py-DSS-interface application.

The impact of the DER on the network hosting capacity was analyzed using the PV and BESS models of OpenDSS. The initial PV system capacity was considered 10 kVA at unity power factor. The PV output was calculated using irradiance and temperature data from the Bureau of Meteorology, Australia. The EV battery specifications for the study were calculated based on the Tesla Model S, with a nominal capacity of 100 kWh and an effective capacity of 95 kWh. The EV charging and discharging were coordinated under grid-to-vehicle (G2V) and vehicle-to-grid (V2G) [38] modes. The rated capacity of BESS was considered 300 kWh and 100 kW, with charging, discharging, and reserve capacities set at 30%, 40%, and 20%, respectively. DER was connected to the distribution network with randomly selected buses.

The hosting capacity analysis began with reading network and DER data. Power flow analysis evaluated the selected constraint violations as shown in Equations (1)–(3). As long as constraints remained within limits, DER capacity was incremented, and constraints were re-evaluated. When constraints were violated, network hosting capacity and required variables for the forecasting model were collected for the hosting capacity analysis and forecasting model training and testing.

(a) Voltage constraint:

$$v_{upper} \geq v_i \geq v_{lower} \;\; such\ that\ \forall i \epsilon \mathbb{N} \tag{1}$$

where the upper and lower limit of the i_{th} node voltage v_i are v_{upper} and v_{lower}, respectively.

(b) Thermal constraint:

$$I_i \leq I_{rated} \tag{2}$$

where I_i is the current passing through the lines i and I_{rated} is the rated current of the conductor.

(c) Transformer overloading:

$$P_{xformer} \leq P_{xformer\_rated} \tag{3}$$

where $P_{xformer}$ is the power delivered by the distribution transformer and $P_{xformer\_rated}$ is the rated capacity of the distribution transformer.

2.2. Time-Series Data Preparation

The historical time-series data for network and DER variables were collected through power flow analysis of the IEEE- 123 test network with DER like PV, EVs, and BESS by the open-sourced network analysis software OpenDSS. The network and DER variables were further analyzed using Python to calculate the network hosting capacity. The py-DSS-interface application integrated the OpenDSS with Python to exchange the data of the power flow analysis for developing the hosting capacity analysis and forecasting model.

2.3. Data Processing

Data consistency, accuracy, and completeness of the data are crucial for better performance of the hosting capacity analysis and forecasting model. The time-series data was cleaned using statistical tools to conform to the model. The 'datetime' module performed time-stamping of the dataset. Following the required feature engineering and removing the missing values, outliers, and duplicates, the data were used for the training and testing of the model.

2.4. Feature Engineering

Feature engineering extracts relevant information from the raw data to improve the performance of the model. The presence of cyclic features affects the cyclic relationship of the model when they are transformed to ordinal, interval, or categorical encoding [39]. The time of the day plays a vital role in accurately predicting the network's capability to integrate renewable resources. The time of day that indicates the hosting capacity, network, and DER variables estimation time was considered an input variable for this study. It was reconstructed into the cyclic feature using the sine-cos encoding technique [40] by Equations (4) and (5).

$$h_{sine} = sine(2\pi * h/24) \tag{4}$$

$$h_{cos} = cos(2\pi * h/24) \tag{5}$$

where h indicates the time of day, h_{sine} and h_{cos} are the sine-cos encoded features.

The hosting capacity of the distribution network depends on the presence of PV output and solar irradiance. There was no power intake from the PV resources into the network at no solar irradiance. Based on this conditionality, the new features were extracted from the PV output and solar irradiance using the One-Hot encoding technique [41]. This technique modified each PV output and irradiance into the binary coded inputs using the conditionality described in Equation (6).

$$f(x_t) = \begin{cases} 1 & if\ x_t > 0; \\ 0 & otherwise \end{cases} \tag{6}$$

where x_t is the value of input variable at time t.

This transformation identifies whether a PV output and solar irradiance actively contribute to the network at a given time. This encoding assigns a binary value of 1 to indicate the presence of PV generation or solar irradiance ($x_t > 0$) and 0 to indicate their absence ($x_t \leq 0$). By converting these variables into binary features, the model can efficiently distinguish between conditions where PV resources contribute to the network and conditions where they do not. This transformation simplifies the input representation, enabling the model to explicitly capture the binary state of PV generation activity and the corresponding network dynamics. The one-hot encoded features enhance the model's ability to distinguish periods with active DER contributions from inactive periods, thus improving the accuracy of hosting capacity forecasting.

2.5. Model Design

The challenges in capturing long-term dependencies and incorporating external environmental factors, such as weather patterns and temporal variations, limit the effectiveness of conventional hosting capacity forecasting models. The SERNN model addresses these challenges by introducing a sensitivity gate that dynamically adjusts the model's response to time-sensitive inputs of the distribution network, integrated DER, and surrounding weather data. By incorporating the previous cell state into the forget gate, the model also overcomes the challenge of managing long-term dependencies, ensuring that critical historical information is retained for accurate forecasting. These innovations enable the model to deliver improved forecasting accuracy, particularly for active management of the complex and dynamic operation of the distribution network with DER.

The goal of the proposed model is to forecast the future value of the network hosting capacity HC_t for time steps T. In the proposed model, the sensitivity gate will capture the impact of contextual variables on the hosting capacity of the distribution network. Also, it will capture the long-term dependency on input variables by feeding the previous cell state into the input gate. The forget gate will help to avoid overfitting by selectively forgetting certain parts of the previous cell states.

Consider the time series data of the network and DER variables $X_{hc} \varepsilon \Re^{Txm}$ where T is the time step and m is the number of feature variables. Therefore, the hosting capacity X_{hc} at time t could be expressed as,

$$X_{hc(t)} = \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,m} \\ x_{2,1} & x_{2,2} & \dots & x_{2,m} \\ \dots & \dots & \dots & \dots \\ x_{T,1} & x_{T,2} & \dots & x_{T,m} \end{pmatrix}.$$

The sensitivity gate input S that will control the gate operation to modify the input variables could be written as $S_{hc} \varepsilon \Re^{Tx1}$, i.e., $S_{hc(t)} = \begin{pmatrix} s_1 \\ s_2 \\ \dots \\ s_T \end{pmatrix}$.

Thus, the future value of the hosting capacity $hc \varepsilon \Re^{Tx1}$ would be a vector of size T, i.e.,

$$HC_t = \begin{pmatrix} hc_1 \\ hc_2 \\ \dots \\ hc_T \end{pmatrix}.$$

The objective function of the proposed model could be written as shown in Equation (7).

$$L = \frac{1}{T} \sum_{t=1}^{T} (\widehat{hc_t} - hc_t)^2 \quad (7)$$

where hc_t is the actual and $\widehat{hc_t}$ is predicted hosting capacity value at time t. T represents the length of the forecasting time step.

2.5.1. Cell Architecture of the SERNN Model

This section outlines the cell architecture of the proposed SERNN model as shown in Figure 2. It consists of three gates: the forget gate, the input gate and the sensitivity gate. The 'sigmoid (σ)' and 'tanh' activation function [42] modify the output at different stages. The following sections summarize the operation of each gate and information flow within the memory cell.

(a) Forget gate (f_t):

The forget gate in the model determines what information to propagate within the cell from one state to another [43]. The previous hidden state h_{t-1}, previous cell state C_{t-1} and current input x_t are concatenated to form a single vector $[x_t, h_{t-1}, C_{t-1}]$. The vector multiplied by a weight matrix W_f, and a bias b_f is added to form the input at the forget gate. The resultant is passed through the 'sigmoid' function $(\sigma(x) = 1/(1 + e^{-x})$ to get the forget gate output f_t. Equation (8) denotes the forget gate output within the cell of the model.

$$f_t = \sigma(W_f[x_t; h_{t-1}; C_{t-1}] + b_f) \quad (8)$$

(b) Input Gate (i_t):

The input gate concatenates the previous cell state (C_{t-1}), previous hidden state (h_{t-1}) and input variables (x_{t-1}) that provide a comprehensive context of the network and DER variables along with the weather information by incorporating the current and historical data. It multiplies the concatenated input vector with the weight factor W_i which adjusts the impact of each variable on the gate output and adds bias b_i. The activation function (σ) helps to achieve the non-linearity and allows the model to learn the complex nature of the input variables pattern. Therefore, this architecture enables the input gate to incorporate new input data and retain relevant past information for accurate forecasting of the network hosting capacity. The input gate equation could be written as shown in Equation (9).

$$i_t = \sigma(W_i[x_t; h_{t-1}; C_{t-1}] + b_i) \tag{9}$$

(c) Sensitivity Gate (s_t):

The sensitivity gate modulates the flow of information into the cell state based on contextual variables, enabling the adaptability and robustness of the model to external factors. Thus, the model gets responsiveness to time-sensitive variables through dynamic adjustment of the gate output. The sensitivity gate maps the sensitivity input x_s by applying the weight matrix W_s and bias factor b_s. The activation function σ helps to regulate the output. The sensitivity gate output s_t could be represented by the Equation (10)

$$s_t = \sigma(W_s x_s + b_s) \tag{10}$$

where σ is the activation function, W_s is the weight matrix, x_s is the sensitivity gate input and b_s is the sensitivity bias.

