energies

MDPI

Article

Predictions for Bending Strain at the Tower Bottom of Offshore Wind Turbine Based on the LSTM Model

Songjune Lee, Seungjin Kang and Gwang-Se Lee *

Wind Energy Research Team, Jeju Global Research Center (JGRC), Korea Institute of Energy Research (KIER), 200 Haemajihaean-ro, Gujwa-eup, Jeju 63357, Republic of Korea; sjlee83@kier.re.kr (S.L.); ksj@kier.re.kr (S.K.)
* Correspondence: lgs@kier.re.kr; Tel.: +82-64-800-2335

Abstract: In recent years, the demand and requirement for renewable energy have significantly increased due to concerns regarding energy security and the climate crisis. This has led to a significant focus on wind power generation. As the deployment of wind turbines continues to rise, there is a growing need to assess their lifespan and improve their stability. Access to accurate load data is crucial for enhancing safety and conducting remaining life assessments of wind turbines. However, maintaining and ensuring the reliability of measurement systems for long-term load data accumulation, stability assessments, and residual life evaluations can be challenging. As a result, numerous studies have been conducted on load prediction for wind turbines. However, existing load prediction models based on 10 min statistical data cannot adequately capture the short-term load variations experienced by wind turbines. Therefore, it is essential to develop models capable of predicting load with a high temporal resolution to enhance reliability, especially with the increasing scale and development of floating wind turbines. In this paper, we developed prediction models with a 50 Hz resolution for the bending strain at the tower bottom of offshore wind turbines by combining SCADA data and acceleration data using machine learning techniques and analyzed the results. The load prediction models demonstrated high accuracy, with a mean absolute percentage error below 4%.

Keywords: offshore wind turbine; bending strain prediction; structural health monitoring; machine learning; long short-term memory

check for updates

Citation: Lee, S.; Kang, S.; Lee, G.-S. Predictions for Bending Strain at the Tower Bottom of Offshore Wind Turbine Based on the LSTM Model. *Energies* **2023**, *16*, 4922. https://doi.org/10.3390/en16134922

Academic Editor: Davide Astolfi

Received: 22 May 2023
Revised: 19 June 2023
Accepted: 20 June 2023
Published: 24 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The demand for wind energy has significantly increased in recent years due to concerns regarding energy security and the global climate crisis. According to statistics from the International Renewable Energy Agency (IRENA), the global capacity for wind power generation has experienced rapid growth, increasing from around 180 GW in 2010 to approximately 824 GW in 2022 [1]. Wind power generation can be broadly classified into onshore and offshore wind power. While the majority of wind power installations have traditionally been located on land, the development of large-scale onshore wind farms has faced various challenges. These include limitations in finding suitable land areas for installation, concerns about environmental impact and ecosystem disruption, as well as noise complaints from nearby communities. To address these challenges, offshore wind power generation is garnering attention as a potential solution. Additionally, offshore wind power generation allows for stable power production due to its relatively consistent wind conditions.

Offshore wind turbines (OWTs) are exposed to not only wind loads but also complex and diverse environmental loads, such as waves and ocean currents. It can cause variability in the performance of power generation and the turbines' lifetime compared to onshore turbines. A significant proportion of accidents involving OWTs are caused by excessive

loads on the towers leading to instability, damage, or collapse of the structures [2]. Therefore, continuous load monitoring for the operation and maintenance of OWTs is essential to enhance their stability, and ensure customer profits throughout their life cycle [3,4].

Generally, in order to evaluate the tower stability or lifespan of an offshore wind turbine, the strain of the tower is measured, converted into a load, and then analyzed. The strain of the tower is measured simultaneously with the operational conditions of the turbine and is used to estimate the damage equivalent load (DEL) for each wind speed bin. This DEL can then be applied to life evaluation according to the statistical wind model. However, strain is typically measured through a bridge composed of a strain gage. During this process, the bridge attached to the tower can be contaminated or damaged by static electricity or lightning strikes that occur along the tower.

