IET Renewable Power Generation

WILEY

The Institution of Engineering and Technology

ORIGINAL RESEARCH OPEN ACCESS

Advanced Harmonic Forecasting in Offshore Wind Farms with Permanent Magnet Synchronous Generators Using a Hybrid Deep and Machine Learning Architecture

Alp Karadeniz

Department of Electrical-Electronic Engineering, Balikesir University, Balıkesir, Türkiye

Correspondence: Alp Karadeniz (akaradeniz@balikesir.edu.tr)

Received: 1 June 2025 | **Revised:** 18 August 2025 | **Accepted:** 15 September 2025

Funding: The author received no specific funding for this work.

ABSTRACT

Wind energy is crucial for reducing fossil fuel dependence and promoting sustainability. Offshore wind farms (OWFs) benefit from higher, stable wind speeds but pose challenges such as harmonic distortion and voltage fluctuations when integrated into power grids. This study develops an advanced model for accurate harmonic forecasting in OWFs using permanent magnet synchronous generators (PMSG). Real meteorological data from Zonguldak and Sinop in the Black Sea region of Turkey were used to simulate power output, voltage, and current waveforms. Harmonic components, including total harmonic distortion for voltage (THDV) and current (THDI), were extracted and predicted. Various machine learning (ML) and deep learning (DL) algorithms were applied, including Linear Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, KNeighbors, LSTM, GRU, and CNN. Additionally, hybrid ML-DL models were explored to enhance forecasting accuracy. A comparative analysis of these models demonstrated their effectiveness in improving harmonic prediction. Results indicate that hybrid models, particularly LSTM+GB and GRU+GB, improve harmonic forecasting accuracy by reducing RMSE by approximately 15% compared to traditional ML methods. This enhancement contributes to better power quality management and grid stability, making offshore wind farms more viable for large-scale renewable energy integration. The findings of this research provide a fundamental basis for future investigations into offshore wind harmonic forecasting.

1 | Introduction

As the world strives for carbon neutrality, renewable energy is becoming increasingly important. Offshore wind, known for its abundance and reliability, is at the forefront of this transition. Energy companies are investing heavily in renewables, and policymakers are supporting the adoption of new technologies. Offshore wind is growing rapidly because of its greater energy potential and less intrusive appearance relative to onshore wind farms [1, 3]. The rise of floating wind turbines is making previously inaccessible wind resources available, further fueling the expansion of offshore wind. Both the number and size of offshore wind farms (OWFs) continue to increase. Wind power plants (WPPs) play a key role in delivering high-quality energy. Ensuring stable voltage and frequency, minimal disturbances, and low harmonics is essential for system reliability [4, 5]. However, integrating WPPs into the main grid raises concerns about harmonic distortion [6]. Type 4 wind turbine generators (WTGs) use controlled back-to-back power electronic converters, which introduce harmonics. These harmonics can cause transformer

Abbreviations: *ML*, machine learning; *DL*, deep learning; *OWF*, offshore wind farms; *WPP*, wind power plants; *WTG*, wind turbine generators; *PMSG*, permanent magnet synchronous generators.

overheating, circuit breaker tripping, and protection device malfunctions. They also reduce equipment lifespan. Minimizing harmonics is crucial for power quality, following IEEE 519-2014 and IEC 61000 standards [7, 8]. Indices such as THDV and THDI measure voltage and current harmonics.

Furthermore, forecasting plays a critical role in mitigating harmonics during the integration of large-scale offshore wind farms. It facilitates the design of harmonic mitigation solutions to ensure compliance with IEEE-519 standards. The objective is to create a novel forecasting model that provides precise and dependable harmonic predictions for offshore wind farms. PMSG configurations are used for this purpose. Real-world meteorological data from Zonguldak and Sinop, two locations in Turkey's Black Sea region, serve as input [9]. Also, the Black Sea presents a high potential for offshore wind energy development, with studies estimating Turkey's technical offshore wind capacity in the region to be 90 GW [10]. Its unique meteorological conditions, including stable wind speeds and minimal wave heights compared to other seas, make it a strategic location for offshore wind farm deployment [11]. Additionally, the World Bank has identified the Black Sea as having 435 GW of total offshore wind potential, emphasizing the need for investment in infrastructure and grid integration [12]. Given these factors, this study focuses on the Black Sea as a representative case study for offshore wind harmonics forecasting and power quality assessment.

Furthermore, wind speed data is used to simulate output power and generate current and voltage waveforms. Harmonics are extracted from these waveforms for analysis, arrangement, and future forecasting. Total harmonic distortion (THD) is specifically predicted for current and voltage. Different artificial intelligence models are applied for THD prediction. Machine learning models include Linear Regression, Decision Tree, Random Forest, Gradient Boosting, AdaBoost, XGBoost, and KNeighbors Regressor. Deep learning models such as long short-term memory, gated recurrent unit, and convolutional neural network are also utilized.

Additionally, to improve prediction performance, hybrid models were introduced. These models integrate machine learning (ML) techniques [13], using outputs from deep learning (DL) models [14, 17] as inputs. This study compared the effects of different artificial intelligence techniques on THDV and THDI estimation. Weather records collected from two locations in Turkey's Black Sea region were utilized. The results highlight the impact of various algorithms on the precision of the forecasts.

This study began with an introduction, discussing the global shift toward carbon neutrality and the power quality challenges of offshore wind farm integration. The literature overview summarized the current state of knowledge. The motivation chapter explained the study's significance. The data preparation chapter covered offshore wind farm simulation, data generation, and harmonic analysis using meteorological data. Deep learning and machine learning models were discussed in relation to harmonic prediction. Experiments and results were presented, including comparative analyses of these models. The conclusion emphasized the potential of hybrid models for accurate harmonic forecasts. The study ended with a references section listing the cited sources.

2 | Literature Overview

In the literature, several studies are presented about harmonic forecasting for power generation systems [18, 26]. In one of the referenced studies [18], a methodology is introduced for forecasting voltage THD using data obtained from a limited number of smart meters. This approach enables the provision of power quality indices to operators without requiring additional infrastructure. Various forecasting algorithms, including artificial neural networks, are applied and compared. The effectiveness of the proposed method is validated using measurements from household devices on a European low-voltage test feeder. Also, in [19], the study presents six hybrid forecasting models combining artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS) for precise harmonics prediction in renewable energy systems (RES). Utilizing real-world data from generators with double-fed induction generator (DFIG) and PMSG, the models aim to improve harmonics forecasting. Results show that model-3 and model-6 are the most effective and consistent in predicting harmonics.

Moreover, the study [20] develops a novel forecasting model for harmonics in RES using wind turbines with DFIG and PMSG combined with Solar PV. Real-world meteorological data are used to simulate and analyze harmonics in voltage and current waveforms. Results show that ANFIS is the most effective for forecasting Total Harmonic Distortion and the 13th harmonic, while 3LCRNNGL excels for the 11th harmonic in both generator models.

Additionally, the study [21] develops a hybrid forecasting model combining the ANFIS and long short-term memory (LSTM) network for precise harmonics prediction in renewable energy systems. Two methods are proposed: one with LSTM first and ANFIS second, and the other with the reverse order. Using harmonics data from DFIG and PMSG systems, the ANFIS-LSTM model consistently outperforms other techniques with lower RMSE values, demonstrating superior accuracy in harmonics forecasting.

Also, another study [22] presents a probabilistic harmonic power flow method that accounts for uncertainties in distributed energy resources (DER) and electrical loads to forecast harmonic distortion. The method uses time-varying stochastic characteristics and a constant-weight point estimate approach based on the Nataf transformation for efficient estimation. Validated with real-field data, this approach effectively analyzes harmonic coupling and provides accurate forecasts of harmonic distortion.

In addition, a study [23] introduces a model for predicting voltage harmonics in wind energy production using measurement data obtained from the Jeffrey's Bay Wind Energy Facility. It utilizes LSTM networks, a type of recurrent neural network (RNN), to analyze 8103 measured voltage harmonics and extract key features. The model, with a single LSTM layer of 128 neurons, effectively forecasts 3800 sample mean values with low root mean square error (RMSE), ensuring cleaner voltage integration into the national grid.

Also, in [24], the study reviews traditional and modern harmonic estimation techniques for power systems affected by

TABLE 1 | Pros & cons of recent studies.

Recent Studies	Pros	Cons
P. Rodríguez-Pajarón et al. (2022)	Utilizes existing infrastructure, practical validation.	Limited to data from smart meters, potentially less generalizable.
F. M. A. Hadi et al. (2023)	Effective and consistent in predictions, hybrid approach improves accuracy.	Complexity in model design and implementation.
H. H. Aly et al. (2023)	High accuracy for specific harmonics, robust model.	May require significant computational resources.
F. M. A. Hadi et al. (2024)	Lower RMSE values, superior accuracy.	Model complexity, potential for overfitting.
Yahui Li et al. (2022)	Accounts for uncertainties, validated with real-field data.	Complexity in implementation, reliance on specific data characteristics.
E. M. Kuyunani et al. (2021)	Low RMSE, effective for large datasets.	Single LSTM layer might limit model flexibility.
Thamer A. H. et al. (2023)	High accuracy, low error rates.	Simulation-based, potential discrepancies in real-world applications.
Y. Chaoying et al. (2024)	High accuracy, improves data coverage.	Complexity in model integration, data requirements.

non-linear devices that cause power quality issues such as harmonics, flicker, and transient instabilities. Harmonics lead to significant challenges like equipment overheating and failure, emphasizing the need for effective mitigation methods. The study highlights the role of artificial intelligence (AI) in improving harmonic estimation, which reduces the reliance on traditional methods requiring complex system parameters and monitors.

In another study [25], the paper proposes an ANN method for estimating harmonic distortions from solar PV inverters in grid-connected systems. The method, incorporating location-specific data, is validated through simulations of a simple power system and tested on the IEEE 34-bus feeder. It demonstrates high accuracy, with individual harmonic component errors below 10% and a median error of 5.4%. Lastly, the study [26] presents a hybrid PCA-ELM-LSTM model for accurate harmonic prediction in power grids, addressing the challenge of sparse harmonic monitoring data. The model integrates principal component analysis (PCA) for dimensionality reduction, LSTM networks for historical information memory, and extreme learning machines (ELM) for enhanced learning accuracy. Validation with actual power grid data demonstrates that this approach provides precise harmonic predictions, improving data coverage and analysis.

Unlike previous studies in Table 1, which primarily focus on onshore wind farms or generalized renewable energy systems, this study specifically targets offshore wind farms utilizing PMSGs with long submarine cables, a critical yet underexplored aspect of offshore wind integration. Furthermore, while prior research often employs traditional harmonic analysis techniques, this study integrates hybrid algorithms to improve THD forecasting accuracy. By leveraging real-world meteorological data from the Black Sea region, our approach also expands offshore wind forecasting research into a geographically underrepresented area, contributing to a more robust understanding of power quality management in offshore energy systems.

3 | Motivation

The literature indicates a growing interest in harmonic forecasting for renewable energy systems, particularly in wind and solar power applications. However, most existing studies have concentrated on onshore wind farms or photovoltaic systems. In contrast, the present study uniquely addresses harmonic prediction for an offshore wind energy system utilizing a PMSG and connected via a long-distance submarine cable. Two specific coastal locations, Zonguldak and Sinop, located along the Black Sea, were selected as representative sites for meteorological data collection. Although offshore wind farms have not yet been deployed in Turkey, several recent studies [27, 32] have investigated the region's offshore wind energy potential.To better contextualize the challenges addressed in this study, it is important to note that the primary difficulty lies in the accurate forecasting of harmonic distortion in offshore wind farms, specifically using PMSGs. Unlike traditional onshore wind applications, offshore wind farms with PMSG units and long-distance submarine cable connections introduce complex harmonic patterns due to high frequency switching harmonics from power electronic converters. Additionally, real-world meteorological variability poses challenges for precise harmonic prediction. With future OWF development expected in the country, this study contributes a valuable harmonic forecasting framework based on real-world wind data, offering practical insights for planning and grid integration of offshore systems.

3.1 | Contributions to the Knowledge

This study makes several significant contributions to the field of harmonic forecasting in OWFs, particularly in the context of Turkey. The key contributions are summarized as follows:

- Focus on Offshore Systems: Unlike previous research, which primarily addresses harmonic forecasting for onshore wind and photovoltaic (PV) systems, this study targets offshore

PMSG systems. By focusing on offshore wind farms with long-distance submarine cables, it fills a gap in the literature concerning harmonic analysis in these specific conditions.

- Application to Black Sea Region: The study uses Zonguldak and Sinop as sample locations within the Black Sea region, offering valuable insights into the harmonic characteristics specific to this area. This is particularly relevant given the lack of existing offshore wind turbine applications in Turkey.
- Integration of Realistic Meteorological Data: By incorporating real-world meteorological wind speed observations provided by the EU Science Hub, the study enhances the accuracy of harmonic forecasting models. This approach ensures that the harmonic predictions are based on realistic and representative data, improving the reliability of the forecasts.
- Comparative Analysis of ML and DL Techniques: The study compares a range of machine learning (ML) and deep learning (DL) techniques, including hybrid models, to determine the most effective methods for harmonic forecasting. This comprehensive analysis contributes to understanding which techniques offer the best performance in predicting total harmonic distortion (THDV and THDI) for offshore wind farms.
- Future Implications for Offshore Wind in Turkey: As Turkey explores the potential for offshore wind power in the Black Sea, this study provides critical insights into harmonic estimation models. The findings offer a foundation for future research and practical applications in the development of offshore wind farms in Turkey.

Overall, this research not only presents an innovative method for harmonic prediction utilizing sophisticated data enhancement and machine learning strategies but also addresses a significant gap in the existing literature concerning the use of these techniques in offshore wind farms in Turkey. The knowledge acquired from this study is anticipated to influence and direct forthcoming investigations in this field.

4 | Data Preparation

This section introduces a model of the OWF designed to replicate actual wind conditions. Figure 1 shows the schematic layout of the analyzed system, which features Type 4 PMSG turbines within the offshore wind farm. The setup includes an electrical grid operating at 230 kV, a 230/60 kV transformer with wye-delta connection rated at 50 MVA, a 60 kV/575 V transformer with delta-wye configuration rated at 10 MVA, and a 36 kilometer long underwater line within the offshore energy system.

In the simulation modeled, an OWF with five turbines, each with a 2 MW capacity, totaling 10 MW, and is connected to a 60 kV distribution network. The generated electricity is transmitted to a 230 kV grid via a 36 km-long submarine cable. Each wind turbine employs a PMSG, which incorporates an induction generator, a DC-DC boost converter, and an IGBT-based pulse width modulation (PWM) converter. The stator winding is directly linked to the 60 Hz grid, whereas the rotor receives a variable frequency supply via the DC-DC boost converter. A detailed schematic of the wind turbine under mask is presented in Figure 2.

The system includes two transformers. The first transformer has a 230 kV to 60 kV wye/delta configuration with a power rating of 50 MVA. The per-unit values for its winding resistances, leakage inductances, and magnetization inductance are 2.66×10^{-3}, 0.08, and 500 p.u., respectively. The following transformer, with a delta/wye configuration, operates at 60 kV/575 V and has a power rating of 10 MVA. The per-unit values for its winding resistances, leakage inductances, and magnetization inductance are 2.666×10^{-3}, 0.008, and 500 p.u., respectively.

In this research, harmonic analysis was performed utilizing the typical meteorological year (TMY) dataset provided by Joint Research Centre (JRC) of the European Commission [9]. This set of data spans from July 2020 through March 2005, offering around 15 years of hourly meteorological records. It contains multiple variables, including wind speed and solar radiation. The previously described model was employed to simulate voltage and current signals, which served as input data for training forecasting methods. Additionally, two pivot points, Zonguldak and Sinop, were selected as sample data locations within the Black Sea region. The coordinates of these points are 41.730° latitude and 31.721° longitude for the Zonguldak area, and 42.090° latitude and 335.241° longitude for the Sinop area. The coordinate locations of the two pivot points are presented in Figures 3 and 4. These two pivot points were selected due to their wind speed values being most suitable for offshore wind farm installation.

The electrical grid was represented as a standard distribution network and simulated for a duration of 100 s. This simulation reflects a scaled version of data collected between March 2005 and July 2020. 8,941 data points were linearly adjusted to fit within the 100-s timeframe, with each point corresponding to approximately 0.0112 s. To emulate actual wind behavior and its effects on the system, the OWF model was implemented to generate related voltage harmonics. Harmonic components appeared in the voltage waveforms, with analysis primarily focusing on the phase 1 voltage due to the system's balanced characteristics. Additionally, harmonic extraction was carried out using fast Fourier transform (FFT). The FFT window spanned two cycles, continuously sampling voltage waveforms over the entire simulation period, yielding 8,941 samples. THD values for both voltage and current signals were calculated based on the simulated results. After the simulation, the harmonic-contaminated 575 V bus voltage and current waveforms for the Sinop and Zonguldak regions, along with their harmonic spectra, are shown in Figures 5–7 and 8. The associated THDV and THDI metrics are also illustrated in Figures 9–11 and 12.