The sensitivity gate addresses the limitations of existing models by capturing the influence of time-sensitive contextual variables. The gate dynamically modulates the input variables (x_s) based on contextual input like 'Hour' to respond to time-sensitive network and DER variables and exogenous factors. The sigmoid activation function (σ) in Equation (10) scales contextual inputs dynamically. This mechanism ensures temporal adaptability by integrating contextual information into the memory cell. Such adaptability enhances prediction accuracy by tailoring model responses to time-sensitive external factors, a distinct advantage over conventional gated architectures. By incorporating the previous cell state into the forget gate, the model overcomes the challenge of managing long-term dependencies of traditional models like LSTM and GRU. Compared to traditional models, which lack external contextual modulation, the sensitivity gate introduces adaptive learning capabilities, enhancing the model's forecasting accuracy for time-varying scenarios.

(d) Cell State Update $(i_{(t,s)})$:

The cell state update $(i_{(t,s)})$ performs the dynamic adjustment based on the input and sensitivity gate. It allows new information integration into the cell state effectively. To determine the current cell state, it also determines which type and how much additional information will be provided for accurate forecasting of the hosting capacity. The cell state update value $(i_{(t,s)})$ is the multiplication of the input gate output (i_t) and sensitivity gate output (s_t), transformed by the non-linear function (tanh). The non-linear function helps to prevent exploding and ensures stability, maintaining the $i_{(t,s)}$ between -1 and 1. The cell state update value $(i_{(t,s)})$ could be calculated as shown in Equation (11).

$$i_{(t,s)} = \tanh\left[\sigma(W_s x_s + b_s)\sigma\left(W_i \begin{bmatrix} x_t \\ h_{t-1} \\ C_{t-1} \end{bmatrix} + b_i\right)\right] \tag{11}$$

where W_s and W_i are the weight matrix, x_s and x_t are the sensitivity and input feature vector, h_{t-1} is the previous hidden state, C_{t-1} is the previous cell state, and b_s and b_i are the biases associated with the sensitivity and input gates, respectively. The inclusion of the sensitivity gate with the cell state update enables the model to dynamically adjust to time-dependent variables, thus enhancing model forecasting accuracy.

(e) Current Cell State (C_t):

The current cell state represents the network's internal memory at a given time step. It allows the model to learn from complex and dynamic data patterns, utilizing a selective memory storage mechanism. To predict the current output, relevant information based on input data and historical context is kept over extended periods, helping the model to capture dependencies and patterns in sequential data. Several key components are involved in the cell state update process: forget gate, input gate, and input modulated transformation. Equation (12) captures the dynamic integration of past and present information in the cell state (C_t). The forget gate output (f_t) determines the proportion of previous memory (C_{t-1}) to retain, ensuring long-term dependencies are preserved. The term ($i_{t,s}$), derived from the input gate (i_t) and sensitivity gate (s_t), introduces new information influenced by external contextual variables. The combination of these terms enables the model to dynamically adjust to time-varying and nonlinear dependencies, ensuring robust and adaptive hosting capacity forecasting.

$$C_t = f_t \odot C_{t-1} + i_{t,s} \tag{12}$$

The current cell state (C_t) represents the internal memory of the SERNN model at time t. It selectively retains historical information ($f_t \odot C_{t-1}$) while dynamically integrating new inputs ($i_{t,s}$). This process is governed by nonlinear activations (σ, $tanh$) that ensure stable gradient propagation and prevent exploding gradients. Equation (12) thus balances memory retention and contextual adaptation, critical for accurate forecasting of hosting capacity in distribution networks with integrated DER. This integration ensures that the model adapts effectively to time-varying inputs.

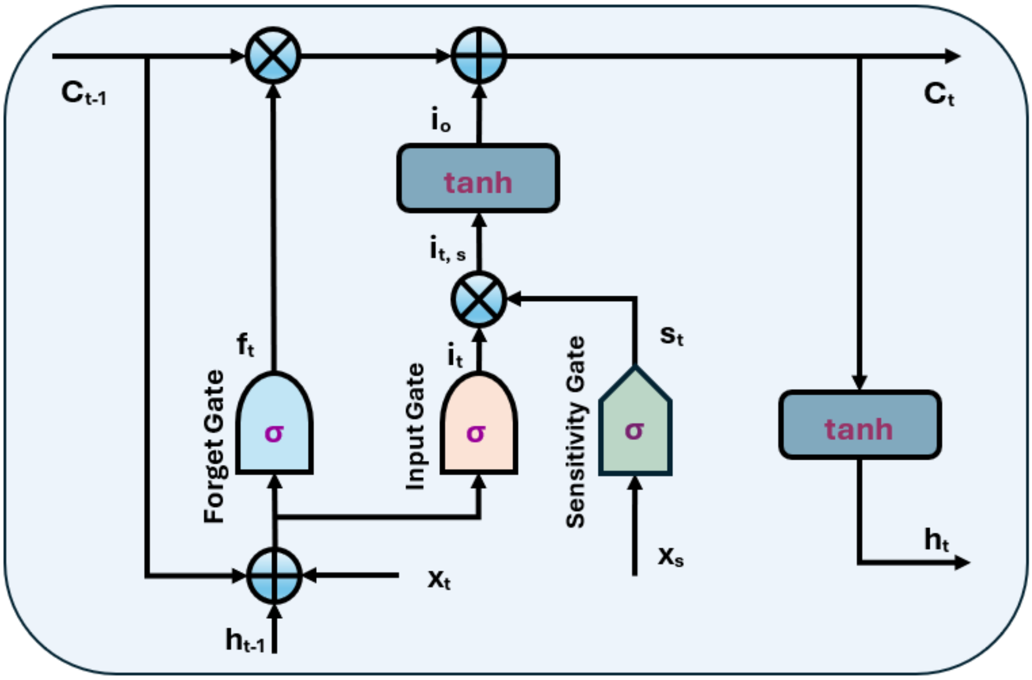


Figure 2. Cell Architecture of the SERNN Model.

(f) Current Hidden State (h_t):

The current hidden state helps the model to determine the future value of the network hosting capacity using trends, seasonal patterns, and other time-dependent dynamics of the input variables. The quality of the hidden state directly impacts the accuracy and reliability of predictions. A well-trained hidden state can capture complex patterns and long-term dependencies, leading to more accurate forecasting of future hosting capacity. Transforming the cell state by a non-linear activation function (tanh) guides the output and makes informed predictions. Equation (13) computes the current hidden state within the model.

$$(h_t) = tanh(C_1) \tag{13}$$

The predicted hosting capacity y_t at time t is calculated by applying the wight (W_o) and bias (b_o) with the current cell state (h_t) by Equation (14).

$$y_t = W_o h_t + b_o \tag{14}$$

2.5.2. Weight and Bias Vectors Update

The weight (W_f, W_i, W_s, W_o) and bias terms (b_f, b_i, b_s, b_o) in Equation (8)–(11) are optimized during model training using backpropagation through time (BPPT) [44]. This process minimizes the loss function as shown in Equation (7). The gradients of the loss function concerning these parameters are calculated and used to update the weight and bias terms iteratively. The forget, input, and sensitivity gates dynamically govern the flow of information, ensuring consistency between the parameter updates and the objective function. For example, the forget gate's weight W_f is updated using the gradient shown in Equation (15).

$$\frac{\partial L}{\partial W_f} = \sum_{t=1}^{T} \frac{\partial L}{\partial h_t} \cdot \frac{\partial h_t}{\partial C_t} \cdot \frac{\partial C_t}{\partial f_t} \cdot \frac{\partial f_t}{\partial W_f} \tag{15}$$

Similar updates occur for W_i, W_s, b_f, b_i, and b_s ensuring that all parameters align with the model's objective to minimize the forecasting error.