Wind turbines have a SCADA (supervisory control and data acquisition) system that measures data needed for power generation and management control. In addition, a condition monitoring system (CMS) can be installed on the drive train, with various vibration sensors installed to monitor its condition. But, CMS often experiences significant failures. While the SCADA system is widely recognized as an essential system, implementing the CMS comes with additional expenses and effort for long-term operation and management as it requires additional data acquisition systems. However, unlike onshore cases, OWTs are relatively difficult to access, and reducing the number of visits is cost-effective and safer for personnel. Therefore, it is essential to have technology that can extract various indicators for wind turbine heal and management based on a minimal additional system to adequately achieve LCOE through reasonable access to O&M.

Choi et al. [5] developed a fatigue load prediction algorithm for offshore wind turbines using polynomial curve fitting, decision trees, regular linear regression, and artificial neural network techniques. They demonstrated that fatigue load prediction using machine learning is superior to the statistical polynomial curve fitting method.

Noppe, N. et al. [6] employed an artificial neural network (ANN) to forecast the damage equivalent load (DEL) of the fore-aft bending moment at the base of the wind turbine tower. Their study demonstrated that partitioning the dataset based on operational modes, specifically standstill, partial load, and full load, enhances the accuracy of the predictions. Among the various feature selection methods investigated, neighborhood components analysis (NCA) exhibited excellent performance in DEL prediction while utilizing the smallest set of features.

A physically informed machine learning approach was applied with a custom loss function (Minkowski logarithmic error) to predict the DEL of a wind turbine tower (Santos, F. et al. [7]). By comparing the accumulated fatigue load at the transition piece interface of three wind turbines with predictions, they demonstrated that the physically informed model estimated the accumulated damage equivalent moment with an error of less than 3%.

Luis Vera-Tudela et al. [8] proposed a metric to predict fatigue load and evaluate the quality of fatigue load prediction for two wind turbines within a wind farm. They evaluated the quality of fatigue damage prediction in the edgewise and flapwise directions of the blades under six different flow conditions for two wind turbines within the wind farm. Although the accuracy decreased somewhat under wake flow conditions, they show that a reasonable monitoring system can be established based on a neural network model regardless of flow condition differentiation.

Cosack and his colleagues [9] proposed a comprehensive methodology to predict load using the standard wind turbine control signal of wind turbines with an ANN.

Francisco d Santos et al. [10] implemented a methodology to determine the DEL of the bending moment for an OWT tower installed on Jacket-foundation using an artificial neural network. They developed the model using practical wind turbine data, indicating its applicability to other turbines via cross-validation. They also showed that not only SCADA signals but also accelerometer sensor signals increase the prediction accuracy of the bending moment DEL of the Jacket-foundation OWT tower.

In most approaches, load prediction models rely on statistical data collected at 10 min intervals, which may not capture the short-term variability of the loads experienced by the turbines. To address this limitation, recent research has explored the use of SCADA data and acceleration data to develop bending strain prediction models that can capture the behavior of the turbines.

SCADA systems are used to monitor and control the operation of offshore wind turbines. The system collects real-time data on various parameters, such as wind speed, blade pitch angle, and power output. The loads experienced by various turbine components can be effectively estimated through bending strain prediction models using the SCADA data.

Acceleration data offer insights into how the turbine components dynamically respond to environmental loads. By utilizing accelerometers placed on the turbine tower, the acceleration data can be measured. This data can then be utilized to create prediction models for bending strain, enabling to capture the turbine's short-term load variability.

In this paper, we developed prediction models with a 50 Hz resolution for the bending strain at the tower bottom of OWTs by combining SCADA data and acceleration data using the machine learning technique. Firstly, we collected the necessary data from the SCADA and CMS systems, ensuring that they were synchronized during the data collection. Secondly, we applied a low-pass filter to the collected data and resampled it to a target frequency of 50 Hz. Thirdly, we developed a prediction model using various machine learning algorithms, including ANN, RNN, LSTM, and GRU, to predict the bending strain of offshore wind turbine towers. We evaluated the accuracy of these models and found that the LSTM algorithm outperformed the others in terms of prediction accuracy for the bending strain.

Along with prediction, mean absolute percentage error (MAPE) is analyzed in detail: (1) MAPE for each prediction; (2) MAPE along time history, and (3) wind speed for three zones including idling (blade pitching to 0 deg.), run-up (blade rotating without generator torque) and normal operation (power production).

Section 2 will provide further details about the measurement campaign. In Section 3, the machine learning theory, characteristics of the training dataset, and the preprocessing steps for predicting the bending strain at a temporal resolution of 50 Hz are discussed. Section 4 presents and analyzes the results of the bending strain prediction. Finally, Section 5 concludes the paper.