For the Sinop area, THDV and THDI values are 2.04% and 4.83%. Also, THDV and THDI values are 2.05% and 5.04% for Zonguldak area. Also, as can be seen from Figure 10 and Figure 12, the system has higher harmonic (2860 Hz and 3120 Hz) values because of OWF's insulated gate bipolar transistor (IGBT)-based PWM converter switching frequency (3000 Hz).

Eventually, the outcomes presented here are derived from the meteorological data sourced from the EU Science HUB, including wind speed. Furthermore, these results are treated as ground truth within the study and are used as a benchmark for comparison following the training of machine learning and deep learning models.

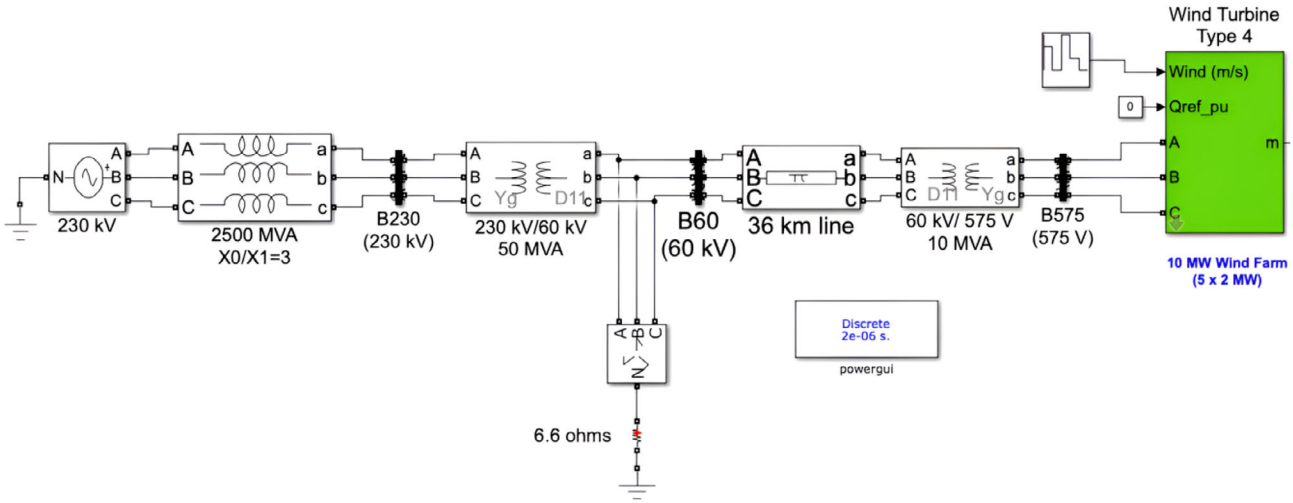

FIGURE 1 | The Simulink model represents the analyzed system with Type 4 PMSG turbines in the offshore wind farm.

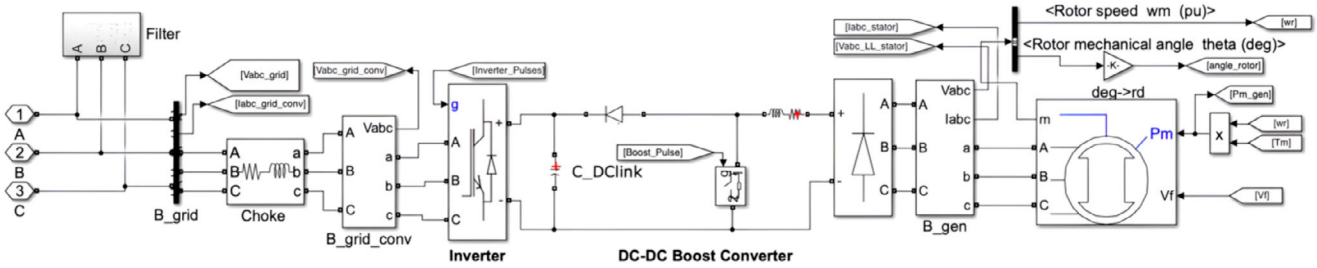

FIGURE 2 | Illustration of the PMSG wind turbine block under the mask.

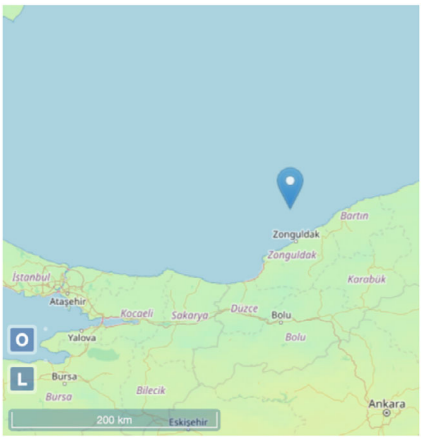

FIGURE 3 | The pivot point of Zonguldak area.

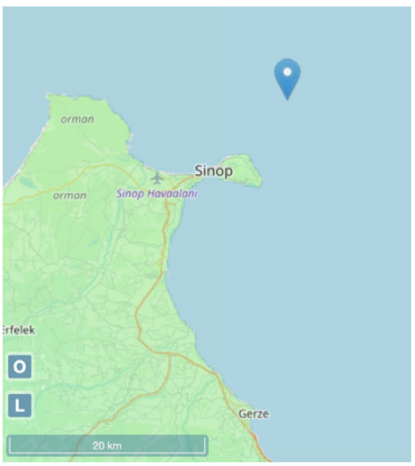

FIGURE 4 | The pivot point of Sinop area.

Also, it should be noted, figures highlight the presence of higher-order harmonics, specifically at 2860 Hz and 3120 Hz, which are characteristic of switching harmonics generated by power electronic converters in offshore wind farms. While these harmonics contribute to overall THDV, their direct impact on forecasting accuracy is minimized by the modeling approach used in this work. The machine learning (ML) and deep learning (DL) models are designed to predict THDV as an aggregate measure, rather than isolating individual harmonic frequencies. Since THDV inherently encapsulates contributions from all harmonic orders, the forecasting models implicitly account for their effect, making a frequency-specific breakdown unnecessary. Moreover, previous studies have shown that higher-order harmonics tend to be less dominant in influencing system-wide power quality compared to lower-order harmonics (e.g., 5th, 7th, 11th) [33]. While further investigation into frequency-specific forecasting could be explored in future work, the current approach provides a comprehensive and practical framework for offshore wind harmonic prediction.

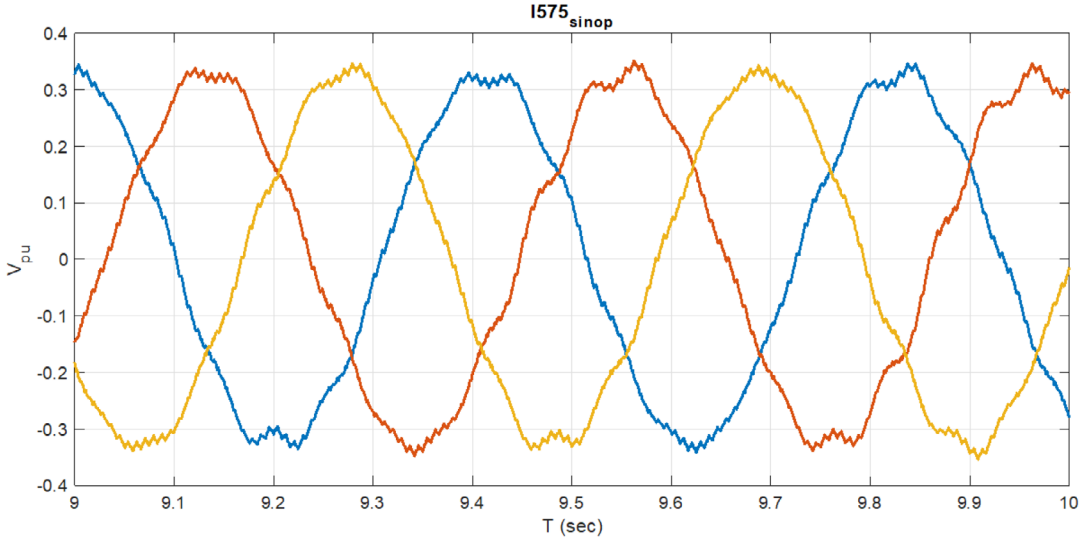

FIGURE 5 | Current waves at the 575 V bus for the Sinop.

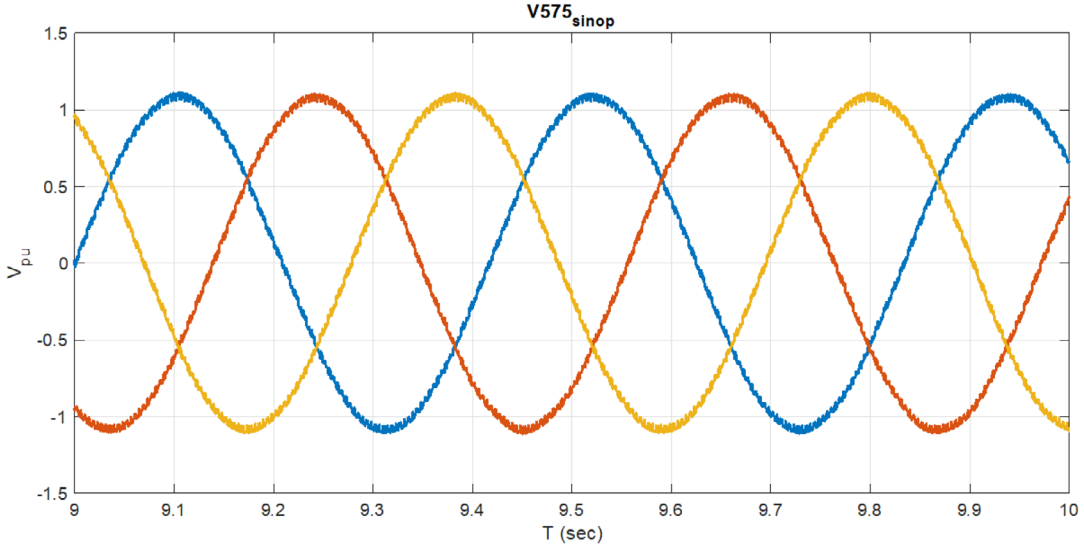

FIGURE 6 | Voltage waves at the 575 V bus for the Sinop.

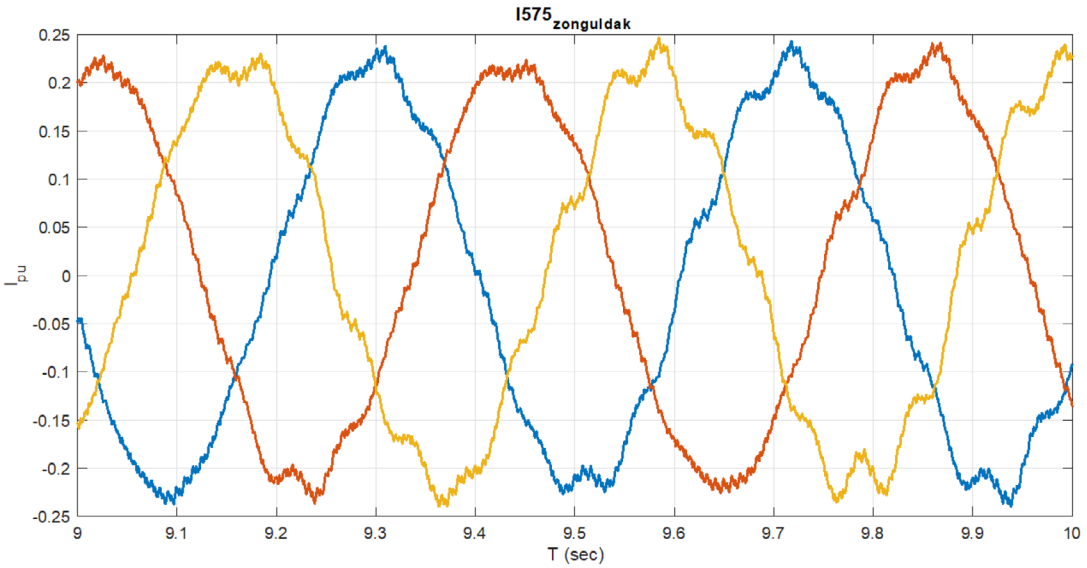

FIGURE 7 | Current waves at the 575 V bus for the Zonguldak.

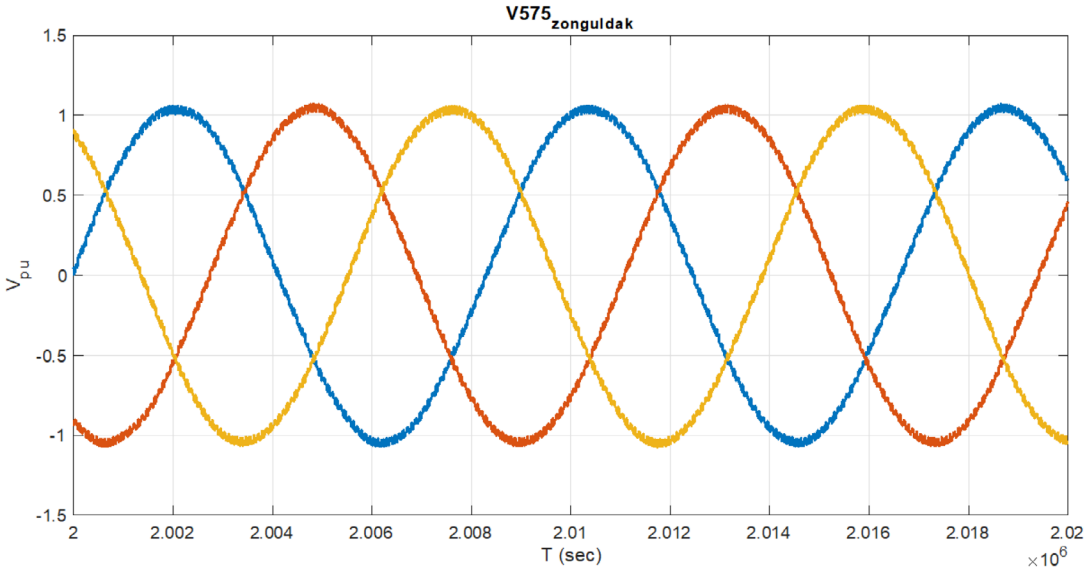

FIGURE 8 | Voltage waves at the 575 V bus for the Zonguldak.

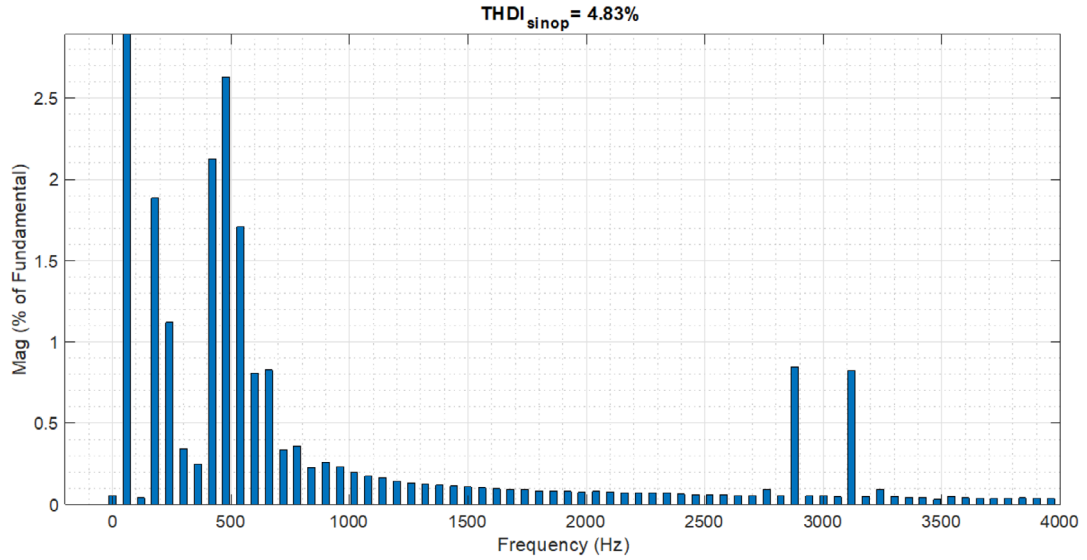

FIGURE 9 | Total harmonic distortion for current at the 575 V bus for the Sinop.