The optimization objective function (Equation (7)) minimizes the squared error across the forecasting time horizon T, ensuring precise predictions of the network hosting capacity. Using backpropagation through time (BPTT), gradients are calculated for all model parameters, including sensitivity gate weights (W_s) and biases (b_s). The additional gradient term for the sensitivity gate, $\frac{\partial L}{\partial s_t}$, ensures that the model dynamically adjusts to contextual variables during training. The update for the sensitivity weight matrix can be expressed as:

$$Ws \leftarrow W_s - \eta \frac{\partial L}{\partial W_s}, \tag{16}$$

where η is the learning rate, ensuring convergence to the optimal parameter set. This approach incorporates the contextual variability that conventional gated networks like LSTM overlook. The iterative approach by incorporating sensitivity input gradients $(\frac{\partial L}{\partial s})$ enables the network to adjust dynamically to time-sensitive variables. This theoretical framework underpins the SERNN's superior accuracy by harmonizing long-term dependencies and temporal sensitivity in hosting capacity forecasting.

2.5.3. Model Evaluation

The model performance was evaluated using the performance indices: mean absolute error (MAE), root mean squared error (RMSE) and R-squared (R^2) as shown in Equation (17), Equation (18) and Equation (19), respectively [45].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|u_i - \hat{u}_i| \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|u_i - \hat{u}_i|^2} \tag{18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(u_i - \hat{u}_i)^2}{\sum_{i=1}^{n}(u_i - \bar{u}_i)^2} \tag{19}$$

where u_i represents the original data, $\hat{u}_i$ the predicted values, $\bar{u}_i$ the mean of actual data and n is the total number of data points.

To ensure fair comparisons, the tuning for all models, including the baseline models, was conducted using the Keras Tuner framework with a random search over a consistent hyperparameter search space. The search spaces included learning rates (1×10^{-2}, 1×10^{-3}, 1×10^{-4}), number of layers (1–3 for LSTM/GRU, 2 for ConvLSTM/Stacked LSTM) and units per layer 32–128. Optimizer 'Adam' and 'MSE' as the loss function ensured uniformity in evaluation. Table 2 summarizes the best hyperparameters obtained for each model. This tuning process ensured that each model achieved its optimal configuration under identical conditions for a fair and robust comparison across all models.

Table 2. Hyperparameter Settings for Baseline Models.

Model	Learning Rate	Number of Layers	Units per Layer	Optimizer
LSTM	1×10^{-3}	2	64, 128	Adam
ConvLSTM	1×10^{-4}	1	32	Adam
BiLSTM	1×10^{-3}	2	64, 64	Adam
Stacked LSTM	1×10^{-4}	3	64, 96, 128	Adam
GRU	1×10^{-2}	2	128, 64	Adam

Algorithm 1 outlines the proposed forecasting model with the SERNN model. The algorithm integrates data processing and advanced feature engineering techniques to prepare the time-series data gathered through power flow analysis of the distribution networks with integrated DER. The SERNN model processes these data through sensitivity gate modulation architecture to capture the temporal dependencies of the data. The adaptive memory cell architecture combines forget, input, and sensitivity gates to manage long-term dependencies while dynamically incorporating new information. Using Backpropagation Through Time (BPTT), the algorithm iteratively optimizes weights and biases to minimize forecasting errors, ensuring robustness and adaptability. Adapting to time-sensitive inputs enhances the model's predictive capabilities, achieving superior performance metrics compared to baseline models.

Algorithm 1 SERNN for Hosting Capacity Forecasting

Require: Contextual Input (network, DER, and weather variables) and Sensitivity Input (Time of the day)
Ensure: Forecasted Hosting Capacity (HC_t)

1: **Step 1: Data Preprocessing and Feature Engineering**
- Data cleaning and temporal alignment
- Cyclic encoding of time-related variables ($Hour$) using sine-cosine transformations ($h_{\sin}, h_{\cos}$)
- Binary encoding of time-dependent variables
- Normalizing all input variables to ensure data consistency

2: **Step 2: Initialize Parameters**
- Initialize weight matrices W_f, W_i, W_s, W_o and biases b_f, b_i, b_s, b_o.
- Set hyperparameters

3: **Step 3: Sensitivity Gate Integration**
- Compute sensitivity gate output to modulate time-sensitive variables: $s_t = \sigma(W_s x_s + b_s)$

4: **Step 4: Memory Cell Operations**
- Compute forget gate output: $f_t = \sigma(W_f[X_t; h_{t-1}; C_{t-1}] + b_f)$
- Compute input gate output: $i_t = \sigma(W_i[X_t; h_{t-1}; C_{t-1}] + b_i)$
- Update cell state dynamically: $C_t = f_t \odot C_{t-1} + tanh(i_t \odot s_t)$

5: **Step 5: Hosting Capacity Forecasting**
- Compute the hidden state to capture temporal dependencies: $h_t = \tanh(C_t)$
- Generate hosting capacity forecast: $\hat{y}_t = W_o h_t + b_o$

6: **Step 6: Backpropagation and Parameter Optimization**
- Minimize the loss function L using Backpropagation Through Time (BPTT)
- Update weights iteratively: $W \leftarrow W - \eta \frac{\partial L}{\partial W}$
- Ensure gradients are dynamically adjusted for sensitivity gate weights W_s and biases b_s.

7: **Step 7: Model Evaluation and Validation**
- Evaluate forecasting performance (MAE, RMSE, and R^2)

8: **Step 8: Forecast Refinement**
- Adjust hyperparameters based on evaluation metrics.

9: **END**

3. Results and Discussion

To verify the effectiveness of the proposed forecasting model, the IEEE-123 distribution network was considered to conduct numerical solutions. DER like PV, EVs, and BESS were connected to the selected buses. The temperature and irradiance data required for calculating the PV output were collected from the Bureau of Meteorology (BoM) Australia. Figures 3 and 4 show the sample of irradiance and temperature data used for the study.

All simulations were conducted on a laptop with an 11th Gen Intel(R) Core(TM) i7-1185G7 CPU @ 3.00 GHz and 16 GB RAM. OpenDSS version 9.6.1.1 (64-bit build) was used for test network simulations and interfaced with Python 11 through the Py-DSS-Interface 2.0.4 interface library in the PyCharm Community Edition environment. For the IEEE-123 distribution network, the over-voltage and under-voltage limits were considered 1.05 pu and 0.95 pu, the thermal limit as the rated current passing through the conductors and transformer overloading as the rated power of the distribution transformers connected in the networks. During the simulation, the total circuit power was measured at the secondary terminals of the main distribution transformer delivering the total load of the test networks. Renewable resources like PV, EVs, and BESS were connected to the selected buses to reflect real scenarios during network simulation. The PV, EVs, and BESS were operated following the OpenDSS model architecture described in [37].

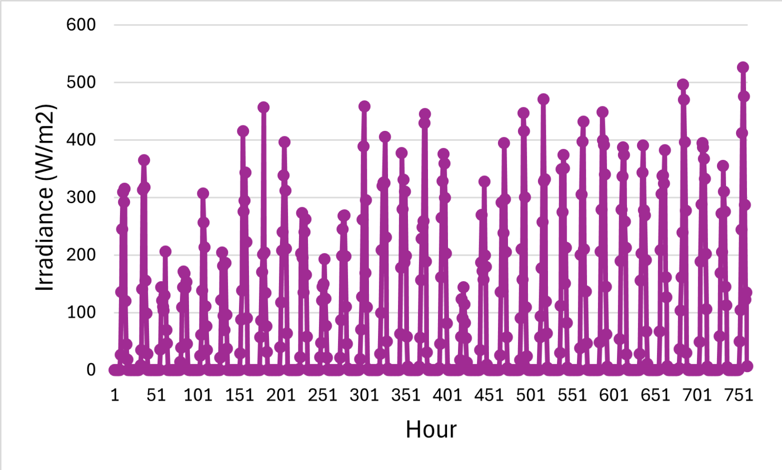


Figure 3. Irradiance.