2. Measurement Campaign

Figure 1a shows the target offshore wind turbine and configuration of strain gauges as a full bridge for measuring the bending strain. A data acquisition system (DAS) is installed in offshore wind turbine with a suction bucket foundation. The OWT has a rated power of 4.2 MW at a wind speed of 11.3 m/s. Two 3-axis accelerometers have spectral range of 0~2.4 kHz with 10% error and locate to measure acceleration at the tower top and bottom with sampling rates of 1.5 kHz. Both accelerometers are positioned by a magnetic base with additional epoxy bonding. A 90-degree strain gauge as full-bridge configuration measures the bending strain at the tower bottom with 50 Hz sampling rates, as shown in Figure 1b. Each SCADA data in this paper are acquired with 0.5 Hz sampling rates. All of the quantity measured and acquired is synchronized at the start of measurement. The motivation for each measurement is that (1) accelerations of higher sampling rates can characterize short-term time series, (2) SCADA with lower sampling rates affect the long-term behavior of time response, and (3) bending measurement synchronized with other measurements is for verification of predictions.

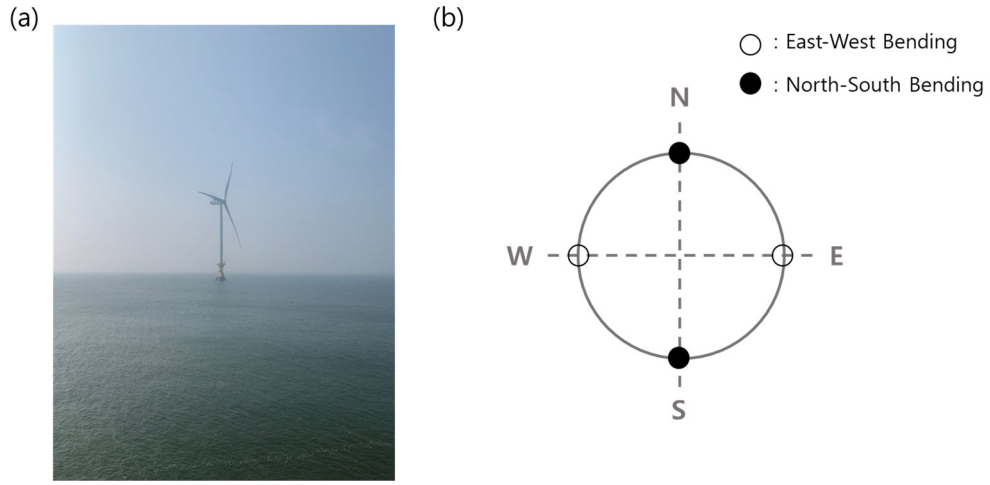


Figure 1. (**a**) Target offshore wind turbine; (**b**) configuration of strain gauges as a full bridge for measuring bending strain.

3. Data Description and Methodology

In order to predict the bending strain of a wind turbine tower, a machine learning model was employed. As mentioned in the introduction, machine learning has been widely used in various studies [11–14], with notable models including ANN, recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU). In this research, ANN, RNN, LSTM, and GRU models were utilized to predict the bending strain of the wind turbine tower, considering its high temporal resolution. The obtained results were compared and analyzed. The following provides a brief explanation of each machine learning model.

3.1. Theory of Neural Networks

Deep learning, a field of machine learning, is currently the subject of extensive research due to its ability to address the problem of overfitting [15,16], its ability to increase processing speed, and due to the development of data science technologies, such as big data. ANNs form the foundation of deep learning algorithms and are composed of perceptrons that imitate the information transmission mechanisms of neurons [17]. Typically, the structure of ANNs consists of three layers: the input layer, the hidden layer, and the output layer. Each layer's weights are sequentially updated from the input layer through the hidden layer to the output layer to minimize the error between predicted values and actual values through a method known as feedforward neural network (FFNN) [18]. When an ANN has two or more hidden layers, it is called a multilayer neural network [19], while a network with more than two hidden layers is known as a deep neural network or deep learning. Figure 2 depicts the typical structure of ANNs, and the mathematical formulation is as follows [20].