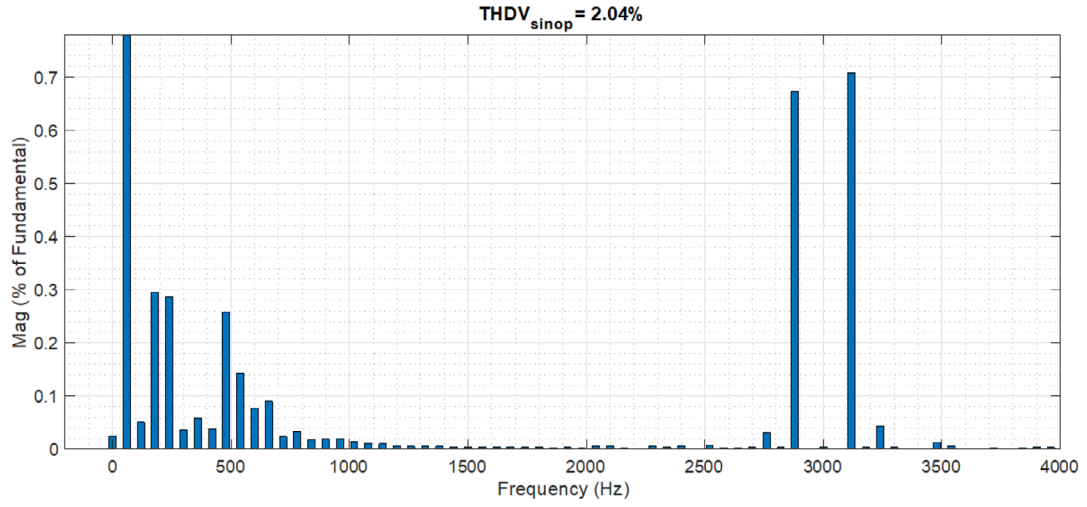

FIGURE 10 | Total harmonic distortion for voltage at the 575 V bus for the Sinop.

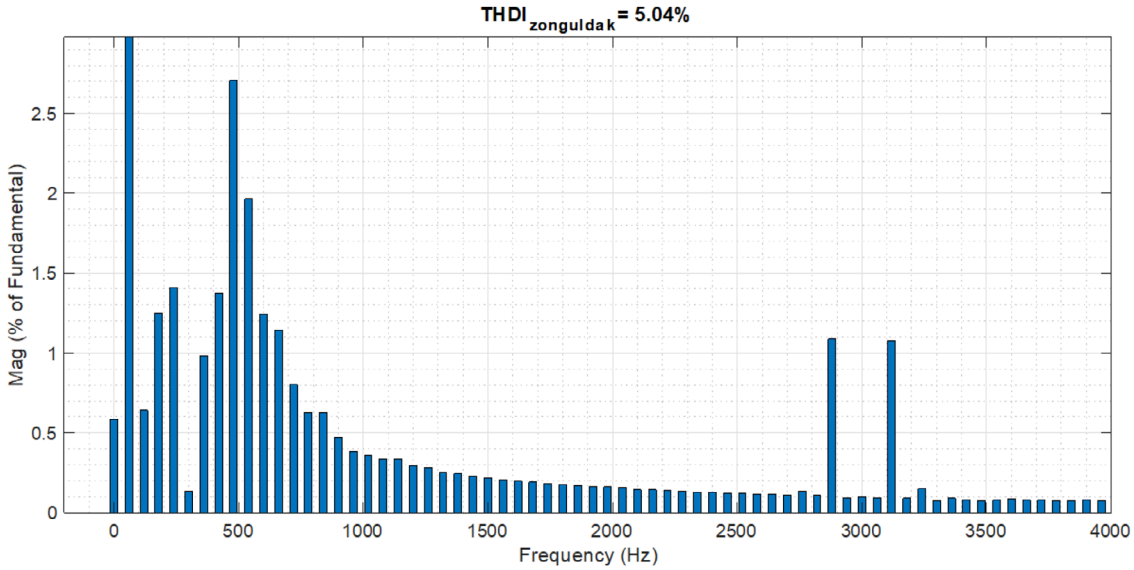

FIGURE 11 | Total harmonic distortion for current at the 575 V bus for the Zonguldak.

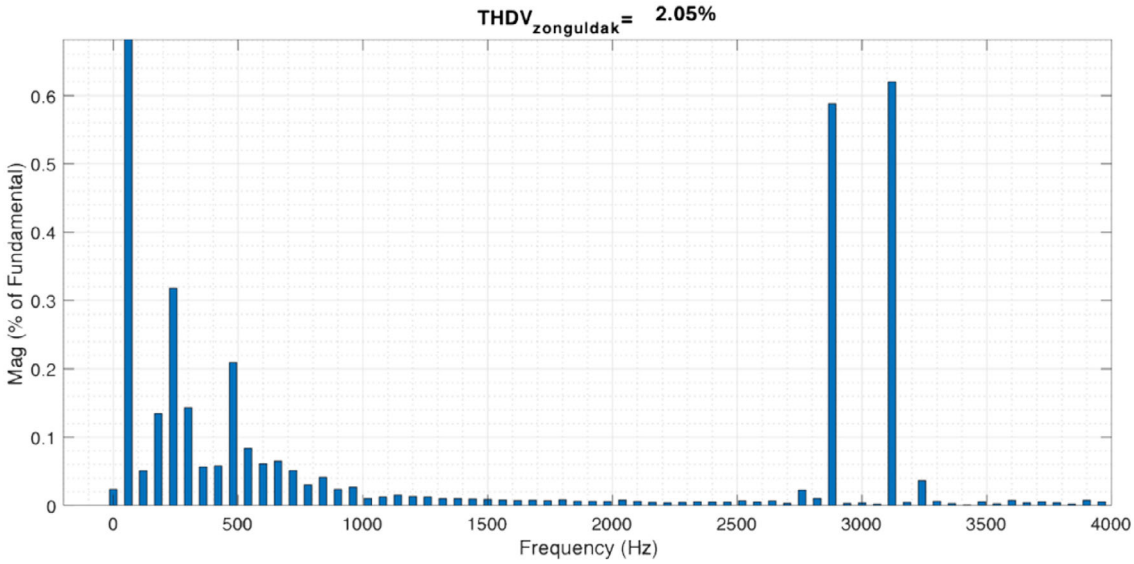

FIGURE 12 | Total harmonic distortion for voltage at the 575 V bus for the Zonguldak.

5 | Machine Learning Algorithms

As a subset of artificial intelligence (AI), machine learning specializes in developing models and algorithms that allow systems to learn from data and enhance their performance over time without direct programming. Its main objective is to build models capable of making predictions or decisions by identifying patterns within data. Machine learning for forecasting involves using various algorithms and techniques from the field of computer science to predict future values of a time series or sequence data. In this study, several ML algorithms are employed.

5.1 | Linear Regression

Linear regression is a fundamental quantitative algorithm used to model the relationship between a dependent variable (target) and several independent variables (features) [34]. The primary premise underlying this method is that there exists a linear interaction between predictor variables and the outcome variable, which is generally represented in Equation (1).

$$y = \beta_0 + \beta_1 x + \epsilon \tag{1}$$

The variable y denotes the target or dependent element, while x represents the predictor or independent variable. The parameter β_0 corresponds to the intercept, indicating where the regression line crosses the y-axis, and β_1 is the slope coefficient that defines the line's inclination. ϵ captures the difference between the observed and predicted values, serving as an indicator of the model's accuracy. These components collectively create the basis for linear regression, allowing for the modeling and understanding of the connections.

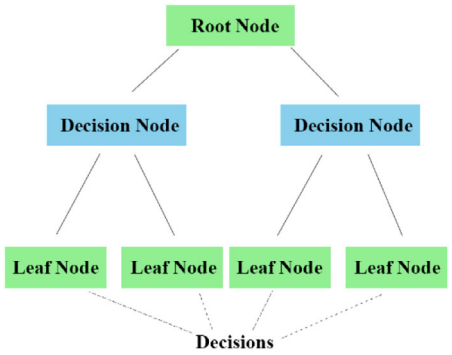

FIGURE 13 | Decision Tree model structure [36].

5.2 | Decision Tree

In Figure 13, a decision tree can be understood as a hierarchical structure that partitions data by evaluating feature values through a series of binary decisions. Each node [35] in the tree signifies a decision point, while the branches represent the possible outcomes leading to either further questions or final classification results. Building the tree requires choosing the most suitable features and split thresholds to divide the dataset, aiming to reduce impurity or enhance information gain at every split. Multiple criteria [36] are applied to identify the optimal partition points, including:

- **Gini Impurity**: This measure estimates the likelihood that a randomly selected sample would be incorrectly classified if its label were assigned based on the class distribution within the node. The Gini impurity for a node is computed using the class probabilities p_i for each i class, as expressed in Equation (2).

$$\text{Gini}(S) = 1 - \sum_{i=1}^{c} p_i^2 \tag{2}$$

- **Entropy**: Entropy measures the degree of randomness or uncertainty in the distribution of classes within a dataset. In Equation (3), the $H(S)$ entropy for a S set is defined.

$$H(S) = -\sum_{i=1}^{c} p_i \log_2 p_i \tag{3}$$

- **Information Gain**: It represents the decrease in entropy or Gini impurity resulting from partitioning the data based on a specific feature. Gain is computed as the gap between the $H(D)$ parent node entropy and the mean entropy of the children with assigned weights after the split, as expressed in Equation (4).

$$IG(D,F) = H(D) - \sum_{v \in \text{Values}(F)} \frac{|D_v|}{|D|} H(D_v) \tag{4}$$

When estimating the target value of an unseen input, the model follows a path starting at the root of the decision tree, moving through internal nodes by evaluating the data point's feature values. The process continues until a terminal leaf node is reached, which provides the final prediction for that input.

5.3 | Random Forest

Random Forest creates a large number of decision trees during the learning phase and combines their predictions. The predictions are averaged for regression problems while the most frequent prediction is chosen for classification problems [37]. The term "random" comes from two main aspects. First, the training data is randomly selected with replacement for each tree, known as bootstrap sampling. Second, at each split in the tree, only a randomly chosen subset of features is considered. The process begins by generating diverse training subsets. Then, each tree is trained on these subsets while evaluating only a limited set of features at each node. Finally, the predictions from all trees are combined to produce the final output.

From a mathematical perspective, predictions for new or unseen samples $\mathbf{x}'$ are obtained through averaging the outputs of each tree as shown in Equation (5).

$$\hat{f}(\mathbf{x}') = \frac{1}{B} \sum_{b=1}^{B} f_b(\mathbf{x}') \tag{5}$$

To estimate the uncertainty in predictions, the std of the predictions from every tree for an input $\mathbf{x}'$ is computed as in Equation (6).

$$\hat{\sigma}(\mathbf{x}') = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(f_b(\mathbf{x}') - \hat{f}(\mathbf{x}')\right)^2} \tag{6}$$

B (number of trees) in the model is a tunable hyperparameter, which depends on the complexity and size of the dataset. The optimal value of B can be assessed through error metrics or cross-validation techniques.

5.4 | Gradient Boosting

This algorithm is a powerful ensemble method commonly used for regression problems [38]. It sequentially builds weak learners, combining them to form a strong predictive model. At each iteration, a new weak learner is trained to minimize the residual errors made by the current ensemble by fitting to the negative gradient calculated from the loss function.

Let the training set be represented as (x_i, y_i). x_i indicates the inputs and y_i the associated outputs. The aim is to identify a $F(x)$ that reduces a chosen L loss function. At the m-th iteration, a $h_m(x)$ weak learner is tuned to the negative gradient of the loss function computed at the current model estimate $F_{m-1}(x)$, which is given by:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \tag{7}$$

This $h_m(x)$ weak learner is then trained to approximate the r_{im} residuals, thereby aiming to reduce the loss. The model is updated as:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{8}$$

Shrinkage factor γ_m regulates the influence of the current weak learner on the ensemble. Lower values of γ_m generally improve generalization but necessitate more iterations M. The final model prediction after completing M iterations can be expressed as:

$$F_M(x) = F_0(x) + \sum_{m=1}^{M} \gamma_m h_m(x) \tag{9}$$

5.5 | XGBoost

XGBoost stands out as a robust and scalable machine learning method that works by sequentially combining decision trees and correcting errors made by previous models [39]. Through Gradient Boosting, Extreme Gradient Boosting minimizes a specified loss function to optimize model performance. XGBoost integrates regularization techniques to prevent overfitting and supports parallel computing for swift processing of large datasets. By pruning trees and controlling their size, it mitigates overfitting risks. Its versatility makes it popular across various tasks like regression, classification, and ranking.

5.6 | K-Nearest Neighbors (KNN)

K-nearest neighbors (KNN) is a straightforward yet robust machine learning algorithm for classification and regression [40]. It predicts the target value of a new data point by considering the k nearest neighbors in the training set based on a chosen distance metric like the Euclidean, Manhattan, and Minkowski distances given in Equations (10)–(12).

$$d_{\text{Euclidean}}(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{10}$$

$$d_{\text{Manhattan}}(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{11}$$

$$d_{\text{Minkowski}}(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}} \tag{12}$$

where x_i and y_i are the ith dimensions of the two points. p is a parameter that determines the type of distance. KNN memorizes the dataset without learning during training. The hyperparameter k controls the neighbor count, affecting model flexibility. While effective for smaller datasets, KNN is computationally intensive and sensitive to k and distance metric choices. Yet, with careful tuning, it remains versatile across tasks.

5.7 | AdaBoost

Adaptive Boosting (AdaBoost) is an ensemble learning method that uses weak classifiers trained iteratively on weighted data to focus on misclassified examples, with final predictions aggregated through weighted voting. The algorithm iteratively improves its performance by adjusting weights (Equation 13) and selecting classifiers (Equation 14) that perform well [41]. The final classifier is a weighted combination (Equation 15) of all weak classifiers.

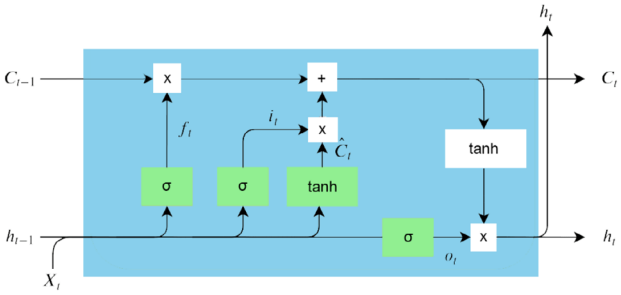

FIGURE 14 | LSTM neural network architecture [42].

$$\epsilon_t = \frac{\sum_{i=1}^{N} w_i^{(t)} \cdot I(y_i \neq h_t(x_i))}{\sum_{i=1}^{N} w_i^{(t)}} \tag{13}$$

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \tag{14}$$

$$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x) \right) \tag{15}$$

AdaBoost combines weak classifiers $h_t(x)$ to form a strong one by iteratively adjusting weights and classifier importance. Key formulas include the weighted error of a weak classifier (ϵ_t), its importance weight (α_t), and the updating and normalization of example weights. The final classifier $H(x)$ is a weighted combination of weak classifiers. The sign function returns +1 for positive values and −1 for negative values, effectively making a binary classification.

6 | Deep Learning Architectures

Deep learning architectures for time series data, such as LSTM networks, gated recurrent units (GRUs), and convolutional neural networks (CNNs), excel at capturing temporal dependencies within sequential data by processing inputs sequentially and maintaining evolving hidden states. These architectures are widely used in applications like forecasting, anomaly detection, and pattern recognition across various domains due to their ability to automatically learn complex temporal patterns from data. Several DL architectures are used in this research.

6.1 | Long Short-Term Memory (LSTM)

LSTM is an advanced variant of RNNs, developed to overcome the vanishing gradient problem encountered in training deep sequential models. This model can effectively capture long-range relationships in sequential inputs. LSTMs are commonly employed in a broad range of tasks, including natural language processing, speech recognition, and time-series forecasting [42]. The schematic of an LSTM cell is illustrated in Figure 14.

The fundamental idea of LSTMs lies in preserving a cell state that flows across time steps, allowing the network to control which information to keep and which to forget selectively. This regulation is performed by three main gates (forget, input, and

output) implemented as small neural networks that control the flow of information within the cell.

The LSTM structure comprises three main gates that regulate the information transfer across time steps. The forget gate decides which portion of the previous cell state C_{t-1} should be discarded or retained. It takes the current input x_t and the previous hidden state h_{t-1} as inputs, producing a forget gate vector f_t with elements ranging between 0 and 1. This vector performs an component wise multiplication with the prior cell state, effectively filtering out the information that is no longer needed. The input gate manages how new data is added to the cell state by generating an i_t input gate vector and a $\tilde{C}_t$ candidate cell state, also based on x_t and h_{t-1}. The candidate state is modulated by the input gate vector before being incorporated into the current cell state. Finally, the output gate determines the hidden state h_t of the LSTM cell. It applies a tanh activation function to the updated cell state C_t and multiplies the result by the output gate vector o_t, producing the output hidden state for the current time step.