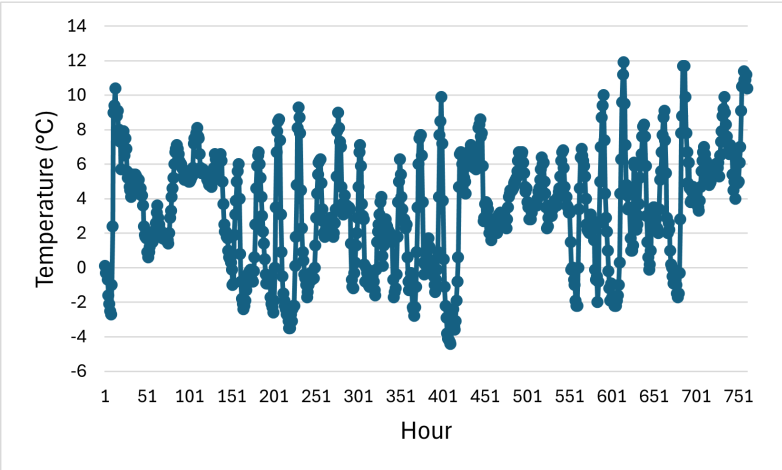


Figure 4. Temperature.

Figure 5 shows the total circuit power delivered by the secondary terminal of the main distribution transformer and the generated power by the PV units connected with the IEEE-123 distribution network. The hosting capacity of the network was calculated using network variables like circuit power, voltage profile, current passing through the conductors, and transformer power. The impacts of PV, EVs, and BESS were evaluated using the hosting capacity analysis framework. The hosting capacity was estimated at hourly intervals using the open-sourced software Python based on the power flow analysis performed by OpenDSS. The py-DSS-interface application package exchanged data between OpenDSS and Python. After completing the data preparation and feature engineering processes described in Sections 2.3 and 2.4, the data was fitted into the proposed hosting capacity forecasting model. The study used 80% data as the training set for the model's training. The remaining data was divided into validation (50%) and testing (50%) sets for validating and testing the model, respectively. To maintain the integrity of the time-series data, no data shuffle was allowed throughout the simulation.

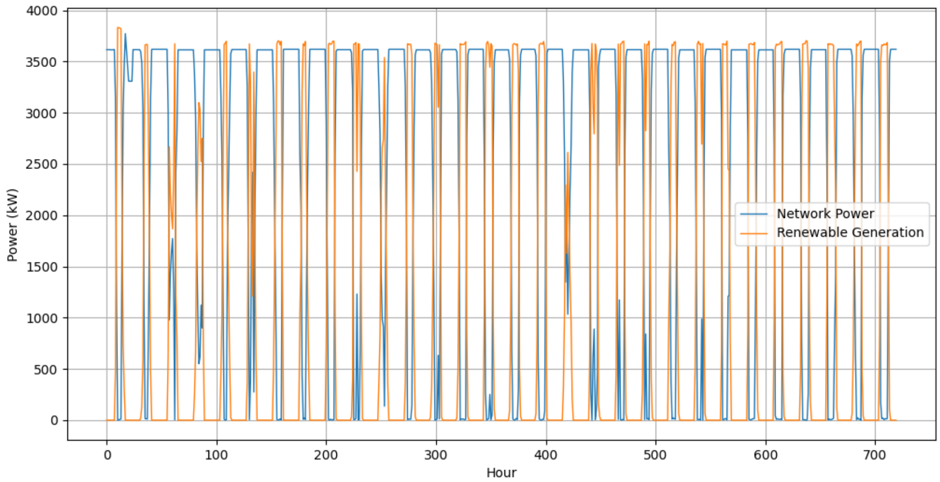


Figure 5. Total Network Power and Renewable Generation (IEEE-123).

The performance of the proposed hosting capacity forecasting model was evaluated using three different case studies. The outcomes and findings of these case studies have been outlined as follows:

3.1. Case Study 1: Impact of Feature Engineering on Hosting Capacity Analysis and Forecasting

Feature engineering is the critical stage for any prediction model. This case study evaluated the impact of advanced feature engineering techniques on the model performance. The model was trained and tested under different combinations of the network, DER, and weather-related input variables collected through power flow analysis and features engineering as described in Section 2 and evaluates their impact on the model's predictive performance. The model was assessed at three key stages: before data processing, after data processing but before feature engineering, and after feature engineering. In the first stage, the model was trained and tested with the raw data while in the second stage, the data were processed to remove duplicate, missing values, outliers, and data inconsistencies. The new features were incorporated into the dataset in the third stage using cyclic encoding and binary representation techniques. Table 3 presents the impact of these steps on the predictive performance of the hosting capacity forecasting model, using key metrics such as MAE, RMSE, and R-squared value.

Table 3. Impact of Feature Engineering on the Forecasting Performance of the Model.

Indices	Before Data Processing	Before Feature Engineering	After Feature Engineering
MAE	0.6301	0.5501	0.2030
RMSE	0.9801	0.9408	0.3884
R-squared	0.8569	0.9145	0.9854

Table 3 shows significant improvements at each stage. During data processing, the removal of outliers, duplicates, and inconsistencies enhanced data quality, leading to a reduction in MAE from 0.6301 to 0.5501, RMSE from 0.9801 to 0.9408, and an increase in R-squared value from 0.8569 to 0.9145. These changes underscore the importance of addressing data inconsistencies and ensuring high-quality input data for robust model performance. Feature engineering further enhanced the model's predictive capabilities by introducing new input variables such as cyclic time encoding and binary conditional features. These transformations reduced MAE from 0.5501 to 0.2030 and RMSE from 0.9408 to 0.3884, while increasing the R-squared value to 0.9854. The result demonstrates that the inclusion of new feature variables through advanced feature engineering techniques like cyclic encoding of time variables and binary representation of time-dependent variables significantly improves the model's ability to capture complex temporal patterns and enhance forecasting accuracy through better representation of input features. This analysis validates the critical role of thorough data preparation and feature engineering in improving the predictive performance of the hosting capacity forecasting model. This study provides the foundation for impact analysis of the sensitivity gate as discussed in the following section.

3.2. Case Study 2: Impact of Sensitivity Gate Input on the Hosting Capacity Analysis and Forecasting

The inclusion of the sensitivity gate in the SERNN model directly addresses the limitations of the conventional, AI-based and model-free hosting capacity analysis and forecasting models. By dynamically adjusting the model response to the contextual and time-sensitive network and DER variables, it demonstrates enhanced capability to capture temporal response and improved forecasting accuracy as shown in Table 4. The SERNN model was trained with and without the sensitivity gate using identical datasets and hyperparameters. The test evaluates the impact of the sensitivity gate on model performance and the capability to capture time-dependent dynamics and contextual variability of the distribution network with DER.

Table 4. Performance Comparison with and without Sensitivity Gate.

Model Variant	MAE	RMSE	R-Squared Value
SERNN (with Sensitivity Gate)	0.2030	0.3884	0.9854
SERNN (without Sensitivity Gate)	0.2283	0.4570	0.9152

The results in Table 4 show that the MAE drops from 0.2283 to 0.2030 and the RMSE declines from 0.4570 to 0.3884 while the R-squared value improves from 0.9152 to 0.9854 due to the inclusion of the sensitivity gate with the model. This demonstrates the sensitivity gate's capability to capture dynamic, time-sensitive characteristics of the network and DER variables, and improve forecasting accuracy. The sensitivity gate ensures that contextual variability, such as PV output variability tied to irradiance and temperature, is effectively captured, significantly reducing MAE and RMSE values.

Different inputs of the sensitivity gate, like time of day, and network and DER variables, impact the network hosting capacity differently. The performance of the proposed model was measured using the MAE, RMSE, and R-squared Value. Figure 6 demonstrates the impact of different sensitivity gate input variables, like time of day (Hour), net circuit power with DER (CktPower), Maximum node voltage (V_{Max}), Minimum node voltage (V_{Min}), Maximum current passing through the conductors (I_{max}) and Power output from the connected PV resources (PV(kW)), on the hosting capacity analysis and forecasting performance of the model.