$$\hat{y}(x, w) = f\left[\sum_{j=1}^{n_h} w_j \cdot h\left(\sum_{i=0}^{n_d} w_{ji} x_i + b_{kj}\right) + b_0\right] \quad (1)$$

where, w represents the parameter to be optimized, h represents the activation function and a linear function f for those in the output layer and subscript k is the number of perceptron in the hidden layer. Figure 3 illustrates a typical structure of a RNN. As proposed by Rumelhart et al. (1986) [21], RNN is a type of deep neural network (DNN) that differs from the primary FFNN, where activation signals only move from the input layer to the output layer. Instead, RNN employs a parallel chain structure, connecting previous neural network results to the current network's learning process. This structure performs excellently processing continuous data, such as time series data [22]. However, RNN is highly dependent on the previous state, as the gradients in each output unit rely heavily on the current state. When there are numerous neuron units or input units, the error

value can exponentially increase (gradient exploding) or rapidly converge to zero (gradient vanishing) due to the weights being repeatedly multiplied through the learning capability of the past. To address these drawbacks, the LSTM algorithm was proposed [23].

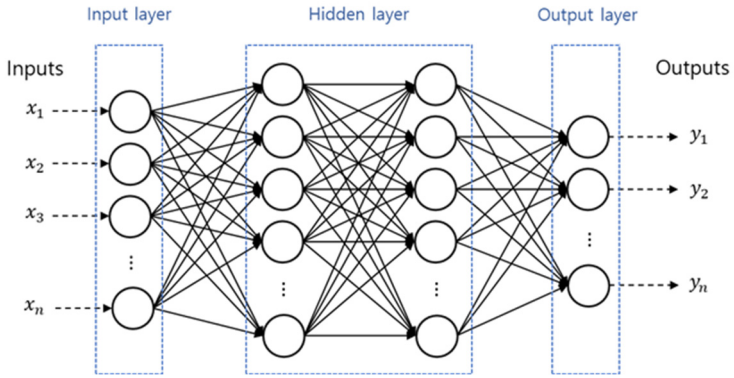


Figure 2. Structure of artificial neural network.

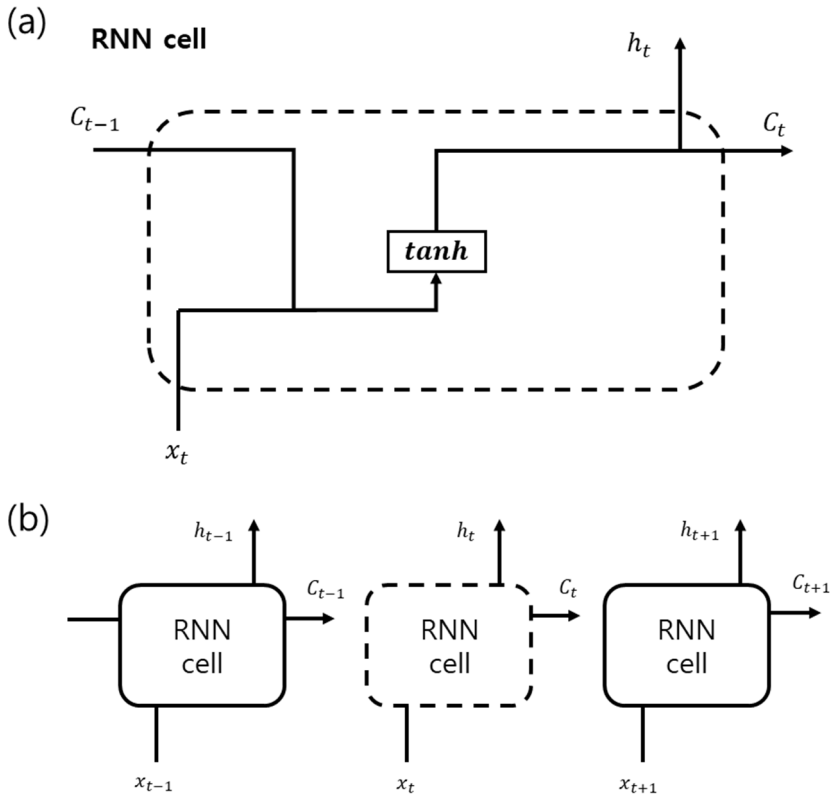


Figure 3. Structure of RNN: (**a**) unit structure and (**b**) timeline expansion diagram.