These operations are mathematically described as follows:

$$\begin{aligned} f_t &= \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \\ i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh\left(W_c[h_{t-1}, x_t] + b_c\right) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t &= o_t \odot \tanh(C_t) \end{aligned} \tag{16}$$

Here, W_f, W_i, W_c, W_o are weight matrices, b_f, b_i, b_c, b_o are bias vectors, σ denotes the sigmoid activation function, and $\odot$ represents element-wise multiplication. The vector $[h_{t-1}, x_t]$ indicates the joining of the earlier hidden state and the recent data.

By leveraging these gating mechanisms, LSTMs effectively learn long-range temporal dependencies and mitigate the vanishing gradient problem, making them highly effective for sequential data modeling.

6.2 | Gated Recurrent Units (GRU)

GRUs operate primarily through hidden states that act as memory cells, storing information from earlier steps in a sequence. A key feature setting GRUs apart from other neural networks is their gating structures, which control how information is passed through the network. These mechanisms consist of two main gates: the update gate (Equation 17) and the reset gate (Equation 18).

The update gate z_t controls how much of the previous hidden state needs to be updated or preserved based on the current input. This update gate helps the model decide which past information is relevant for the current time step.

Conversely, the reset gate r_t controls how much of the previous hidden state is reset or discarded based on the current input:

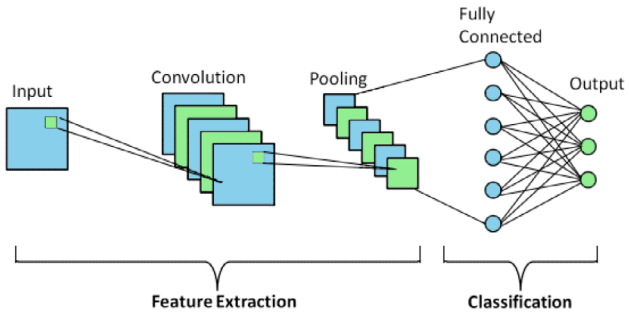

FIGURE 15 | Visual representation of CNN layers [44].

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \tag{17}$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \tag{18}$$

After computing the update and reset gates, the GRU generates a candidate hidden state $\tilde{h}_t$ by combining the reset-modified previous hidden state with the current input (Equation 19). The candidate state represents the model's modified view of the current sequence step within the context:

$$\tilde{h}_t = \tanh\left(W[r_t \odot h_{t-1}, x_t]\right) \tag{19}$$

Finally, the update gate z_t blends the candidate hidden state and the previous hidden state used to generate the updated state h_t for the current time step (Equation 20). This weighted combination allows the model to preserve essential information and eliminate obsolete or irrelevant data, enabling it to adapt efficiently to sequential inputs:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{20}$$

The GRU model is mathematically defined using trainable components like W weight matrices and σ activation functions. Through training, these values are optimized, enabling GRUs to retain information across long sequences and capture sequential patterns [43].

6.3 | Convolutional Neural Networks (CNN)

CNNs are a cornerstone in the domain of deep learning, particularly renowned for their effectiveness in image-related tasks [44]. Convolutional layers in Figure 15 use small filters sliding across input data, extracting features via convolutions. Each convolutional layer is typically followed by an activation function for ReLU, which introduces non-linearity into the network and enables it to model intricate relationships within the data. Equation (21) calculates the convolution by summing over all possible translations of the kernel function, multiplying the corresponding values of feature map and kernel, and then summing up these products.

$$(f * g)(i, j) = \sum_m \sum_n f(m, n) \cdot g(i - m, j - n) \tag{21}$$

f represents the input image or feature map, g represents the convolutional filter (also called kernel or weight), (i, j) indicates the position in the output feature map, (m, n) denotes the position

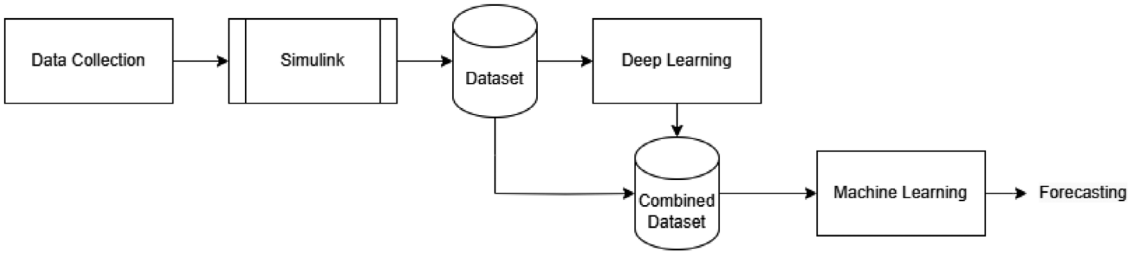

FIGURE 16 | Model stacking forecasting procedure diagram illustrating the integration of deep learning (DL) and machine learning (ML) models.

within the filter, and $(i-m, j-n)$ corresponds to the related position in the input image or feature map.

Pooling layers downsample feature maps, retaining key information. Max pooling selects the maximum value from each patch, compressing the feature maps. This controls complexity and prevents overfitting. Fully connected layers integrate high-level features for global predictions, connecting all neurons. CNNs optimize with backpropagation and gradient-based algorithms, minimizing loss. They excel at extracting features from raw data, which is ideal for images. Their modular architecture enables effective pattern abstraction.

7 | Deep and Machine Learning Hybrid Models

Hybrid models combine deep learning and traditional machine learning approaches to take advantage of the advantages of each approach when handling challenging problems. These models integrate features that are automatically extracted from raw data through deep learning techniques with handcrafted features developed through domain knowledge. This fusion results in a more comprehensive and detailed representation of the underlying information. To achieve this, feature fusion methods are employed to merge insights obtained from both deep and machine learning, leading to better understanding of the data and improved prediction accuracy.

The performance of these hybrid models is further enhanced by using model stacking techniques. In this process, the outputs generated by the machine learning models are used as input features for the deep learning models, allowing the system to benefit from the complementary strengths of both approaches. In addition, hybrid models increase transparency by combining the explainable features produced by machine learning algorithms with the powerful forecasting abilities of deep learning models. This synergy enables hybrid models to effectively solve a wide variety of problems while optimizing both predictive performance and model transparency.

The hybrid architecture based on model stacking for forecasting purposes is illustrated in Figure 16. The process starts with inputting meteorological data, which is then used to train deep learning models to extract relevant features from the raw input. The output features obtained from these deep learning models serve as inputs to the machine learning models. These ML models utilize the features to perform the final prediction of harmonic distortion.

The selection of machine learning and deep learning algorithms for hybrid models in this study is based on their effectiveness in time-series forecasting and pattern recognition, which are essential for accurate THDV prediction in offshore wind farms. LSTM networks were chosen due to their ability to capture long-term temporal dependencies in wind speed and harmonic distortion trends, making them highly suitable for sequential data modeling in power systems [45]. Similarly, GRU offer comparable sequence-learning capabilities but with reduced computational complexity, making them advantageous for large-scale forecasting applications [46]. In addition to deep learning approaches, Gradient Boosting (GB) and Random Forest (RF) were selected for their strong performance in modeling nonlinear relationships and handling structured meteorological data, as previously demonstrated in harmonic forecasting studies [47]. By integrating hybrid ML-DL models, the study aims to combine the advantages of both methodologies, yielding more accurate results in predicting harmonic distortions. This multi-model approach ensures a robust and data-driven solution for power quality assessment in offshore wind farms.

8 | Data Preprocessing

The target variables of this study, THDV and THDI, were calculated using voltage (V) and current (I) data obtained from the TMY dataset with the Simulink model. The simulation results were exported as a comma-separated values (CSV) file. Multiple preprocessing procedures were performed to ready the dataset for use with algorithms. First, the dataset was loaded from the CSV file, providing a foundation for further processing and analysis. Next, the columns WS10m, THDV, and THDI were selected, as these were considered important for training and prediction. The columns that were not relevant were eliminated to keep only the data crucial for the learning process.

Two additional columns were created to capture the THDV and THDI values from the 2 preceding days. This step provided important contextual data for time series forecasting. Rows containing missing values due to the shifting process were excluded to preserve data quality. Moreover, the target variable, whether THDV or THDI, was relocated to the last column in the dataset for easier handling during model training. The division of training and testing datasets is shown in Figure 17.

The arrays for data entry and produced outcome were organized for training the model. For THDV prediction, the input features comprised wind speed ('WS10m') along with THDV measurements from the prior 2 days, while the output comprised solely

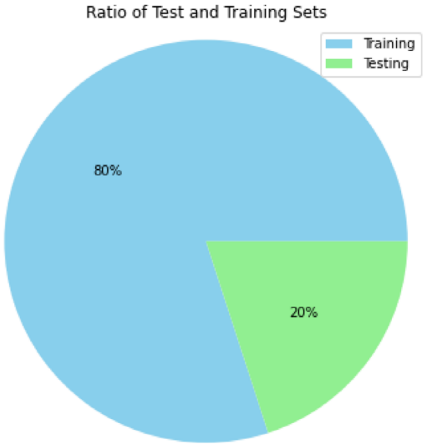

FIGURE 17 | Visualization of data allocation for training and testing.

of the response THDV variable. Similarly, for THDI, the input data contained wind speed ('WS10m') and THDI values from the previous 2 days, with the output comprising only the THDI target variable. The dataset was divided into training and test sets. Eighty percent of the data was used for training. The twenty percent was reserved for testing the model. Splitting the data into 80% for training and 20% for testing allows assessment of how good the model performs on unfamiliar data and assists in detecting potential problems.

9 | Model Implementation

Initially, various traditional regression algorithms were implemented. These models include Linear Regression, Decision Tree, Random Forest, Gradient Boosting, AdaBoost, K-Nearest Neighbors, and XGBoost. No changes or improvements were made to the vanilla machine learning models during this process, as they were already highly optimized. The implementation was carried out using the Scikit-learn library of the Python programming language [48].

Each model was trained using training data (Xtrain, ytrain), and predictions were made on the testing data (Xtest, ytest). Performance evaluation was done using RMSE and mean absolute error (MAE), which are common metrics for assessing model accuracy. The percentage values of RMSE and MAE indicated how well the model's predictions matched the test set. The results were used to compare the overall performance of the models.

An LSTM-based model was implemented using Keras, a widely used open-source deep learning library. With 100 and 50 units, the model was built using the Sequential API with two LSTM layers. After these layers, a Dense layer containing a neuron is incorporated. The initial layer was constructed to return the full sequence of outputs, allowing the second layer to process the complete temporal information. The input dimensions were adapted to match the structure of the training dataset. For optimization, the Adam algorithm was employed along with mean squared error (MSE) as the loss metric.

In parallel, a GRU network was developed, also utilizing the Sequential model structure. It featured two layers, each composed of 50 units. The network has a Dense output layer with one neuron. The input dimensions were established according to the time series characteristics of the dataset. Also, similar to the LSTM model, the Adam optimizer and MSE loss function is utilized.

TABLE 2 | Proposed hybrid forecasting models.

Deep learning		Machine learning
LSTM	+	Linear Regression
LSTM	+	Random Forest
LSTM	+	Gradient Boosting
LSTM	+	XGBoost
GRU	+	Linear Regression
GRU	+	Random Forest
GRU	+	Gradient Boosting
GRU	+	XGBoost

Lastly, a convolutional neural network (Conv1D) model was defined. In the first layer, a Conv1D layer with 64 filters and a kernel size of 3x1 was added. ReLU is used as the activation function [49]. The input shape settings of the model were adapted to match the dimensions of the training data. Then, the MaxPooling1D layer was added to reduce the size of the feature map. Subsequently, the Flatten layer was used to obtain a flattened feature map. After that, a hidden layer with 50 neurons and an output layer were added. The model was configured using the Adam optimizer and the mean squared error as the loss function [50].

LSTM, GRU, and CNN models were trained using the training data (Xtrain, ytrain) for 100 epochs with a 32 batch size, incorporating early stopping and validation data (Xtest, ytest). Following training, each model generated predictions on the test dataset (Xtest). The RMSE and MAE percentage values for LSTM, GRU, and CNN models were calculated. These values indicated the performance of each model's predictions.

The proposed hybrid models aim to predict harmonic values by integrating different algorithms. In this strategy, structured datasets are initially evaluated by means of LSTM and GRU architectures. The findings are subsequently utilized as inputs for traditional machine learning algorithms. This integration is intended to combine the advantages of both deep learning and classical approaches, thereby improving prediction performance. The forecasting framework incorporates combinations of LSTM and GRU with models such as Linear Regression, Random Forest, Gradient Boosting, and XGBoost, as outlined in Table 2. These specific algorithms were selected due to their reliable performance in earlier evaluations. The primary aim is to enhance forecast accuracy by merging deep learning-derived predictions with the original input features employed in the training of the machine learning models.

A pseudocode representation was developed to illustrate a hybrid regression method that integrates deep learning with traditional machine learning techniques. Initially, the DL and ML models

ALGORITHM 1 | Hybrid model for regression.

Input: Training data (X_{train}), Test data (X_{test}), Training target values (y_{train}), Test target values (y_{test})
Output: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE)
BEGIN
    Step 1: Initialize DL and ML models
    Step 2: Train the DL model using X_{train} and y_{train}
    Step 3: Predict using the DL model on X_{train}
    Step 4: Predict using the DL model on X_{test}
    Step 5: Combine initial data with DL predictions
        Concatenate X_{train} with the DL training predictions
        Concatenate X_{test} with the DL test predictions
    Step 6: Train the ML model using the combined training data
    Step 7: Predict using the ML regressor on the combined test data
    Step 8: Calculate RMSE and MAE percentage between y_{test} and the predictions
END

were set up. The DL model was trained using the training data and target values. Subsequently, predictions were made with the DL model on both the training and test datasets. These DL predictions were then combined with the original training and test features. The ML model was trained using this augmented training data, which included both the original features and the DL predictions for THDV and THDI. The trained ML model was then used to make predictions on the augmented test data. Finally, the performance of the hybrid model was evaluated by calculating the root mean squared error and mean absolute error percentages by comparing the hybrid model's predictions to the actual test target values.

10 | Experiments and Results

Initially, machine learning models were trained using data collected from the Zonguldak and Sinop regions. The evaluation of these models in predicting THDV and THDI for both areas delivers thorough analysis of their respective performance aspects. In Table 3, the regression results of the models are summarized.

Linear Regression, which demonstrates reasonable accuracy for THDV of the Zonguldak data (RMSE: 1.711%, MAE: 1.317%), faces challenges in accurately modeling the complexity of the Sinop data (RMSE: 1.838%, MAE: 1.416%). Despite its simplicity, Decision Tree Regressor shows larger error values for THDI, particularly indicating overfitting in the Sinop region (RMSE: 18.067%, MAE: 14.105%). Conversely, Random Forest, Gradient Boosting, and XGB regressors exhibit strong performance by effectively capturing complex patterns within the data in both regions. While AdaBoost Regressor is competent, it shows slightly higher error rates compared to the top-performing algorithms. Importantly, the KNeighbors Regressor displays the least favorable performance due to its limited ability to capture fundamental relationships in the dataset.

TABLE 3 | Results of regression analysis for machine learning algorithms. RMSE and MAE values are given in percentage.

	THDV		THDI	
Algorithms	**RMSE**	**MAE**	**RMSE**	**MAE**
Zonguldak				
Linear Regression	1.711	1.317	17.317	13.781
Decision Tree Regressor	2.355	1.797	19.880	15.239
Random Forest Regressor	1.751	1.335	14.510	11.172
Gradient Boosting Regressor	1.708	1.305	14.295	11.120
AdaBoost Regressor	1.815	1.425	16.238	13.048
KNeighbors Regressor	2.981	2.296	28.040	21.439
XGB Regressor	1.716	1.315	14.885	11.466
Sinop				
Linear Regression	1.838	1.416	15.279	11.831
Decision Tree Regressor	2.540	1.931	18.067	14.105
Random Forest Regressor	1.886	1.433	13.118	10.329
Gradient Boosting Regressor	1.821	1.393	13.133	10.346
AdaBoost Regressor	1.980	1.600	14.880	12.169
KNeighbors Regressor	3.196	2.523	25.153	20.087
XGB Regressor	1.829	1.402	13.234	10.368

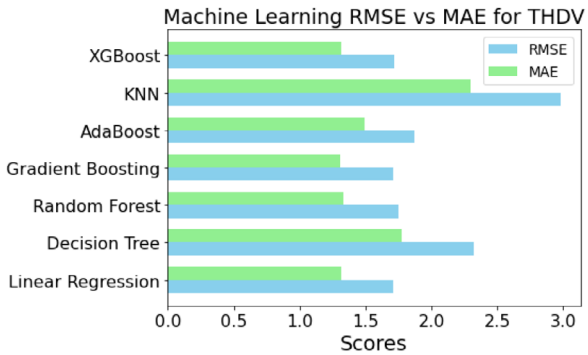

FIGURE 18 | Zonguldak dataset RMSE vs. MAE for THDV with ML algorithms.