The Figure 6 shows that the 'Hour' input achieves the lowest MAE and RMSE values and the highest R-squared value, demonstrating its superior ability to capture the dynamic variability of network and DER conditions. In contrast, 'PV(kW)' yields the lowest performance, reflecting its limited influence in representing temporal characteristics. This superior performance can be attributed to the inherent temporal dependencies between hosting capacity and time of the day. The result indicates that the time of the day ('Hour') as sensitivity gate input captures the dynamic response of the network and DER variables for calculating the future value of the network's hosting capacity better than other input variables. For instance, solar irradiance and air temperature strongly influence PV output and BESS charging behavior, which varies predictably with time, making 'Hour' an ideal contextual input to capture the temporal response of the network by the SERNN model. The performance of the SERNN model with the sensitivity gate has been compared with the baseline models in the following section.

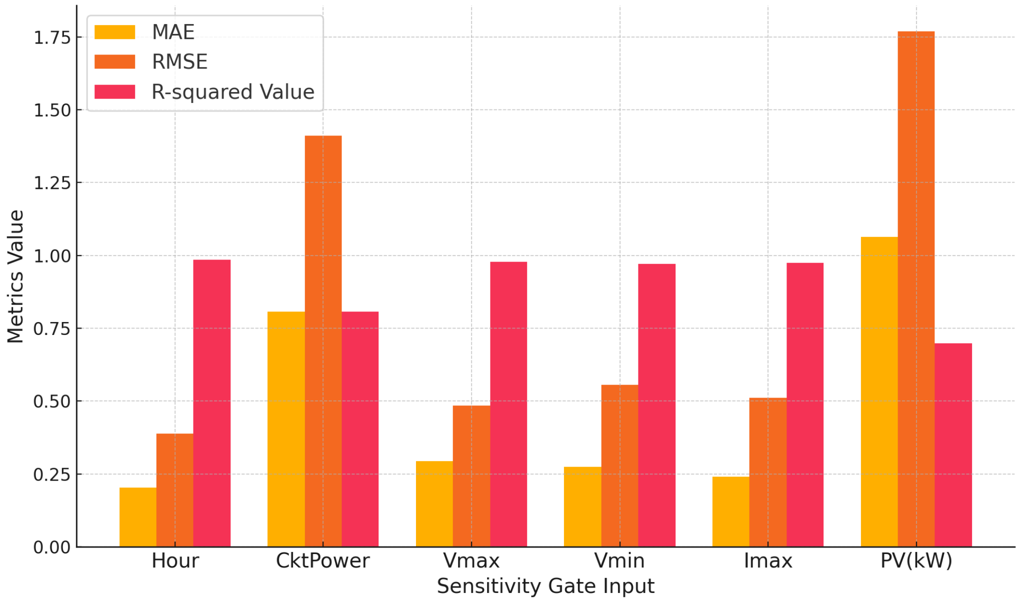


Figure 6. Impact Analysis of Sensitivity Gate Input on Model Performance.

3.3. Case Study 3: Comparison of SERNN with Baseline Models

The superior performance of the SERNN model compared to baseline models demonstrates its effectiveness in addressing the limitations of traditional hosting capacity forecasting approaches. Specifically, the model's architecture resolves challenges in capturing long-term dependencies and contextual variability, enabling a more robust and accurate prediction framework. This is reflected in the significant improvement in key performance metrics (MAE, RMSE, R-squared Value) over existing models. The model is implemented in TensorFlow/Keras with the following key elements:

(1) Input layers: It consists of LSTM layers that dynamically process the feature inputs, retaining long-term dependencies via the forget gate.
(2) Sensitivity Layers: It processes the input variables based on the sensitivity input that enables the model to respond to time-sensitive variables.
(3) Dense Layers: It calculates the s_t that enables the model to modulate new information.
(4) Concatenation: It combines the outputs of the LSTM layer and sensitivity gate for integrated processing in subsequent LSTM layers.

The sensitivity and dense layers dynamically update the cell state, reflecting Equation (12), and compute the final hidden state used for hosting capacity forecasting. This architecture ensures that the cell state (C_t) selectively retains past information ($f_t \odot C_{t-1}$) while incorporating new contextually modulated information ($i_{t,s}$). By leveraging TensorFlow/Keras's robust backpropagation mechanisms, all parameters ($W_f, W_i, W_s, b_f, b_i, b_s$) are iteratively optimized to align with the objective function (Equation (7)). The model captures both short-term variability and long-term dependencies in hosting capacity forecasting.

Figure 7 shows the data flow architecture within the proposed SERNN forecasting model. The SERNN model architecture consists of 60 unit LSTM layers to capture the sequential dependencies across time steps and return the full sequence output. The specialized sensitivity input layer processes the sensitivity variable through a Dense layer with 10 units using the ReLU activation function. The sensitivity layer transformed the sensitivity input into a more expressive representation. The output from the LSTM layer is concatenated with the processed sensitivity input along the feature dimension, allowing the model to simultaneously process both the primary input features and the sensitivity input. With 60 units, the second LSTM layer consolidates the combined data into a single sequence before the output layer. Finally, a Dense layer with one unit produces the hosting capacity prediction. The model is compiled with the Adam optimizer, MSE as the loss function, and MAE as an evaluation metric to evaluate forecasting accuracy, providing an optimized framework to capture the influence of dynamic, time-sensitive variables in distribution networks with DER.

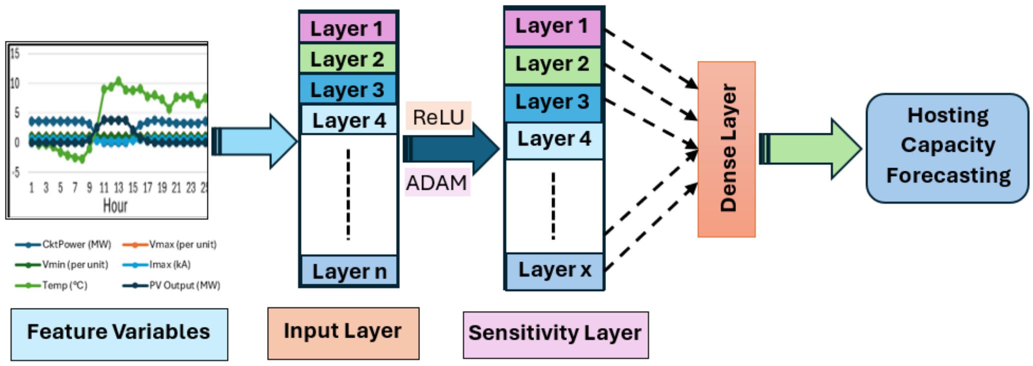


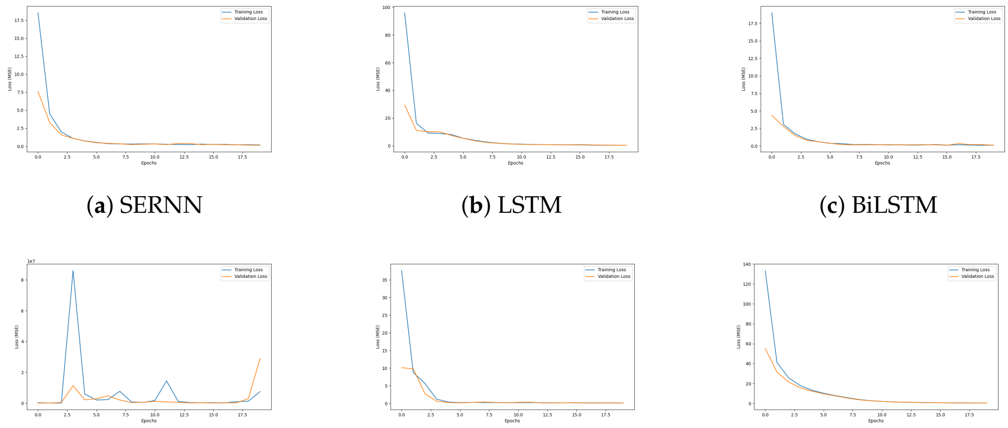
Figure 7. Data Flow in the SERNN Model.