Figure 4 illustrates the typical structure of an LSTM. In the LSTM unit, a memory cell is used to maintain its state over a long time, and three nonlinear gate structures, namely the input, output, and forget gates, regulate the flow of data inside and outside the cell. That is, LSTM introduces the concept of a cell (C_t) to update the state (h_t) at a specific time, deciding whether to update the internal information based on the input and the state up to the current time. The forget gate (f_t), expressed in Equation (2), applies the previous cell's output (h_{t-1}) and the current input (x_t) to a sigmoid activation layer to obtain a value between 0 and 1. This value is then multiplied element-wise with the current state to determine whether to retain or discard this information.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (2)$$

where, σ is the activation function, and W_f and b_f are the weight of the forget gate and the bias term, respectively. For the input gate i_t, as shown in Equation (3), the decision on what information to store in the cell is determined through two distinct steps as follows: (1) utilizing the sigmoid function to determine what to update, and (2) generating the candidate cell state ($\widetilde{C}_t$) for updating the new cell state using the hyperbolic tangent function as shown in Equation (4).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

where, W_i and W_c represent the weights of the input gate and the candidate cell, respectively, while b_i and b_c represent the biases of the input gate and the candidate cell, respectively. Next, the current cell state (C_t) is updated by combining the past cell state (C_{t-1}) with the cell candidate ($\widetilde{C}_t$), as shown in Equation (5).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{5}$$

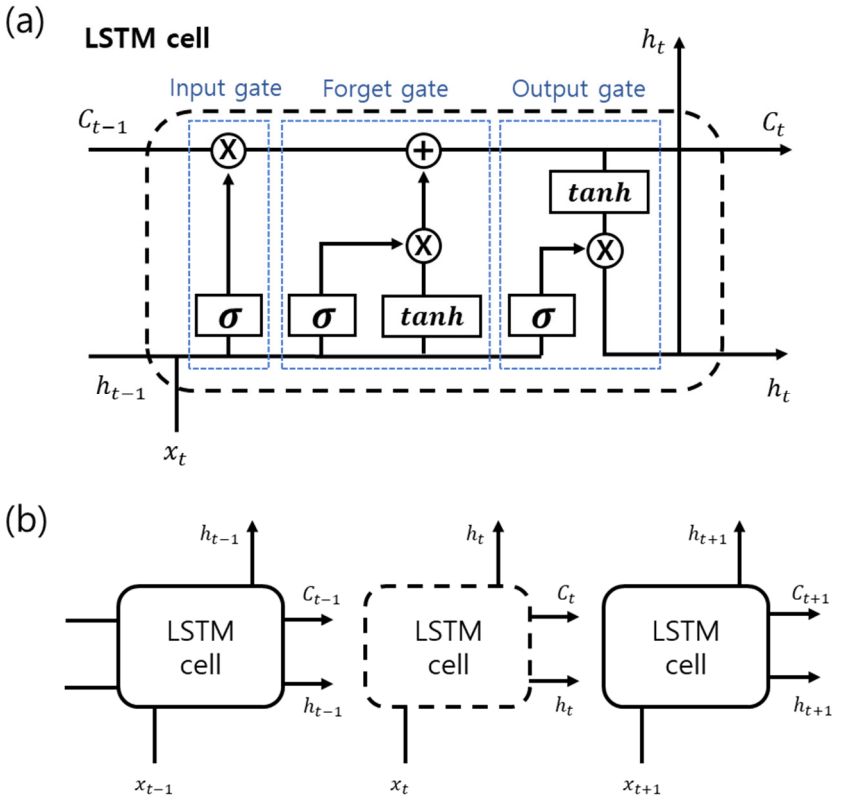


Figure 4. Structure of LSTM: (**a**) unit structure and (**b**) timeline expansion diagram.

The output gate (o_t) determines which parts of the cell state to output using the sigmoid function, as shown in Equation (6). Lastly, by multiplying the activated cell state (C_t) by the hyperbolic tangent function, as depicted in Equation (7), the state at a specific time (h_t) is updated.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{7}$$

where, W_o represents the weights of the output gate, and b_o represents the bias.