The findings in Figures 18–21 underscore the importance of algorithm selection tailored to the characteristics of the dataset; Random Forest, Gradient Boosting, and XGB Regressors emerge as preferred options for accurate prediction of THDV and THDI in both regions. Figures 22–25 display the graphs of daily observed and forecasted values covering 1 month in July 2020 from the Zonguldak and Sinop datasets are presented. The results show that the predicted data align closely with the real measurements.

Subsequently, the regression performances of deep learning algorithms were tested. All deep learning models were set to run for 100 epochs. To prevent overfitting and optimize training efficiency, "EarlyStopping" was implemented during model training [51]. The stopping criterion was based on monitoring

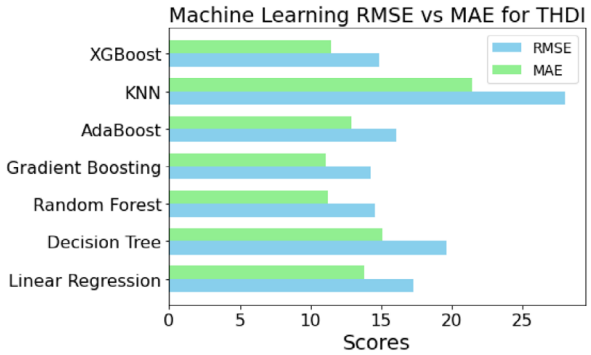

FIGURE 19 | Zonguldak dataset RMSE vs. MAE for THDI with ML algorithms.

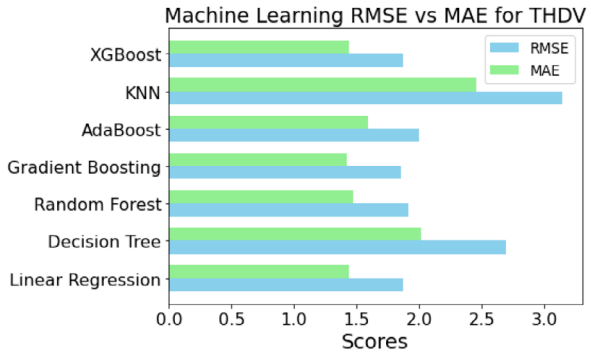

FIGURE 20 | Sinop dataset RMSE vs. MAE for THDV with ML algorithms.

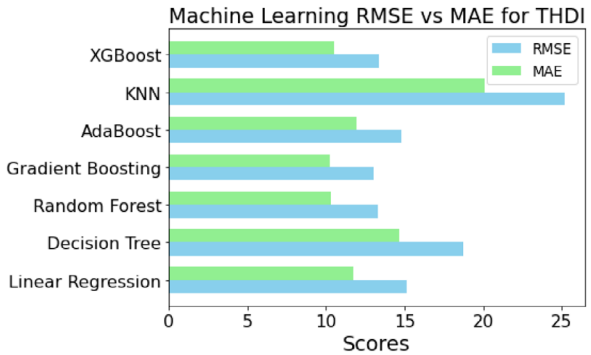

FIGURE 21 | Sinop dataset RMSE vs. MAE for THDI with ML algorithms.

the validation loss, and training was halted if no improvement was observed for 10 consecutive epochs (patience = 10). This threshold was determined through empirical tuning, where different patience values (e.g., 5, 10, 15 epochs) were tested, and 10 epochs provided the best balance between training stability and generalization performance. The early stopping mechanism helps avoid excessive training while ensuring that the models achieve optimal performance without overfitting to the training data.

Also, Table 4 shows the prediction results of deep learning models. In this table, lower values for both RMSE and MAE show higher performance. From Table 4, the LSTM and GRU models

TABLE 4 | Deep learning models regression analysis findings.

	THDV		THDI	
Models	**RMSE (%)**	**MAE (%)**	**RMSE (%)**	**MAE (%)**
Zonguldak				
LSTM	1.716	1.315	14.885	11.466
GRU	2.865	2.189	17.550	13.765
CNN	3.332	2.506	17.500	13.722
Sinop				
LSTM	1.829	1.402	13.234	10.368
GRU	3.084	2.418	16.260	12.836
CNN	3.642	2.728	16.059	12.624

have relatively low RMSE (1.716%, 2.865%) and MAE (1.315%, 2.189%) values, indicating that they provide better predictions compared to the CNN model (3.332% RMSE, 2.506% MAE) for THDV. The LSTM model has the lowest RMSE (1.716%) and MAE (1.315%) values among the three models, indicating its superior performance in predicting the data from the Zonguldak and Sinop locations. The GRU model also performs well, with slightly higher values compared to the LSTM model. The CNN model obtains significantly higher RMSE (17.5%) and MAE (13.722%) values compared to the other models, suggesting that it might not be the best choice for this dataset or may require further optimization. GRU and LSTM gave successful results, but CNN did not succeed in the regression task. One possible reason for this difference in performance could be that CNNs are more adept at handling image data than tabular data. Overall, from Figures 26–28 and 29, it can be said that, the LSTM model appears to be the optimal choice for predicting the harmonics, followed closely by the GRU model. To further clarify the theoretical basis behind the hybrid model's performance improvement, it should be emphasized that our approach leverages the strengths of both model types. Deep learning models (LSTM, GRU) are capable of learning temporal dependencies and nonlinear sequence dynamics, producing outputs that capture the underlying trends in harmonic distortion. When these outputs are used as additional features for machine learning models (e.g., Gradient Boosting), the ML component effectively performs residual learning, capturing relationships and correcting patterns not fully modeled by the DL component alone. This architecture is not a simple stacking but a theoretically grounded fusion that enhances overall prediction accuracy through complementary capabilities.

Finally, the prediction accuracies of the recommended hybrid models have been tested. Hybrid models are formed by combining LSTM or GRU with traditional regression algorithms (LR: Linear Regression, RF: Random Forest, GB: Gradient Boosting, XB: XGBoost). The hybrid models consistently achieve lower RMSE and mean absolute error (MAE) values compared to the standalone LSTM and GRU models, indicating good predictive performance. Among the hybrid models which is given in Table 5, GRU combined with Gradient Boosting (GRU+GB) and LSTM combined with Gradient Boosting (LSTM+GB) have the lowest RMSE (1.702%, 1.704% for THDV; 14.117%, 14.136% for THDI) and MAE (1.302%, 1.3% for THDV; 10.953%, 10.985% for

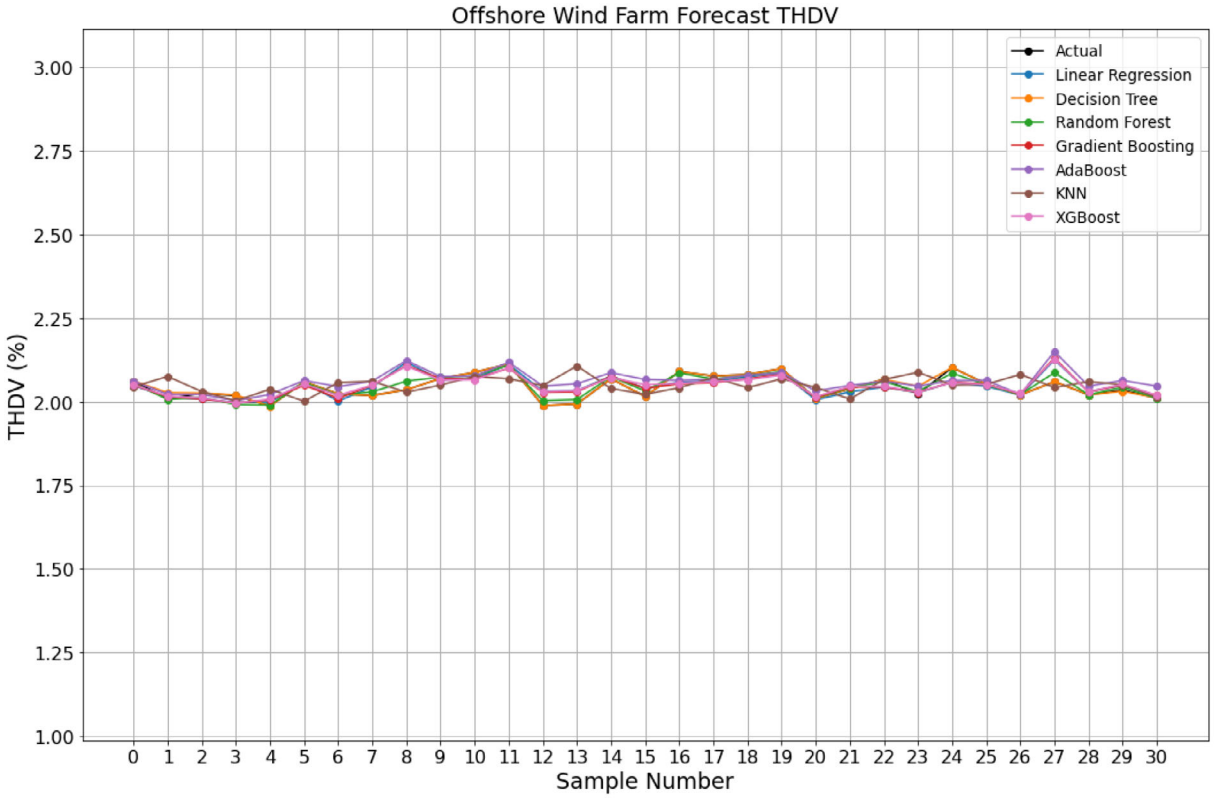

FIGURE 22 | Observed vs. forecasted THDV graph for ML algorithms (Zonguldak).

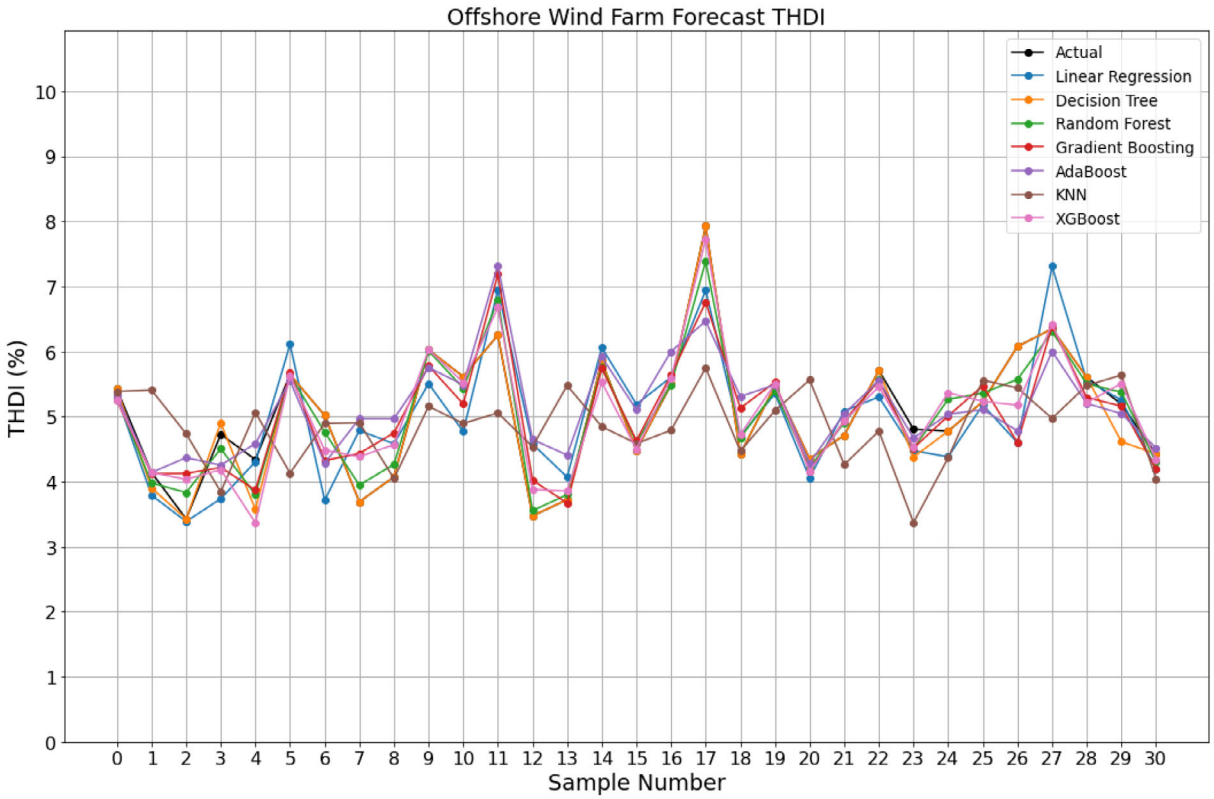

FIGURE 23 | Observed vs. forecasted THDI graph for ML algorithms (Zonguldak).

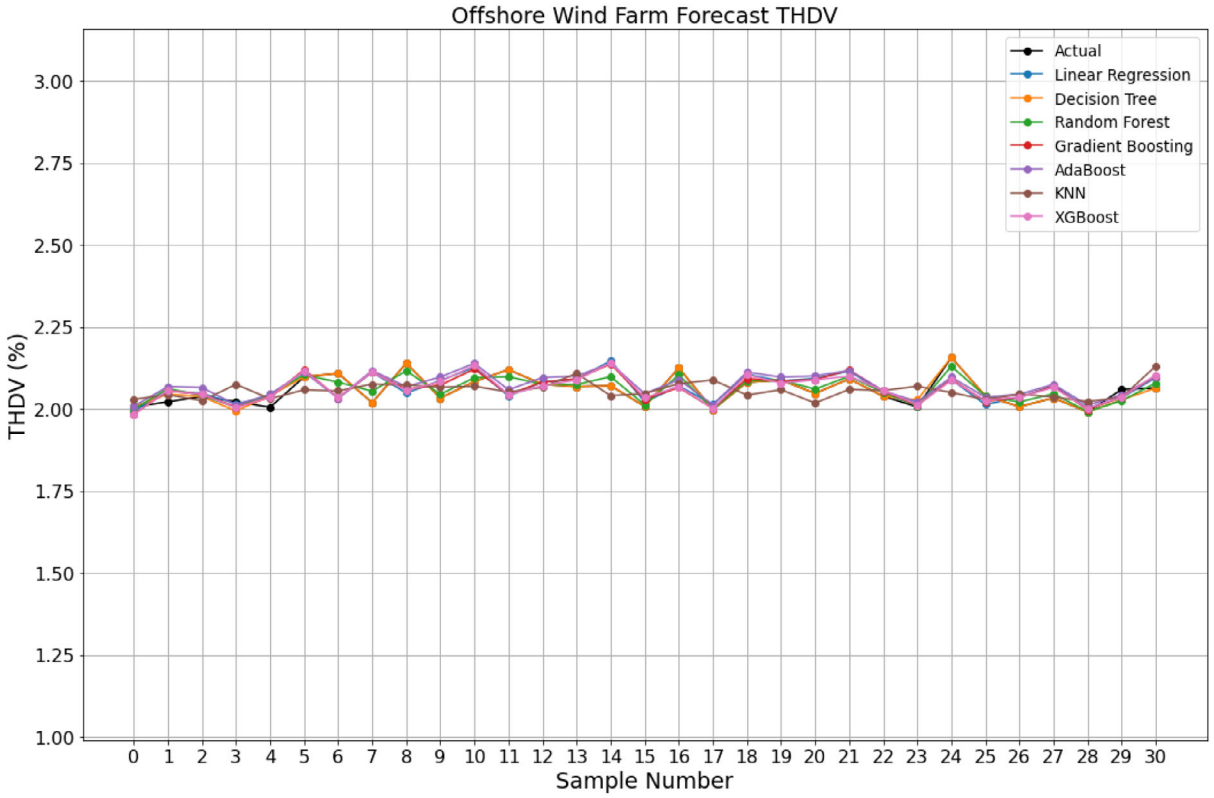

FIGURE 24 | Observed vs. forecasted THDV graph for ML algorithms (Sinop).