To establish a performance benchmark for the SERNN model, several baseline models were executed for comparative analysis. The baseline LSTM model comprises a single LSTM layer with 60 units, followed by a Dense output layer with one unit, capturing sequential patterns in the data. Like the LSTM, the GRU model includes a single GRU layer with 60 units and a Dense output layer, designed to capture temporal dependencies with potential efficiency gains. The ConvLSTM model consists of a Conv1D layer with 32 filters and a kernel size of 3, followed by a 60-unit LSTM layer, combining convolutional feature extraction with sequential processing. The Bidirectional LSTM model utilizes a bidirectional wrapper around a 60-unit LSTM layer, enabling the model to capture both forward and backward dependencies, thus enhancing temporal pattern recognition. The Stacked LSTM model comprises two LSTM layers with 60 units each, with the first layer returning sequences, thereby enhancing the extraction of temporal features through a deeper structure.

Each model was compiled using the Adam optimizer with MSE as the loss function, ensuring consistency in training parameters across models. This setup allows for a robust comparison framework, highlighting the advantage of the SERNN model's sensitivity input integration over conventional approaches. The comparative analysis of the SERNN and the baseline models underscores the enhanced forecasting accuracy achieved through the inclusion of sensitivity inputs, offering a refined approach to address the dynamic requirements of hosting capacity forecasting in next-generation distribution networks. To compare the forecasting performance of the SERNN with baseline models, all models were simulated with the same dataset and model parameters.

Figure 8 compares the training and validation loss curves for the SERNN model and baseline models, including LSTM, Bi-LSTM, Convolution LSTM, Stacked LSTM, and GRU. Each subfigure highlights the convergence behavior and consistency of the models during the training and validation phases. Among the baseline models, the LSTM and BiLSTM show moderate performance, showing steady convergence but relatively higher loss values. The Convolution LSTM, Stacked LSTM, and GRU demonstrate slower convergence and higher loss values, indicating challenges in capturing the complexities of the dataset. The

SERNN model demonstrates faster convergence and lower loss values in both the training and validation phases. The result indicates the proposed model's capability to respond to dynamic network conditions and intricate characteristics of the DER and weather variables, confirming its effectiveness in forecasting network hosting capacity with higher precision.



(**a**) SERNN (**b**) LSTM (**c**) BiLSTM

(**d**) Convolution LSTM (**e**) Stacked LSTM (**f**) GRU

Figure 8. Training and Validation Loss of Different Models.

Table 5 compares the hosting capacity analysis and forecasting performances of the proposed SERNN model with other deep learning algorithm-based forecasting models like LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM, and GRU. Among other deep learning models, the Bidirectional LSTM performs lowest, while the ConvLSTM performs better. The proposed SERNN outperforms all other models regarding performance indices like MAE, RMSE, and R-squared value. These results highlight that the sensitivity gate and dynamic architecture of the SERNN model allow superior adaptability to time-sensitive variables, significantly improving prediction accuracy and outperforming conventional deep-learning models.

Table 5. Hosting Capacity Analysis and Forecasting Performance Comparison of the Proposed Model with Other Standard Models.

Models	**MAE**	**RMSE**	**R-Squared Value**
Proposed Model			
SERNN	0.2030	0.3884	0.9854
Conventional Deep Learning Models			
LSTM	0.8604	1.6985	0.7214
ConvLSTM	0.6892	1.3664	0.8197
Bidirectional LSTM	0.9861	1.7402	0.7076
Stacked LSTM	0.6889	1.4317	0.8021
GRU	0.8738	1.4218	0.8048

Figure 9 shows the actual and predicted hosting capacity comparison of different models. Compared to baseline forecasting models, the SERNN model demonstrates better performance. The figure illustrates the SERNN model's superior alignment with actual hosting capacity values, showcasing its advanced ability to capture complex dependencies compared to conventional models.

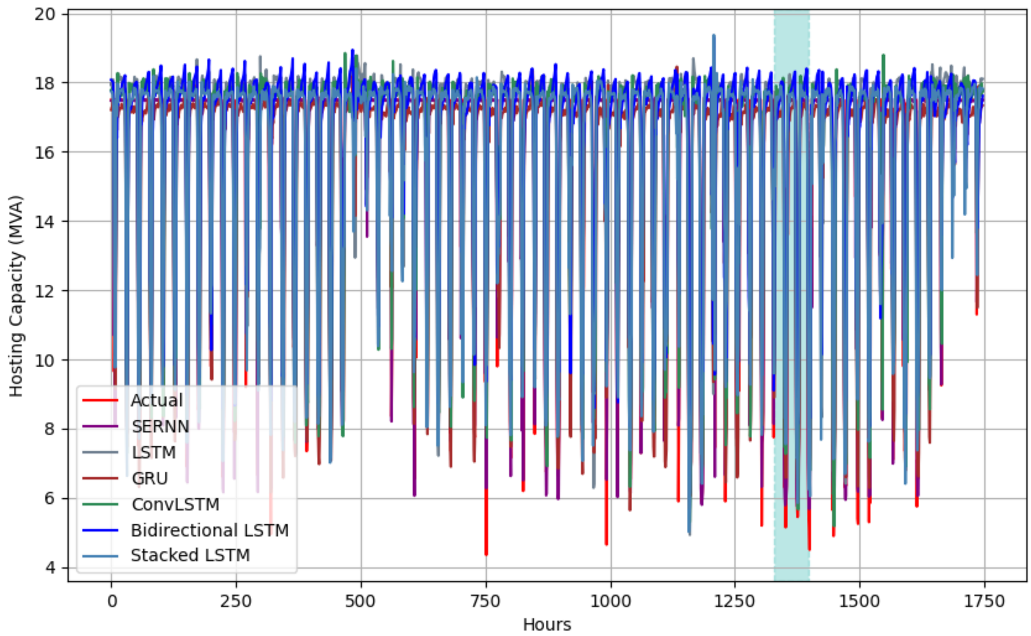


Figure 9. Actual vs. Predicted Hosting Capacity Analysis and Forecasting data.

Figure 10 compares the performance indices of the different forecasting models. It indicates that among LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM and GRU models, the Stacked LSTM gets lower MAE (0.6889) and ConvLSTM achieves lower RMSE (1.3664) and R-squared (0.8197) values. However, the SERNN model outperforms (R-squared: 0.9854) other deep learning algorithm-based forecasting models like LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM, and GRU.

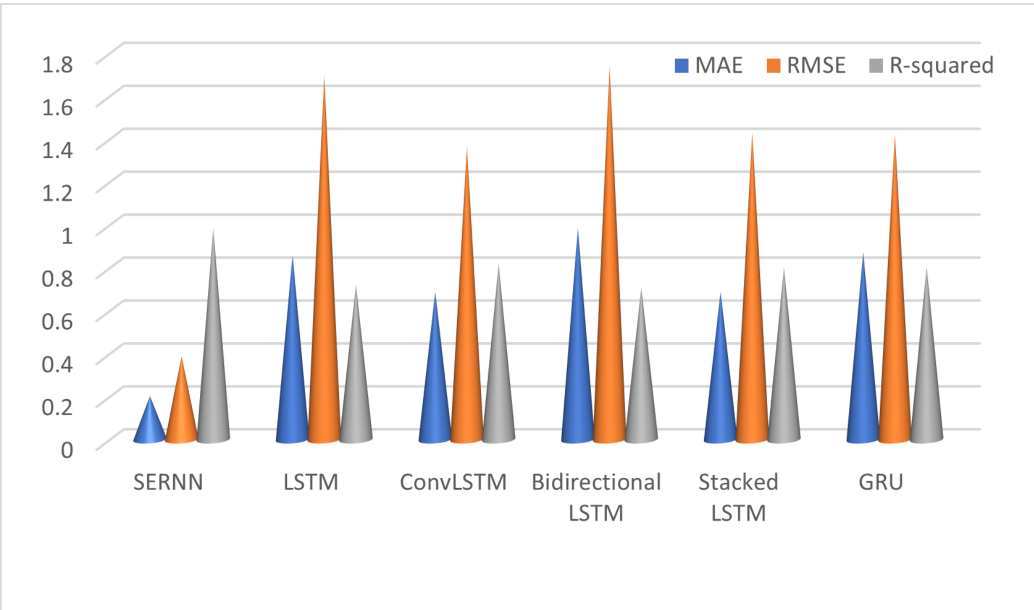


Figure 10. Comparison of Proposed Model Performance with Baseline Models.