Figure 5 illustrates the general structure of a GRU, which was developed as a solution to address the limitations of initial recurrent neural networks (RNNs), similar to LSTM. Unlike LSTM, which incorporates both a hidden state and a cell state vector, GRU is solely composed of a hidden state vector. It performs computations on input data through the following sequential steps [24]:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{8}$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{9}$$

$$\tilde{h}_t = tanh\left(W_h \cdot \left[r_t \bigotimes h_{t-1}, x_t\right] + b_h\right) \tag{10}$$

$$h_t = r_t \bigotimes h_{t-1} + (1 - z_t) \bigotimes \tilde{h}_t \tag{11}$$

where, r_t, and z_t represents the coefficient of reset gate and update gate, respectively, and $\tilde{h}_t$ is the candidate of current cell state. The computation of $\tilde{h}_t$ is performed by combining the previous time step's hidden state, h_t, and the input data, x_t, through the reset gate. The update gate, z_t, controls the information flow from the previous hidden state and the candidate information to determine the hidden state. GRU stands out for its compact architecture, characterized by fewer gates and parameters compared to LSTM. This design choice enables GRU to exhibit faster processing speed and improved performance during training. Machine learning model used in this paper is trained using Keras v. 2.11.0 [25].

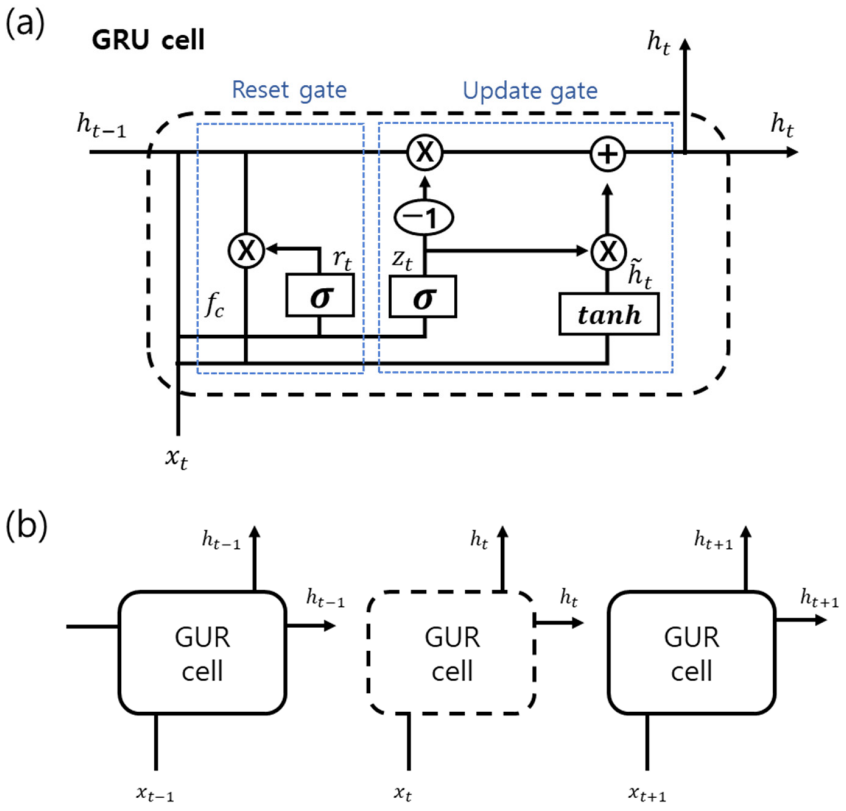


Figure 5. Structure of GRU: (**a**) unit structure and (**b**) timeline expansion diagram.

3.2. Data Description

Table 1 presents the list of SCADA, acceleration, and bending strain data used for wind turbine tower bending strain prediction.

Table 1. Description of datasets from the measurement campaign.