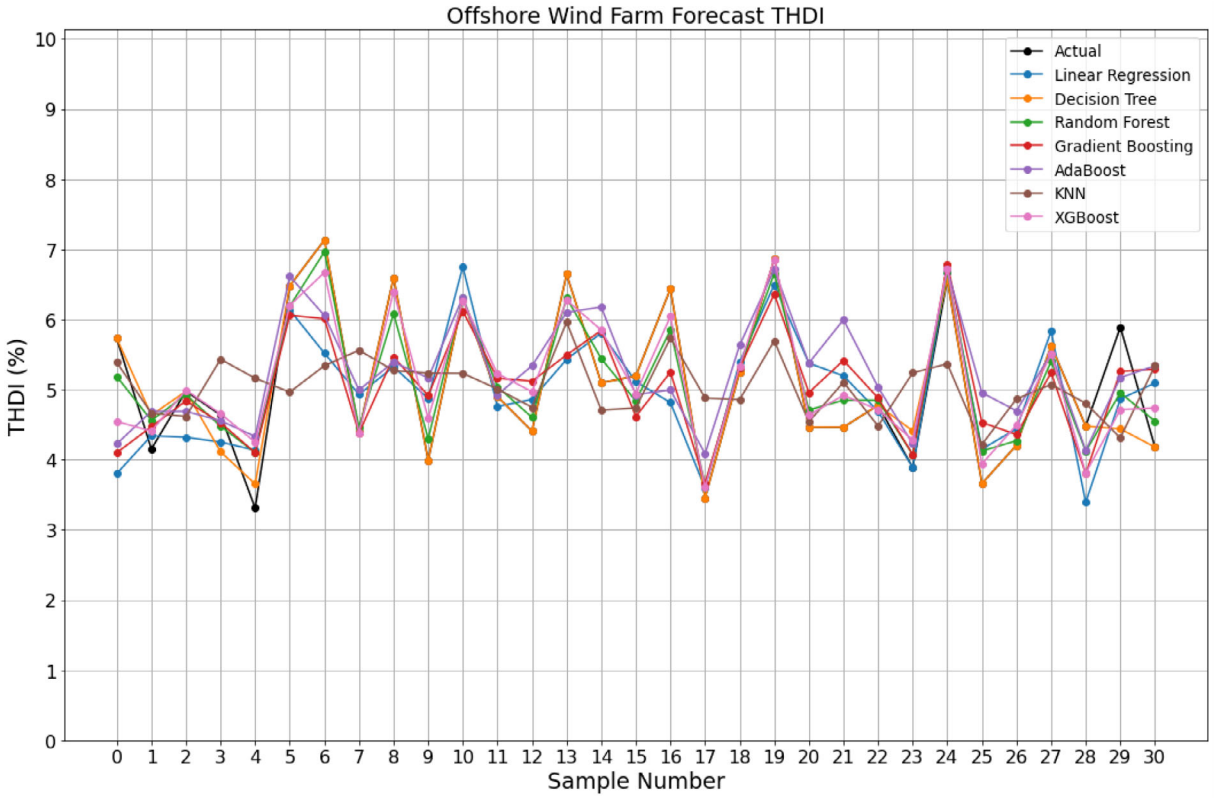

FIGURE 25 | Observed vs. forecasted THDI graph for ML algorithms (Sinop).

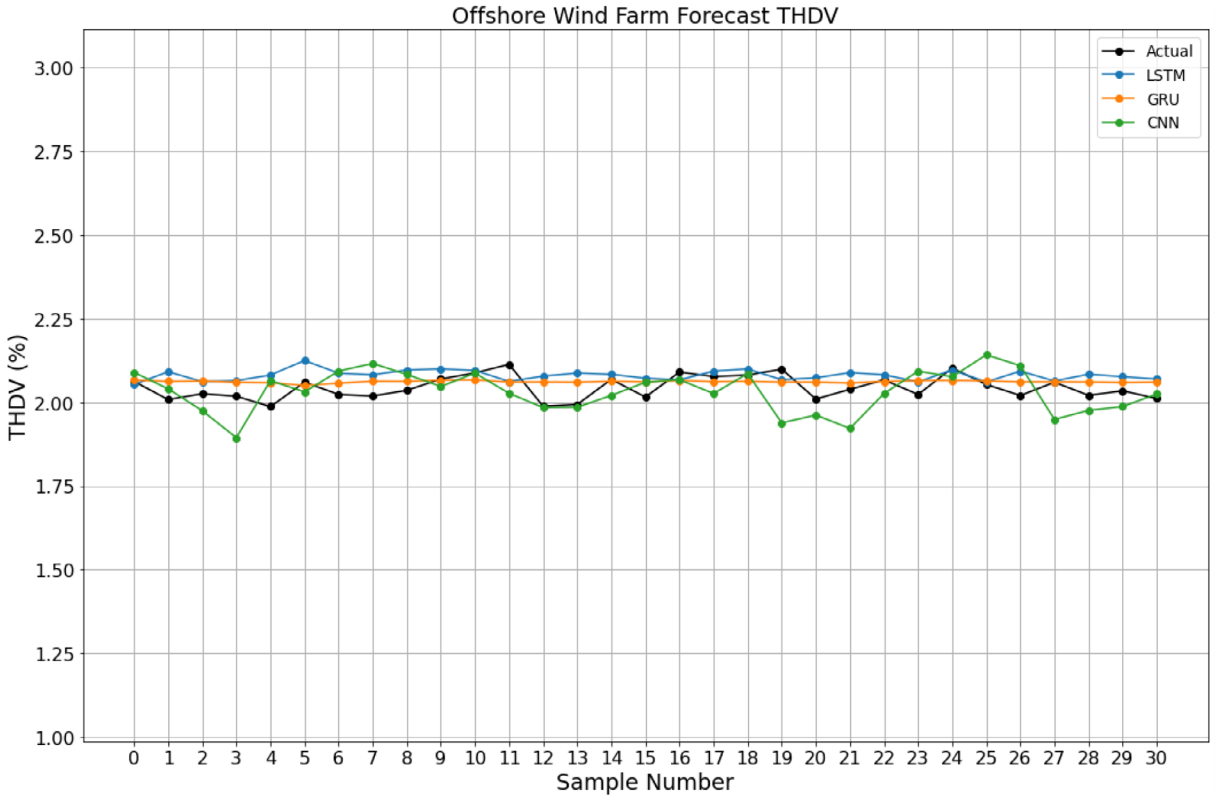

FIGURE 26 | Observed vs. forecasted THDV graph for DL architectures (Zonguldak).

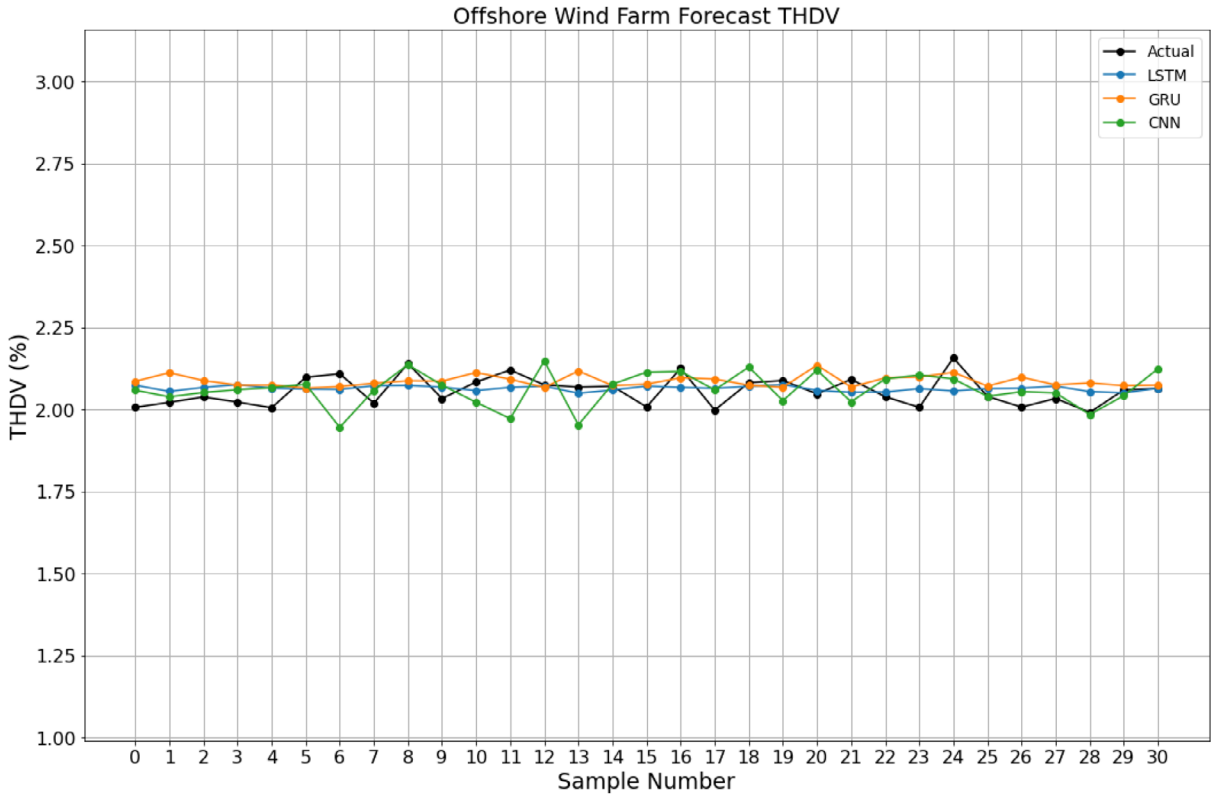

FIGURE 27 | Observed vs. forecasted THDV graph for DL architectures (Sinop).

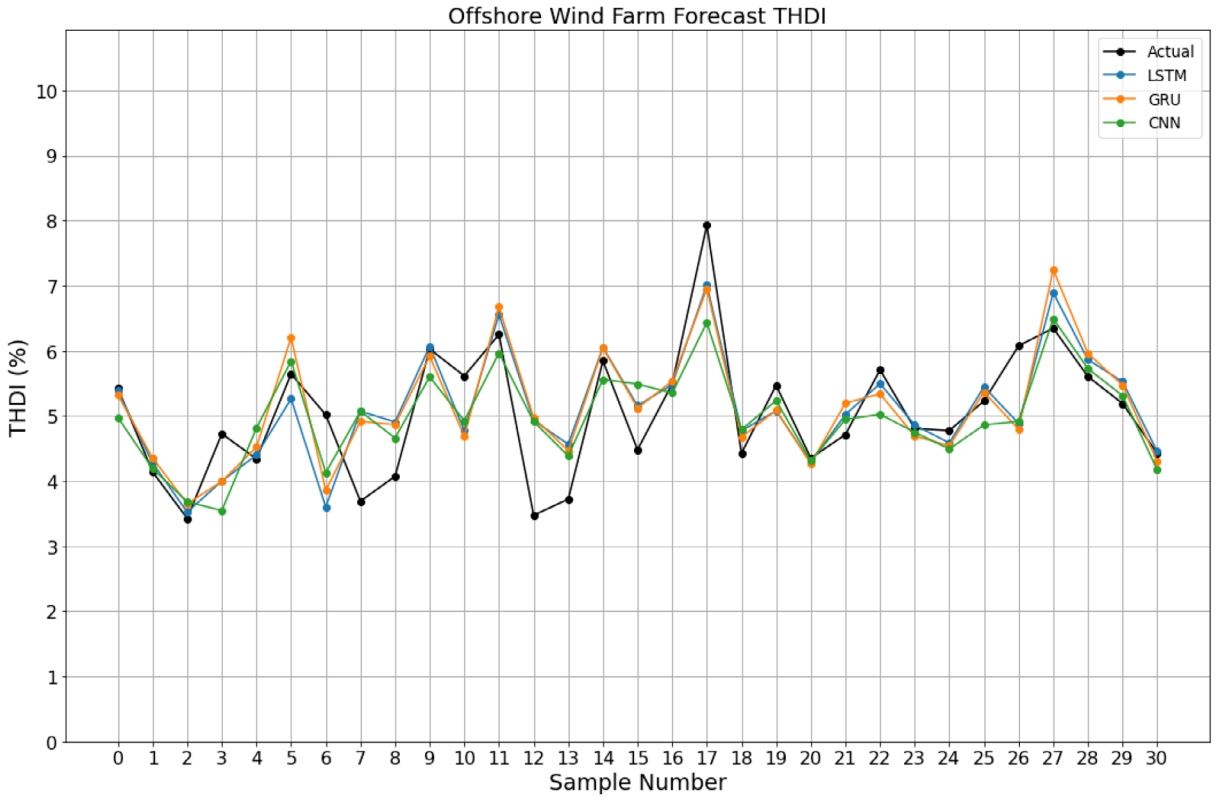

FIGURE 28 | Observed vs. forecasted THDI graph for DL architectures (Zonguldak).

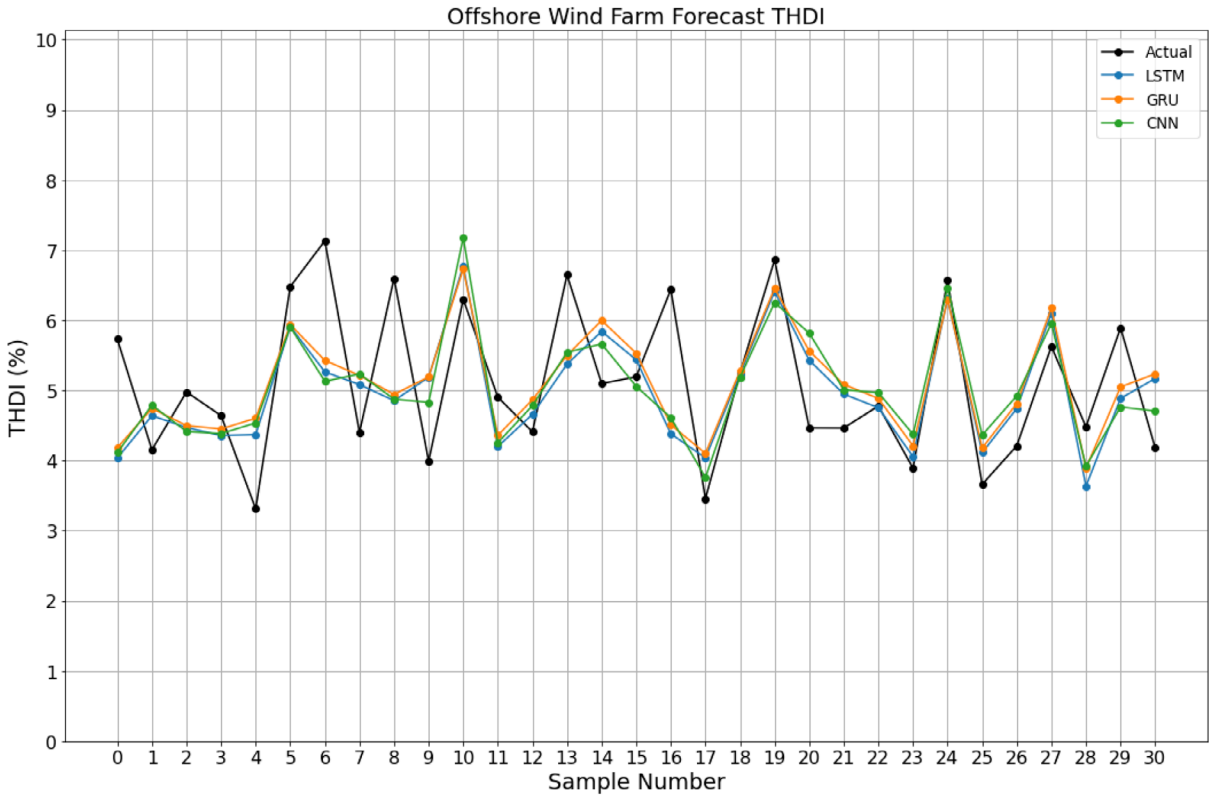

FIGURE 29 | Observed vs. forecasted THDI graph for DL architectures (Sinop).

TABLE 5 | Harmonic prediction outcomes from hybrid models (Zonguldak dataset).

	THDV		THDI	
Models	**RMSE (%)**	**MAE (%)**	**RMSE (%)**	**MAE (%)**
LSTM+LR	1.710	1.317	16.000	12.543
LSTM+RF	1.737	1.325	14.361	11.143
LSTM+GB	1.704	1.300	14.136	10.985
LSTM+XB	1.718	1.318	14.815	11.360
GRU+LR	1.710	1.317	15.868	12.484
GRU+RF	1.741	1.328	14.375	11.145
GRU+GB	1.702	1.302	14.117	10.953
GRU+XB	1.716	1.315	14.654	11.274

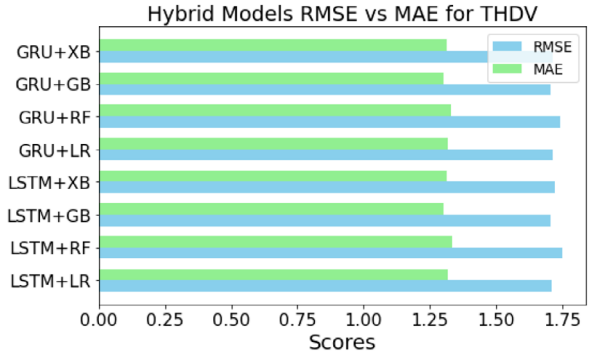

FIGURE 30 | Zonguldak dataset RMSE vs. MAE for THDV with hybrid models.