The proposed model has focused on the impact of exogenous factors on the model performance. The sensitivity gate captures the impact of both long-term and short-term exogenous feature variables. Also, the previous cell state has modified the forget gate input to capture the long-term dependency on the forecasting horizon. Study 1 shows (Table 3 that introducing new feature variables using feature engineering techniques considerably improves the model's performance. Before applying the feature engineering techniques, the model shows the MAE, RMSE, and R-squared values 0.5501, 0.9408, and 0.9145. However, introducing new features using feature engineering techniques improves the model's performance indices MAE, RMSE, and R-squared values 0.2322, 0.4085, and 0.9839, respectively. Study 2 reveals that different sensitivity gate inputs impact the model's hosting capacity prediction performance differently. Figure 6) shows that the time of the day ('Hour') as sensitivity gate input provides a better model performance, achieving the MAE, RMSE, and R-squared values 0.2030, 0.3884, and 0.9854, respectively. Moreover, the sensitivity gate inputs like 'Max_V' and 'Min_V' have a moderate impact, while the 'PV_KW' has the lowest influence on the model performance among all sensitivity gate inputs. Compared to the standard models like LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM and GRU, the proposed SERNN model outperforms in terms of accuracy for hosting capacity analysis and forecasting of the distribution network with DER. The results of Case Study 3 demonstrate the model's superior performance compared to the standard models.

The sensitivity gate dynamically adjusts the model response to contextual variables. PV and BESS output is directly correlated with the time of the day ('Hour'). In this study, 'Hour' was used as sensitivity input to capture the temporal dependencies and dynamic characteristics of the network and DER variables. However, the sensitivity input may need to be redesigned for other DER resources like wind energy. The flexibility of the modular design of the sensitivity gate supports input selection based on the critical variable of the DER, allowing it to adapt to various distributed network scenarios.

4. Conclusions

Reliable integration and active management of the DER with distribution networks necessitates precise forecasting of network hosting capacity. Beyond the inherent complexities of network and DER variables, accurate prediction of hosting capacity depends on various exogenous factors. The study addresses the challenges of the exogenous variables and the dynamic nature of the network and DER variables by introducing a sensitivity gate within a novel memory cell architecture, enabling dynamic adjustments, enhancing adaptability and forecasting accuracy. Moreover, by incorporating the previous cell state into the forget gate input, the model effectively captures long-term dependencies, significantly impacting the model output. The study also investigates the role of feature engineering in enhancing the accuracy of hosting capacity forecasts. Findings suggest that the timing of the network analysis for estimating hosting capacity can significantly enhance model forecasting precision.

The case studies demonstrate the model's superior performance in hosting capacity forecasting. Incorporating advanced feature engineering significantly enhances the model's performance, reducing the MAE from 0.5501 to 0.2030, RMSE from 0.9408 to 0.3884 and increasing the R-squared value from 0.9145 to 0.9854. These enhancements validate the critical role of feature engineering techniques in improving model performance. Furthermore, incorporating the sensitivity gate further enhanced the model's ability to capture time-sensitive and contextual variability, reducing MAE to 0.2030 and RMSE to 0.3884, with an R-squared value of 0.9854, compared to 0.2283, 0.4570, and 0.9152, respectively, without the sensitivity gate. The sensitivity gate demonstrated its ability to dynamically adjust responses to contextual variables like time of the day, significantly improving forecasting accuracy. The results also revealed the critical role of sensitivity gate inputs, with the 'Hour' variable outperforming others by achieving the lowest MAE (0.2030) and RMSE (0.3884) and the highest R-squared value (0.9854). This demonstrates the model's adaptability in capturing temporal dependencies and contextual variability, making it suitable for dynamic scenarios in distribution networks. The study shows that the proposed model, enhanced with a sensitivity gate, outperforms traditional models like LSTM, ConvLSTM, Bidirectional LSTM, Stacked LSTM, and GRU. The SERNN achieved an R-squared value of 0.9854, significantly higher than ConvLSTM (0.8197) and LSTM (0.7214), while reducing RMSE by 76.4 per cent and MAE by 70.5 per cent relative to LSTM. These results underline the SERNN model's ability to capture complex dependencies and adapt to dynamic network conditions, making it a robust tool for forecasting network hosting capacity.

This research provides valuable insights for researchers, network operators and engineers. The innovative sensitivity gate mechanism substantially enhances the accuracy of network hosting capacity forecasts, facilitating more effective active management of integrated DER. This advancement supports more intelligent planning and reliable integration of DER into next-generation distribution networks. Future research could focus on extending the SERN model by integrating other types of DER like wind power and advanced storage systems. Future research can explore extending the SERNN model to incorporate other DER types, such as wind power and other renewable resources, and optimizing hyperparameters to enhance accuracy further. Testing the model under various distribution network scenarios, including tariff structures and pricing modalities, could generalize its robustness, making it a valuable tool for grid operators managing next-generation distribution networks.

Author Contributions: All authors have contributed equally, read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, L.; Yang, Z.; Xiao, Q.; Guo, Y.; Ying, Z.; Hu, T.; Xu, X.; Khan, S.; Li, K. Distributed scheduling for multi-energy synergy system considering renewable energy generations and plug-in electric vehicles: A level-based coupled optimization method. *Energy AI* **2024**, *16*, 100340. [CrossRef]
2. Mulenga, E.; Bollen, M.H.; Etherden, N. Adapted stochastic PV hosting capacity approach for electric vehicle charging considering undervoltage. *Electricity* **2021**, *2*, 387–402. [CrossRef]
3. Strezoski, L.; Padullaparti, H.; Ding, F.; Baggu, M. Integration of Utility Distributed Energy Resource Management System and Aggregators for Evolving Distribution System Operators. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 277–285. [CrossRef]
4. Ismael, S.M.; Aleem, S.H.A.; Abdelaziz, A.Y.; Zobaa, A.F. State-of-the-art of hosting capacity in modern power systems with distributed generation. *Renew. Energy* **2019**, *130*, 1002–1020. [CrossRef]
5. Paliwal, P.; Patidar, N.; Nema, R. Planning of grid integrated distributed generators: A review of technology, objectives and techniques. *Renew. Sustain. Energy Rev.* **2014**, *40*, 557–570. [CrossRef]
6. Zain ul Abideen, M.; Ellabban, O.; Al-Fagih, L. A review of the tools and methods for distribution networks' hosting capacity calculation. *Energies* **2020**, *13*, 2758. [CrossRef]
7. Islam, M.T.; Hossain, M. Artificial Intelligence for Hosting Capacity Analysis: A Systematic Literature Review. *Energies* **2023**, *16*, 1864. [CrossRef]
8. Munikoti, S.; Abujubbeh, M.; Jhala, K.; Natarajan, B. A novel framework for hosting capacity analysis with spatio-temporal probabilistic voltage sensitivity analysis. *Int. J. Electr. Power Energy Syst.* **2022**, *134*, 107426. [CrossRef]
9. Aydin, M.S. Strategies for increasing hosting capacity in PV-rich LV feeders via radiality-imposed reconfiguration. *Sustain. Energy Grids Netw.* **2024**, *38*, 101288. [CrossRef]
10. Taheri, S.I.; Salles, M.B.; Nassif, A.B. Distributed energy resource placement considering hosting capacity by combining teaching—Learning-based and honey-bee-mating optimisation algorithms. *Appl. Soft Comput.* **2021**, *113*, 107953. [CrossRef]
11. Chihota, M.J.; Bekker, B.; Gaunt, T. A stochastic analytic-probabilistic approach to distributed generation hosting capacity evaluation of active feeders. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107598. [CrossRef]
12. Barbosa, T.; Andrade, J.; Torquato, R.; Freitas, W.; Trindade, F.C. Use of EV hosting capacity for management of low-voltage distribution systems. *IET Gener. Transm. Distrib.* **2020**, *14*, 2620–2629. [CrossRef]