	Sensor	Description	Sampling Frequency	Units
Input	SCADA	Generator speed	0.5 Hz	RPM
		Turbine power		kW
		Nacelle position		Deg.
		Wind speed		m/s
		Wind direction		Deg.
		Nacelle outside temperature		°C
		Blade drive current B1		A
		Pitch angle B1		Deg.
		Pitching speed B1		RPM
		Gearbox oil temperature		°C
		Gearbox oil pressure		Bar
		Deviation of wind direction and nacelle position		Deg.
	Accelerometers	Tower bottom X-vibration	1500 Hz	g
		Tower bottom X-vibration		g
		Tower bottom X-vibration		g
		Main bearing X-vibration		g
		Main bearing X-vibration		g
		Main bearing X-vibration		g
Target	Strain gauges	Tower bottom EW bending moment	500 Hz	mV/V
		Tower bottom NS bending moment		mV/V

The SCADA data were measured at a sampling rate of 0.5 Hz, while the acceleration data were captured at 1500 Hz and the bending strain data at 500 Hz. In order to predict the bending strain at a target frequency of 50 Hz, a re-sampling process was conducted to harmonize the sampling rates of these different signals. The SCADA data underwent up-sampling, while the bending strain and acceleration data were down-sampled, resulting in a consistent sampling rate of 50 Hz for all datasets. Figure 6 illustrates the resampling method. Linear interpolation was applied to the internal values of the resampling period for up-sampling, and decimation was performed on the internal values for down-sampling.

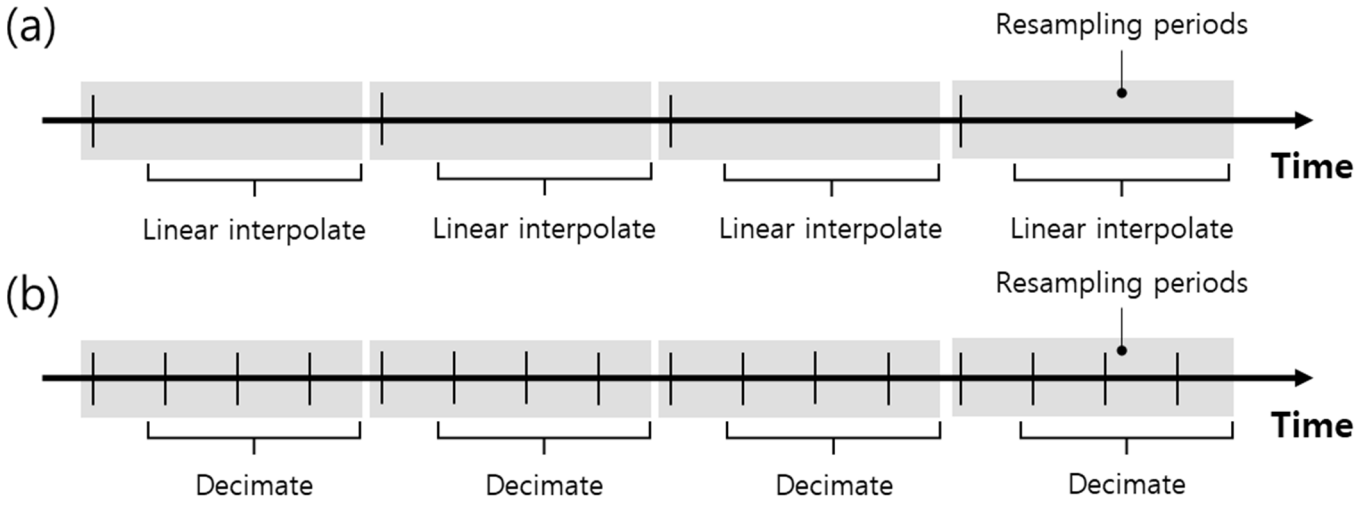


Figure 6. Schematic description of resampling method: (**a**) up-sampling and (**b**) down-sampling.

Figure 7 depicts representative SCADA, acceleration, and bending strain plots that have completed the resampling process. The data were resampled to a frequency of 50 Hz and utilized for training purposes of the bending strain prediction model.