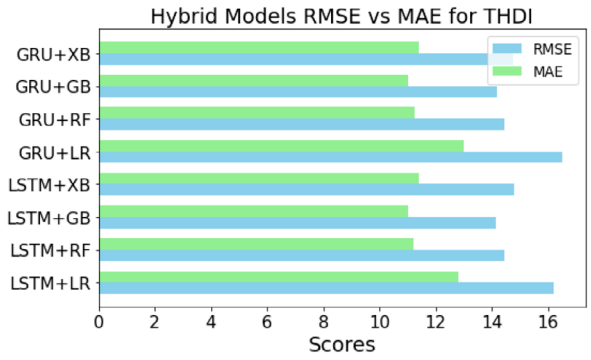

FIGURE 31 | Zonguldak dataset RMSE vs. MAE for THDI with hybrid models.

THDI) values, suggesting that these combinations provide the best predictive accuracy for the dataset. Results indicated that LSTM-based models generally exhibited marginally lower RMSE and MAE compared to their GRU counterparts (Figures 30 and 31), showcasing the superior sequential learning capabilities of LSTM networks. Notably, ensemble methods such as Gradient Boosting consistently demonstrated strong predictive accuracy (1.702% RMSE, 1.3% MSE) across both LSTM and GRU architectures. Among the secondary algorithms, Linear Regression emerged as a competitive choice for THDV, demonstrating comparable performance (1.710% RMSE, 1.317% MAE) to GB. Conversely, models incorporating Random Forest exhibited slightly higher THDI error rates, such as 14.361% RMSE and 11.143% MAE, indicating potential limitations in capturing complex nonlinear relationships. The hybrid models offer a promising approach by leveraging the strengths of both DL models (LSTM, GRU) and traditional regression algorithms, resulting in improved predictive accuracy for the dataset. Figure 32 demonstrates a close alignment between real and predicted THDV outputs in the hybrid models graph for Zonguldak. Also, Figure 33 reveals a close match between actual and predicted THDI values, affirming the models' effectiveness in harmonic forecasting.

TABLE 6 | Harmonic prediction outcomes from hybrid models (Sinop dataset).

	THDV		THDI	
Models	**RMSE (%)**	**MAE (%)**	**RMSE (%)**	**MAE (%)**
LSTM+LR	1.837	1.415	14.418	11.284
LSTM+RF	1.870	1.421	13.169	10.260
LSTM+GB	1.811	1.386	13.140	10.299
LSTM+XB	1.867	1.434	13.332	10.364
GRU+LR	1.835	1.413	14.252	11.196
GRU+RF	1.865	1.419	13.091	10.236
GRU+GB	1.813	1.384	13.121	10.288
GRU+XB	1.864	1.430	13.314	10.435

Table 6 provides the harmonic forecast results of hybrid models for the Sinop dataset. From this table, the forecast performance of each hybrid model is evaluated using RMSE and MAE metrics, represented as percentages. Among the hybrid methods, combinations such as GRU+RF and GRU+GB exhibit the minimum RMSE (13.091%) and MAE (1.384%) scores (Figures 34 and 35), suggesting superior predictive accuracy. LSTM+GB also shows competitive performance with 1.811% and 13.14% RMSE values. The hybrid models demonstrate low 13.091% RMSE and 10.236% MAE values for THDI, indicating effective forecasting performance for the Sinop dataset. They have yielded better results for THDV in the Zonguldak dataset. These results further emphasize the potential of hybrid models for achieving accurate harmonic forecasts by integrating deep learning with machine learning techniques. Figures 36 and 37 show a strong alignment between actual and predicted THDV and THDI values for Sinop, respectively, highlighting the robustness of the hybrid approach in ensuring consistent predictions.

The superior performance of the LSTM+GB model can be attributed to its ability to capture complex temporal dependencies and non-linear relationships inherent in harmonic distortion data. LSTM networks are adept at modeling long-term temporal dependencies, making them particularly effective in scenarios where harmonic variations are influenced by wind speed fluctuations and other time-dependent factors. Gradient Boosting (GB)complements this by efficiently handling nonlinear interactions within the data, allowing for a more nuanced understanding of how variables such as converter switching

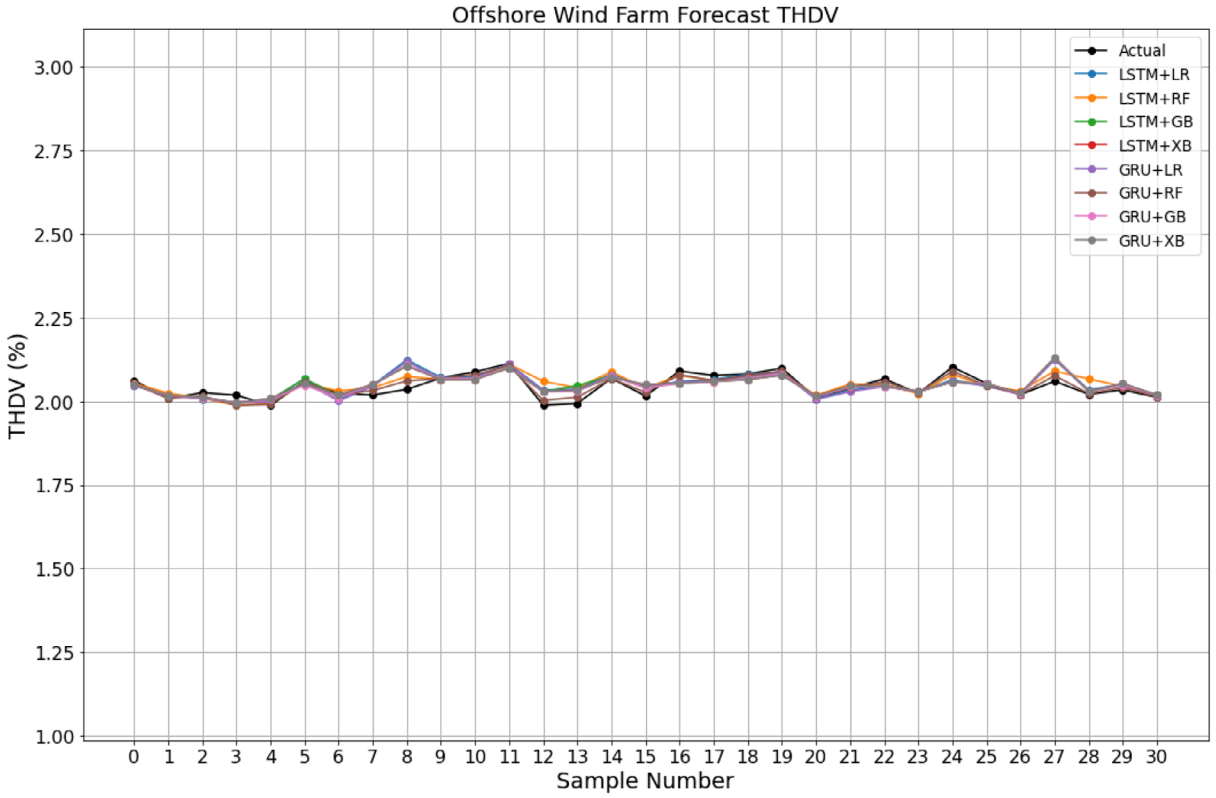

FIGURE 32 | Observed vs. forecasted THDV graph for hybrid models (Zonguldak).

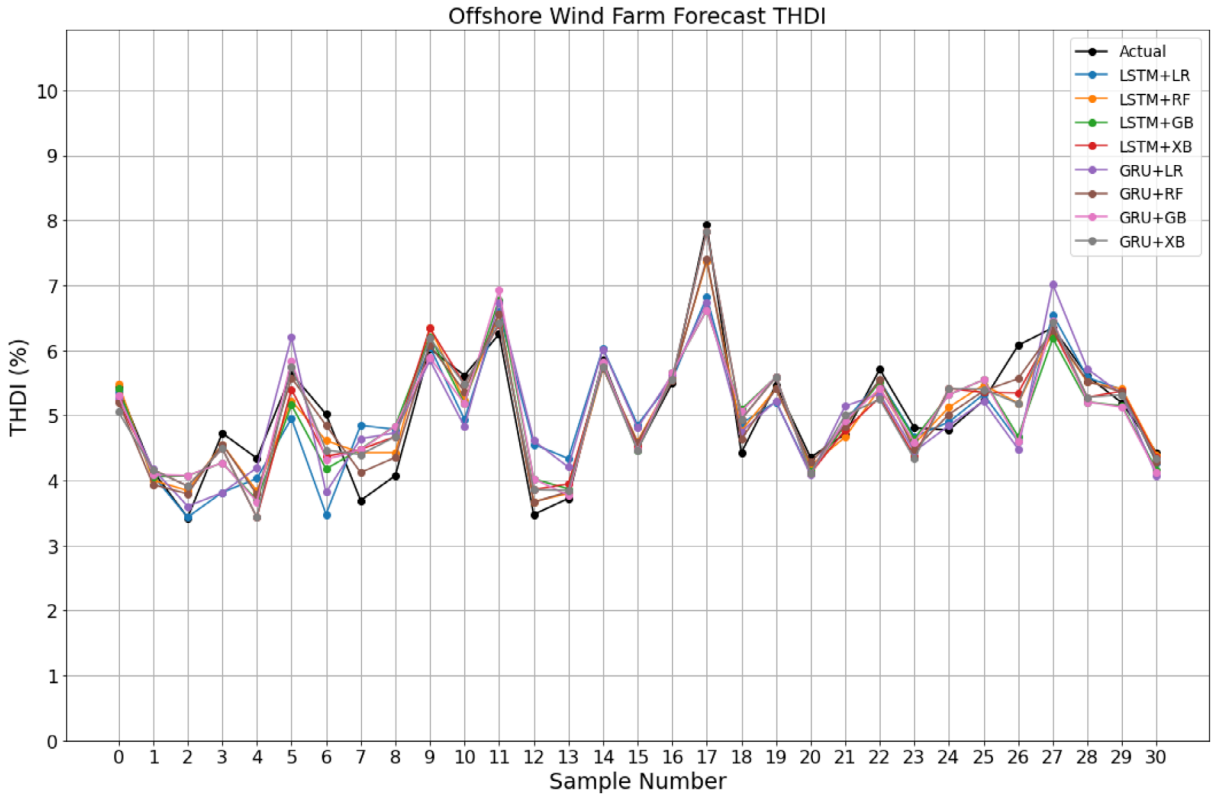

FIGURE 33 | Observed vs. forecasted THDI graph for hybrid models (Zonguldak).

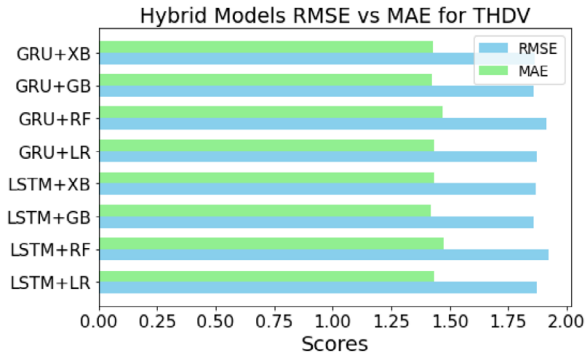

FIGURE 34 | Sinop dataset RMSE vs. MAE for THDV with hybrid models.

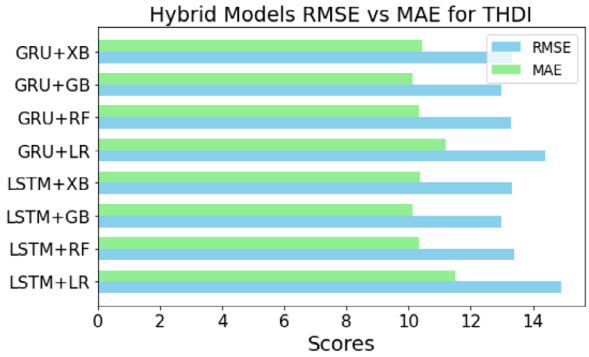

FIGURE 35 | Sinop dataset RMSE vs. MAE for THDI with hybrid models.

operations and submarine cable characteristics impact THD. Linear regression (LR) models, with their inherent assumption of linearity, may fail to capture these complex relationships, leading to less accurate predictions. The interaction between wind speed variability and converter switching can introduce high-frequency harmonics, which are better captured by models capable of learning non-linear patterns. Additionally, the distributed capacitance and inductance of submarine cables can cause harmonic resonance and amplification, phenomena that non-linear models like LSTM+GB are better equipped to predict. Therefore, the integration of LSTM and GB provides a robust framework for modeling the intricate dynamics affecting THD in offshore wind farms [52, 54].

10.1 | Limitations And Delimitations

The study has several limitations and delimitations that affect its applicability and scope. Limitations include the reliance on specific meteorological data from 2005 to 2020, which may not be representative of other regions or datasets, potentially affecting the generalizability of the findings. Additionally, the focus on PMSG systems means that other wind energy technologies are not considered, which could limit the applicability of the results to different turbine types. Moreover, the geographic focus on the Black Sea region, specifically Zonguldak and Sinop, may restrict the relevance of the findings to different environmental and meteorological conditions.

Also, delimitations of the study involve a deliberate confinement to offshore wind farms using PMSG technology, excluding other types of wind turbines and power production methods. The choice of the Black Sea region was intentional to address region-specific issues, with an acknowledgment that the results may not be applicable elsewhere. Furthermore, the use of TMY data from March 2005 to July 2020 ensures analytical consistency but may not account for more recent weather changes.

11 | Conclusion

This study focuses on forecasting harmonic values for a Type 4 offshore wind farm located at two key points, Zonguldak and Sinop, in the Black Sea region of Turkey. It provides a comparison of different machine learning and deep learning methods, individually and combined, to find the best-performing strategy for harmonic forecasting in this setting.

A Matlab/SIMULINK model of an electric power system incorporating offshore wind farms equipped with PMSG turbines was developed. Real wind speed measurements from the EU Science Hub for the Zonguldak and Sinop regions were employed.

Experimental data were generated from the offshore wind farm model to obtain parameters required for training machine and deep learning models. Several traditional learning techniques were applied for forecasting. Also, deep learning techniques such as GRUs and LSTM were used. Furthermore, hybrid approaches integrating machine learning algorithms with predictions generated by deep learning architectures were introduced to enhance forecasting accuracy. The capabilities of machine learning, deep learning, and hybrid methods was assessed for the offshore wind farm application. The key findings from the analyses are outlined below:

- For THDV forecasting, LSTM with GB and GRU with GB consistently yield the lowest RMSE (1.702%) and MAE (1.302%) values.
- THDI predictions benefit most from LSTM with GB and GRU with GB, showing the lowest RMSE (14.117%) and MAE (10.953%) values.
- LR and RF hybrid models perform competitively across both THDV and THDI forecasts, with RMSE ranging from 1.710% to 1.837% and MAE from 1.302% to 1.415%.
- The GB ensemble method exhibits strong predictive accuracy across all hybrid architectures.
- On average, GRU tends to perform slightly (0.01% to 0.1%) better than LSTM for both THDV and THDI in terms of RMSE and MAE.
- Overall, hybrid models, especially those incorporating ensemble methods, prove effective in achieving accurate harmonic forecasts for both locations.

Future studies may explore different types of hybrid models and investigate methods for their efficient integration into renewable energy systems to achieve more consistent forecasting.