13. Sugihara, H.; Funaki, T. Increasing electric vehicle hosting capacity and equality for fast charging stations using residential photovoltaics in medium and low-voltage distribution networks. *IEEJ Trans. Electr. Electron. Eng.* **2020**, *15*, 364–371. [CrossRef]
14. Fokui, W.S.T.; Saulo, M.J.; Ngoo, L. Optimal placement of electric vehicle charging stations in a distribution network with randomly distributed rooftop photovoltaic systems. *IEEE Access* **2021**, *9*, 132397–132411. [CrossRef]
15. Zou, M.; Yang, Y.; Liu, M.; Wang, W.; Jia, H.; Peng, X.; Su, S.; Liu, D. Optimization Model of Electric Vehicles Charging and Discharging Strategy Considering the Safe Operation of Distribution Network. *World Electr. Veh. J.* **2022**, *13*, 117. [CrossRef]
16. Mulenga, E. Solar PV Stochastic Hosting Capacity Assessment Considering Epistemic (E) Probability Distribution Function. *Electricity* **2022**, *3*, 586–599. [CrossRef]
17. Home-Ortiz, J.M.; Melgar-Dominguez, O.D.; Mantovani, J.R.S.; Catalao, J.P. PV hosting capacity assessment in distribution systems considering resilience enhancement. *Sustain. Energy Grids Netw.* **2022**, *32*, 100829. [CrossRef]
18. Rawat, T.; Singh, J.; Pandey, V.C.; Sharma, S. Planning renewables in distribution system with electric vehicles to improve hosting capacity and energy losses: Two-stage stochastic optimization framework. *Sustain. Energy Grids Netw.* **2024**, *39*, 101445. [CrossRef]
19. Fachrizal, R.; Ramadhani, U.H.; Munkhammar, J.; Widén, J. Combined PV–EV hosting capacity assessment for a residential LV distribution grid with smart EV charging and PV curtailment. *Sustain. Energy Grids Netw.* **2021**, *26*, 100445. [CrossRef]
20. Rabiee, A.; Keane, A.; Soroudi, A. Enhanced transmission and distribution network coordination to host more electric vehicles and PV. *IEEE Syst. J.* **2021**, *16*, 2705–2716. [CrossRef]
21. Trinh, P.H.; Chung, I.Y. Integrated Active and Reactive Power Control Methods for Distributed Energy Resources in Distribution Systems for Enhancing Hosting Capacity. *Energies* **2024**, *17*, 1642. [CrossRef]
22. Wu, J.; Yuan, J.; Weng, Y.; Ayyanar, R. Spatial-Temporal Deep Learning for Hosting Capacity Analysis in Distribution Grids. *IEEE Trans. Smart Grid* **2022**, *14*, 354–364. [CrossRef]
23. Breker, S.; Rentmeister, J.; Sick, B.; Braun, M. Hosting capacity of low-voltage grids for distributed generation: Classification by means of machine learning techniques. *Appl. Soft Comput.* **2018**, *70*, 195–207. [CrossRef]
24. Dong, C.; Yu, T.; Pan, Z.; Wu, Y.; Wang, Z.; Wang, Y.; Wang, K. Renewable energy hosting capacity assessment in distribution networks based on multi-strategy improved whale optimization algorithm. *IET Renew. Power Gener.* **2024**. [CrossRef]
25. Qammar, N.; Arshad, A.; Miller, R.J.; Mahmoud, K.; Lehtonen, M. Machine learning based hosting capacity determination methodology for low voltage distribution networks. *IET Gener. Transm. Distrib.* **2024**, *18*, 911–920. [CrossRef]
26. Zulu, E.; Hara, R.; Kita, H. An efficient hybrid particle swarm and gradient descent method for the estimation of the hosting capacity of photovoltaics by distribution networks. *Energies* **2023**, *16*, 5207. [CrossRef]
27. Suchithra, J.; Rajabi, A.; Robinson, D.A. Enhancing PV hosting capacity of electricity distribution networks using deep reinforcement learning-based coordinated voltage control. *Energies* **2024**, *17*, 5037. [CrossRef]
28. Lee, K.; Im, S.; Lee, B. Prediction of renewable energy hosting capacity using multiple linear regression in KEPCO system. *Energy Rep.* **2023**, *9*, 343–347. [CrossRef]
29. Islam, M.T.; Hossain, M.; Habib, M.A. Machine Learning-based Hosting Capacity Analysis and Forecasting in Low-Voltage Networks. In Proceedings of the 2023 IEEE International Future Energy Electronics Conference (IFEEC), Sydney, Australia, 20–23 November 2023; pp. 461–464.
30. Procopiou, A.T.; Liu, M.Z.; Ochoa, L.F.; Langstaff, T.; Harding, J. Smart meter-driven estimation of PV hosting capacity. In Proceedings of the CIRED 2020 Berlin Workshop (CIRED 2020), IET, Berlin, Germany, 22–23 September 2020; Volume 2020, pp. 128–131.
31. Azzolini, J.A.; Reno, M.J.; Yusuf, J.; Talkington, S.; Grijalva, S. Calculating PV hosting capacity in low-voltage secondary networks using only smart meter data. In Proceedings of the 2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 16–19 January 2023, pp. 1–5.
32. Suchithra, J.; Robinson, D.A.; Rajabi, A. A Model-Free Deep Reinforcement Learning-Based Approach for Assessment of Real-Time PV Hosting Capacity. *Energies* **2024**, *17*, 2075. [CrossRef]
33. Bassi, V.; Ochoa, L.F.; Alpcan, T.; Leckie, C. Electrical model-free voltage calculations using neural networks and smart meter data. *IEEE Trans. Smart Grid* **2022**, *14*, 3271–3282. [CrossRef]
34. Talkington, S.; Grijalva, S.; Reno, M.J.; Azzolini, J.A.; Pinney, D. A measurement-based approach to voltage-constrained hosting capacity analysis with controllable reactive power behind-the-meter. *Electr. Power Syst. Res.* **2023**, *221*, 109395. [CrossRef]
35. Klonari, V.; Toubeau, J.F.; Lobry, J.; Vallée, F. Estimating the photovoltaic hosting capacity of a low voltage feeder using smart meters' measurements. In *Smart Metering Technology & Services*; Intech Book; IntechOpen: London, UK, 2016; pp. 131–158.
36. Zhou, Y.; Yang, Z.; Sun, Q.; Yu, C.; Yu, C. An artificial intelligence model based on multi-step feature engineering and deep attention network for optical network performance monitoring. *Optik* **2023**, *273*, 170443. [CrossRef]
37. EPRI Home. 2024. Available online: https://www.epri.com/ (accessed on 6 October 2024).
38. Kasturi, K.; Nayak, C.K.; Nayak, M.R. Electric vehicles management enabling G2V and V2G in smart distribution system for maximizing profits using MOMVO. *Int. Trans. Electr. Energy Syst.* **2019**, *29*, e12013. [CrossRef]

39. Kourentzes, N.; Crone, S.F. Frequency independent automatic input variable selection for neural networks for forecasting. In Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN), IEEE, Barcelona, Spain, 18–23 July 2010; pp. 1–8.
40. Habib, M.A.; Hossain, M. Revolutionizing Wind Power Prediction—The Future of Energy Forecasting with Advanced Deep Learning and Strategic Feature Engineering. *Energies* **2024**, *17*, 1215. [CrossRef]
41. Rodríguez, P.; Bautista, M.A.; Gonzalez, J.; Escalera, S. Beyond one-hot encoding: Lower dimensional target embedding. *Image Vis. Comput.* **2018**, *75*, 21–31. [CrossRef]
42. Sharma, S.; Sharma, S.; Athaiya, A. Activation functions in neural networks. *Towards Data Sci.* **2017**, *6*, 310–316. [CrossRef]
43. Ghimire, S.; Nguyen-Huy, T.; AL-Musaylh, M.S.; Deo, R.C.; Casillas-Pérez, D.; Salcedo-Sanz, S. Integrated Multi-Head Self-Attention Transformer model for electricity demand prediction incorporating local climate variables. *Energy AI* **2023**, *14*, 100302. [CrossRef]
44. Guo, J. *Backpropagation Through Time*; Harbin Institute of Technology: Harbin, China, 2013; Volume 40, pp. 1–6.
45. Habib, M.A.; Hossain, M.; Alam, M.M.; Islam, M.T. A hybrid optimized data-driven intelligent model for predicting short-term demand of distribution network. *Sustain. Energy Technol. Assess.* **2024**, *67*, 103818. [CrossRef]