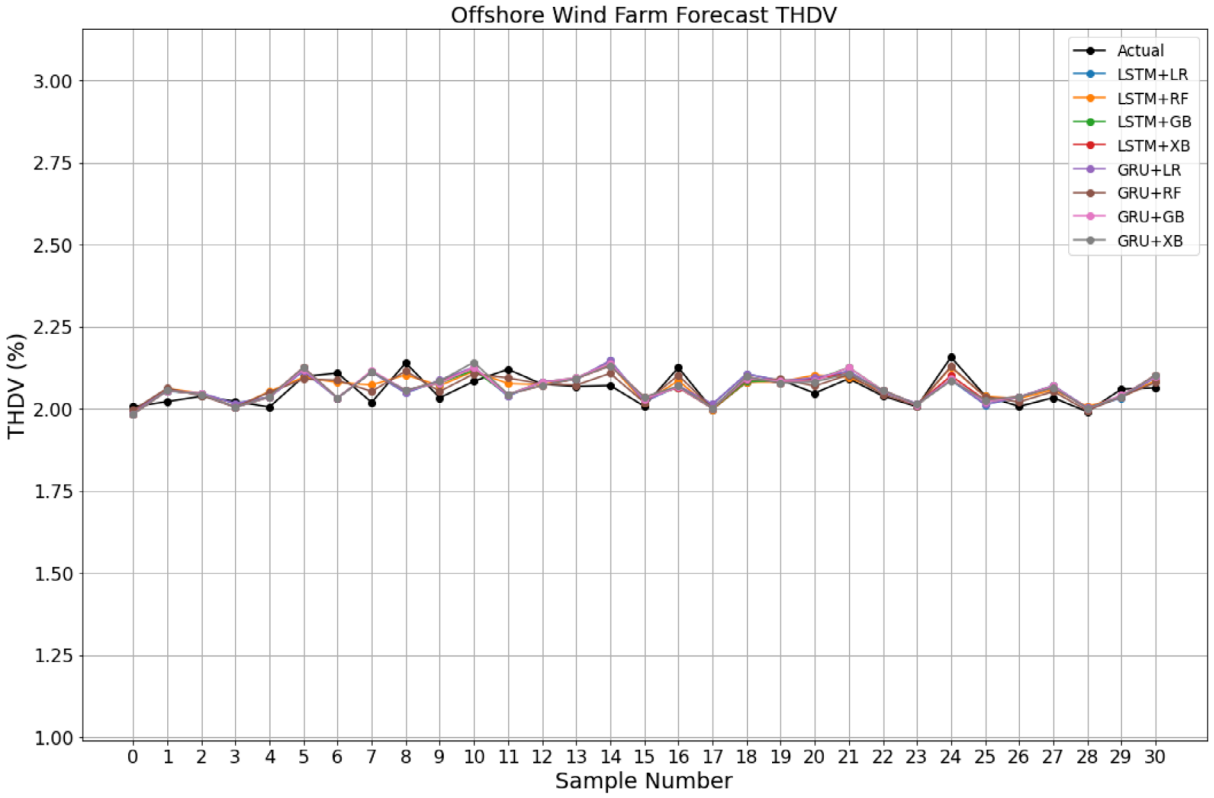

FIGURE 36 | Observed vs. forecasted THDV graph for hybrid models (Sinop).

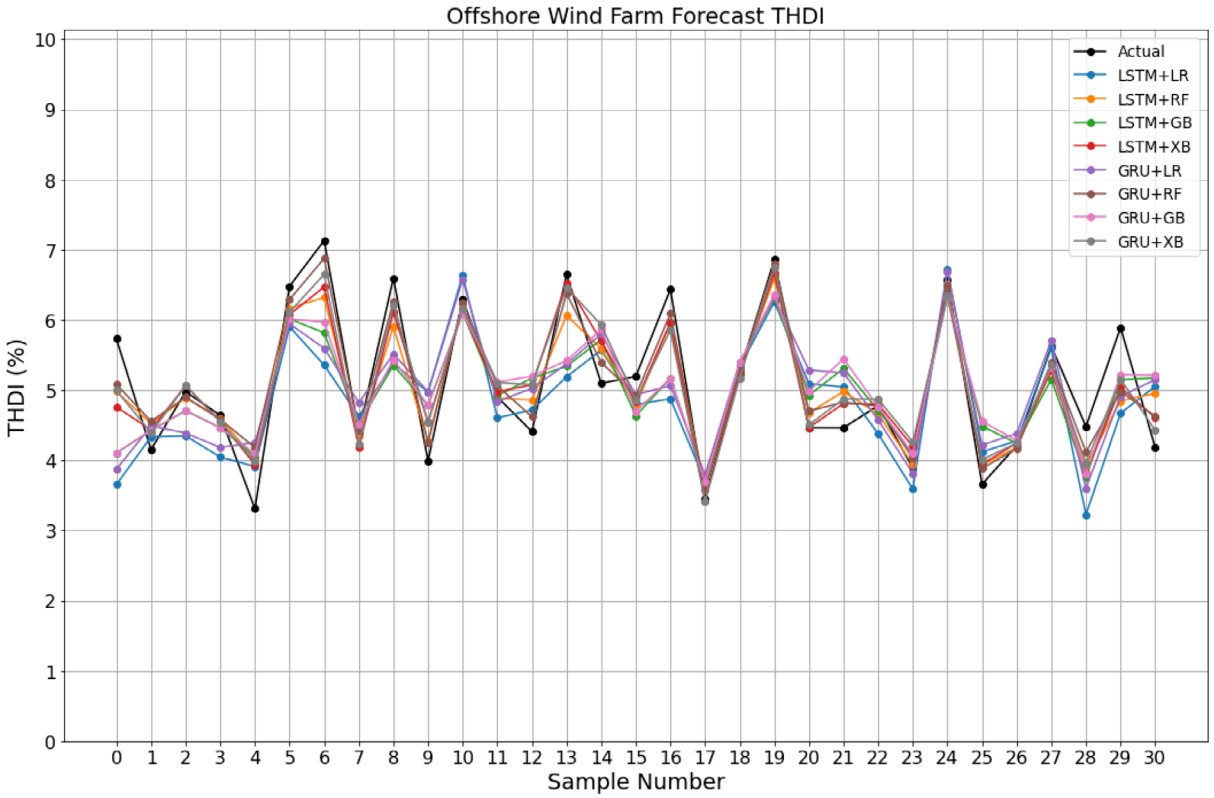

FIGURE 37 | Observed vs. forecasted THDI graph for hybrid models (Sinop).

Author Contributions

Alp Karadeniz: investigation, supervision, methodology, software, writing, validation.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

Data Availability Statement

Research data are not shared.

References

1. G. Van Kuik, B. Ummels, and R. Hendriks, "Perspectives on Wind Energy," in *Sustainable Energy Technologies: Options and Prospects* (Springer, 2008), 75–97.

2. C. Shan, *Harmonic Analysis of Collection Grid in Offshore Wind Installations* (NTNU, 2018).

3. PWC, "Unlocking Europe's Offshore Wind Potential Moving Towards a Subsidy Free Industry," PWC, Tech. Rep. (May 2017).

4. E. Ebrahimzadeh, F. Blaabjerg, X. Wang, and C. L. Bak, "Harmonic Stability and Resonance Analysis in Large PMSG-Based Wind Power Plants," *IEEE Transactions on Sustainable Energy* 9, no. 1 (2017): 12–23.

5. Ł. H. Kocewiak, B. L. Ø. Kramer, O. Holmstrøm, K. H. Jensen, and L. Shuai, "Resonance Damping in Array Cable Systems by Wind Turbine Active Filtering in Large Offshore Wind Power Plants," *IET Renewable Power Generation* 11, no. 7 (2017): 1069–1077.

6. H. Benbouhenni, N. Bizon, M. I. Mosaad, I. Colak, A. B. Djilali, and H. Gasmi, "Enhancement of the Power Quality of DFIG-Based Dual-Rotor Wind Turbine Systems Using Fractional Order Fuzzy Controller," *Expert Systems with Applications* 238 (2023): 121695, https://doi.org/10.1016/j.eswa.2023.121695.

7. M. Agazar, G. D'Avanzo, G. Frigo, D. Giordano, C. Iodice, P. S. Letizia, M. Luiso, A. Mariscotti, A. Mingotti, F. Munoz, et al., "Power Grids and Instrument Transformers up to 150 kHz: A Review of Literature and Standards," *Sensors* 24, no. 13 (2024): 4148, https://doi.org/10.3390/s24134148.

8. P. Kalkal and A. V. R. Teja, "A Novel Selective Harmonic Mitigation PWM Technique With THD Minimization Using Second-Order Sliding Modes," *IEEE Transactions on Industrial Electronics* 72, no. 1 (2024): 38–48, https://doi.org/10.1109/TIE.2024.3401205.

9. EU Science HUB, "PVGIS - Photovoltaic Geographical Information System," accessed April 16, 2024, https://re.jrc.ec.europa.eu/pvg_tools/en/#TMY.

10. A. Genç, F. K. Koca, and Ş. T. Azgın, "Offshore Wind Power Potential of the Black Sea Region in Turkey," *Energy Sources, Part B: Economics, Planning, and Policy* 12, no. 7 (2017): 604–613.

11. A. Durakovic, "Huge Offshore Wind Potential in Eastern Europe, Turkey Sits Untapped," published April 16, 2021, *Offshore Wind*.

12. Black Sea Renewable Energy Coalition (BSREC), *Effective Offshore Wind Grid and Infrastructure Development in the Black Sea* (BSREC, 2024).

13. Z.-H. Zhou, *Machine Learning* (Springer Nature, 2021).

14. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature* 521, no. 7553 (2015): 436–444.

15. Y. Yu, X. Si, C. Hu, and J. Zhang, "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures," *Neural Computation* 31, no. 7 (2019): 1235–1270.

16. R. Dey and F. M. Salem, "Gate-Variants of Gated Recurrent Unit (GRU) Neural Networks," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)* (IEEE, 2017), 1597–1600.

17. Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Transactions on Neural Networks and Learning Systems* 33, no. 12 (2021): 6999–7019.

18. P. Rodríguez-Pajarón, A. H. Bayo, and J. V. Milanović, "Forecasting Voltage Harmonic Distortion in Residential Distribution networks using smart meter data," *International Journal of Electrical Power & Energy Systems* 136 (2022): 107653.

19. F. M. A. Hadi, H. H. Aly, and T. Little, "Harmonics Forecasting of Wind and Solar Hybrid Model Based on Deep Machine Learning," *IEEE Access* 11 (2023): 100438–100457, https://doi.org/10.1109/ACCESS.2023.3314742.

20. F. M. A. Hadi, H. H. Aly, and T. Little, "Harmonics Forecasting of Wind and Solar Hybrid Model Driven by DFIG and PMSG Using ANN and ANFIS," *IEEE Access* 11 (2023): 55413–55424, https://doi.org/10.1109/ACCESS.2023.3253047.

21. F. M. A. Hadi and H. H. Aly, "Harmonics Forecasting of Renewable Energy System Using Hybrid Model Based on LSTM and ANFIS," *IEEE Access* 12 (2024): 50966–50985, https://doi.org/10.1109/ACCESS.2024.3386092.

22. Y. Li, Y. Sun, Q. Wang, K. Sun, K.-J. Li, and Y. Zhang, "Probabilistic Harmonic Forecasting of the Distribution System Considering Time-Varying Uncertainties of the Distributed Energy Resources and Electrical Loads," *Applied Energy* 329 (2023): 120298, https://doi.org/10.1016/j.apenergy.2022.120298.

23. E. M. Kuyunani, A. N. Hasan, and T. Shongwe, "Improving Voltage Harmonics Forecasting at a Wind Farm Using Deep Learning Techniques," in *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)* (IEEE, 2021), 1–6, https://doi.org/10.1109/ISIE45552.2021.9576357.

24. A. Taghvaie, T. Warnakulasuriya, D. Kumar, F. Zare, R. Sharma, and D. M. Vilathgamuwa, "A Comprehensive Review of Harmonic Issues and Estimation Techniques in Power System Networks Based on Traditional and Artificial Intelligence/Machine Learning," *IEEE Access* 11 (2023): 31417–31442, https://doi.org/10.1109/ACCESS.2023.3260768.

25. T. A. H. Alghamdi, O. T. E. Abdusalam, F. Anayi, and M. Packianather, "An Artificial Neural Network Based Harmonic Distortions Estimator for Grid-Connected Power Converter-Based Applications," *Ain Shams Engineering Journal* 14 (2023): 101916, https://doi.org/10.1016/j.asej.2022.101916.

26. Y. Chaoying, W. Ling, Z. Yuanzhe, Y. Donghai, and S. Yitao, "Harmonic Prediction Based on PCA-ELM-LSTM Hybrid Model," in *2024 6th Asia Energy and Electrical Engineering Symposium (AEEES)* (IEEE, 2024), 219–224, https://doi.org/10.1109/AEEES61147.2024.10544690.

27. D. Ganea, E. Mereuta, and L. Rusu, "Estimation of the Near Future Wind Power Potential in the Black Sea," *Energies* 11, no. 11 (2018): 3198.

28. M. Argin and V. Yerci, "Offshore Wind Power Potential of the Black Sea Region in Turkey," *International Journal of Green Energy* 14, no. 10 (2017): 811–818.

29. L.-I. Nedelcu and E. Rusu, "An analysis of the wind parameters in the western side of the Black Sea," *Inventions* 7, no. 1 (2022): 21.

30. F. Onea and L. Rusu, "Evaluation of Some State-of-the-Art Wind Technologies in the Nearshore of the Black Sea," *Energies* 11, no. 9 (2018): 2452.

31. I. Koletsis, V. Kotroni, K. Lagouvardos, and T. Soukissian, "Assessment of Offshore Wind Speed and Power Potential over the Mediterranean and the Black Seas under Future Climate Changes," *Renewable and Sustainable Energy Reviews* 60 (2016): 234–245.

32. A. Diaconita, G. Andrei, and L. Rusu, "New Insights into the Wind Energy Potential of the West Black Sea Area Based on the North Sea Wind Farms Model," *Energy Reports* 7 (2021): 112–118.

33. J. C. Das, *Power System Harmonics and Passive Filter Design* (Wiley, 2015).

34. S. Weisberg, *Applied Linear Regression* (John Wiley & Sons, 2005).

35. L. Rokach and O. Maimon, "Decision Trees," in *Data Mining and Knowledge Discovery Handbook* (Springer, 2005), 165–192.

36. S. Tangirala, "Evaluating the Impact of GINI Index and Information Gain on Classification Using Decision Tree Classifier Algorithm," *International Journal of Advanced Computer Science and Applications* 11, no. 2 (2020): 612–619.

37. A. Parmar, R. Katariya, and V. Patel, "A Review on Random Forest: An Ensemble Classifier," in *International Conference on (ICICI)* (Springer, 2018), 758–763.

38. A. Natekin and A. Knoll, "Gradient Boosting Machines, A Tutorial," *Frontiers in neurorobotics* 7 (2013): 21.

39. T. Chen, et al., "Xgboost: Extreme Gradient Boosting," *R package version 0.4-2* 1, no. 4 (2015): 1–4.

40. O. Kramer and O. Kramer, "K-nearest Neighbors," in *Dimensionality Reduction With Unsupervised Nearest Neighbors* (Springer, 2013), 13–23.

41. C. Ying, M. Qi-Guang, L. Jia-Chen, and G. Lin, "Advance and prospects of AdaBoost algorithm," *Acta Automatica Sinica* 39, no. 6 (2013): 745–758.

42. V. Hnamte, H. Nhung-Nguyen, J. Hussain, and Y. Hwa-Kim, "A Novel Two-Stage Deep Learning Model for Network Intrusion Detection: LSTM-AE," *IEEE Access* 11 (2023): 37131–37148, https://doi.org/10.1109/ACCESS.2023.3266979.

43. Z. Zhao, S. Yun, L. Jia, J. Guo, Y. Meng, N. He, X. Li, J. Shi, and L. Yang, "Hybrid VMD-CNN-GRU-Based Model for Short-Term Forecasting of Wind Power Considering Spatio-Temporal Features," *Engineering Applications of Artificial Intelligence* 121 (2023): 105982, https://doi.org/10.1016/j.engappai.2023.105982.

44. Y. Sun, Q. Zhou, L. Sun, L. Sun, J. Kang, and H. Li, "CNN–LSTM–AM: A Power Prediction Model for Offshore Wind turbines," *Ocean Engineering* 301 (2024): 117598, https://doi.org/10.1016/j.oceaneng.2024.117598.

45. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation* 9, no. 8 (1997): 1735–1780.

46. K. Cho, et al., "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," in *Proceedings of EMNLP* (ACL, 2014), 1724–1734.

47. J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *The Annals of Statistics* 29, no. 5 (2001): 1189–1232.

48. J. Hao and T. K. Ho, "Machine Learning Made Easy: A Review of scikit-Learn Package in Python Programming Language," *Journal of Educational and Behavioral Statistics* 44, no. 3 (2019): 348–361.

49. J. Schmidt-Hieber, "Nonparametric Regression Using Deep Neural Networks with ReLU Activation Function," *Annals of Statistics* 48, no. 4 (2020): 1875–1897.

50. Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)* (IEEE, 2018), 1–2.

51. L. Prechelt, "Early Stopping-But When?," in *Neural Networks: Tricks of the Trade* (Springer, 2002), 55–69.

52. Z. Shen, X. Fan, L. Zhang, et al., "Wind Speed Prediction of Unmanned Sailboat Based on CNN and LSTM Hybrid Neural Network," *Ocean Engineering* 254 (2022): 111352.

53. C. Hu, Y. Zhao, H. Jiang, et al., "Prediction of Ultra-Short-Term Wind Power Based on CEEMDAN-LSTM-TCN," *Energy Reports* 8 (2022): 483–492.

54. T. Westerweller, K. Friedrich, U. Armonies, et al., "Trans Bay Cable – World's First HVDC System Using Multilevel Voltage-Sourced Converter," in *CIGRÉ Session* (CIGRÉ, 2010).