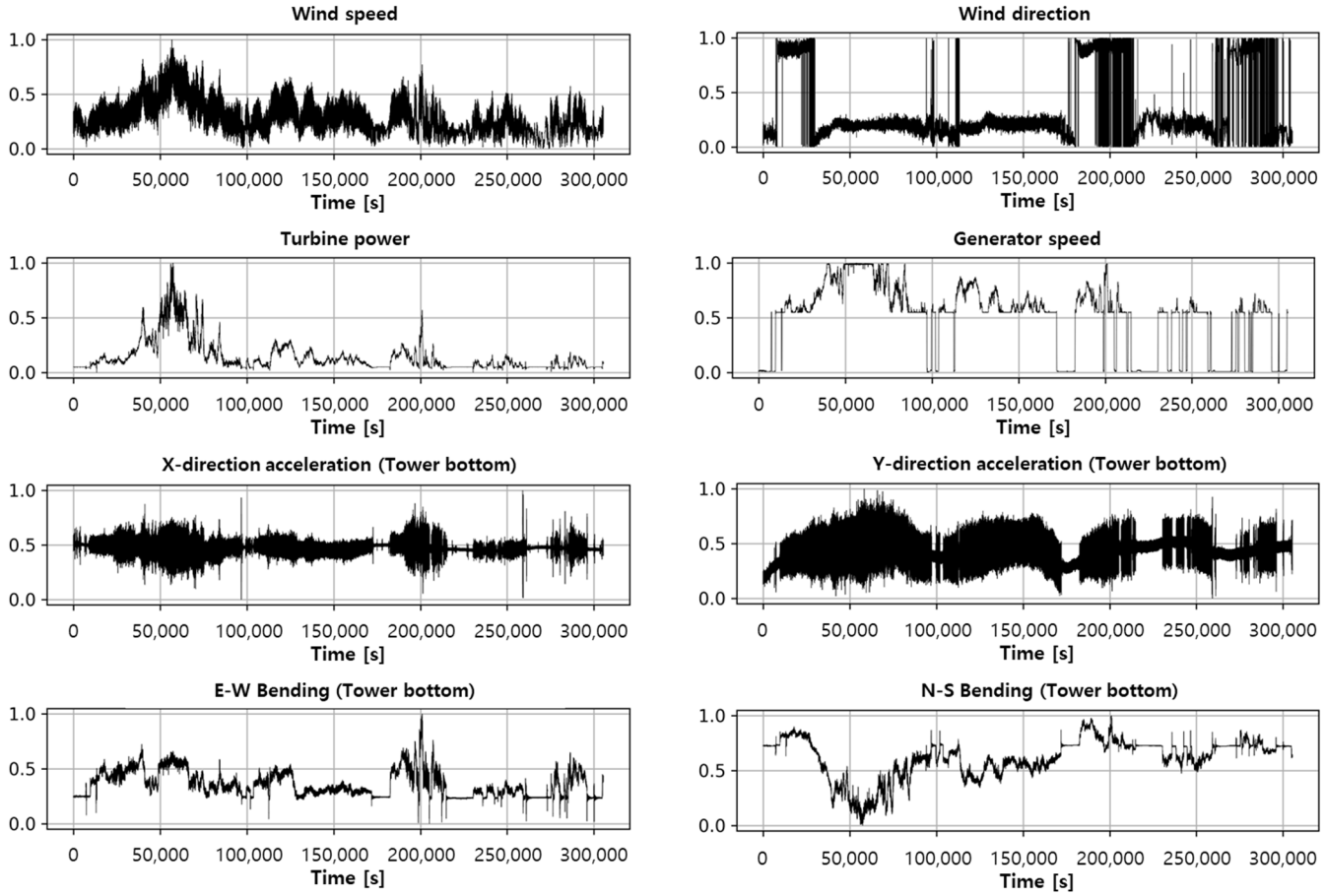


Figure 7. Normalized time series of representative SCADA, acceleration and bending strain for learning.

For training purposes, this study utilized a total of 15,269,998 data points collected at a sampling rate of 50 Hz, spanning a duration of 305,400 s (equivalent to approximately 3.5 days). These data were then partitioned into training data, accounting for 70% of the total, and validation data, comprising 20% of the total. Furthermore, for model testing, another set of 2,159,950 data points, encompassing a time span of 43,199 s (equivalent to 12 h), was employed. In general, the most influential environmental variables for wind turbines are wind speed and direction [26]. The static and dynamic characteristics of wind turbines in normal operating conditions strongly depends on wind speed and direction. Therefore, a preliminary analysis of wind speed and wind direction information at the time of data measurement is an essential prerequisite for improving the accuracy of predicting the bending strain of the wind turbine tower.

In Figure 8, the wind speed histogram is illustrated in conjunction with the Weibull probability density function. The wind speed is distributed with 0.2 m/s intervals. The most frequent occurrence of wind speeds within the measurement range is observed at 2.4 m/s, while wind speeds exceeding 12 m/s are extremely rare. The Weibull distribution parameters indicate a shape parameter of 2.27 and a scale parameter of 4.31.

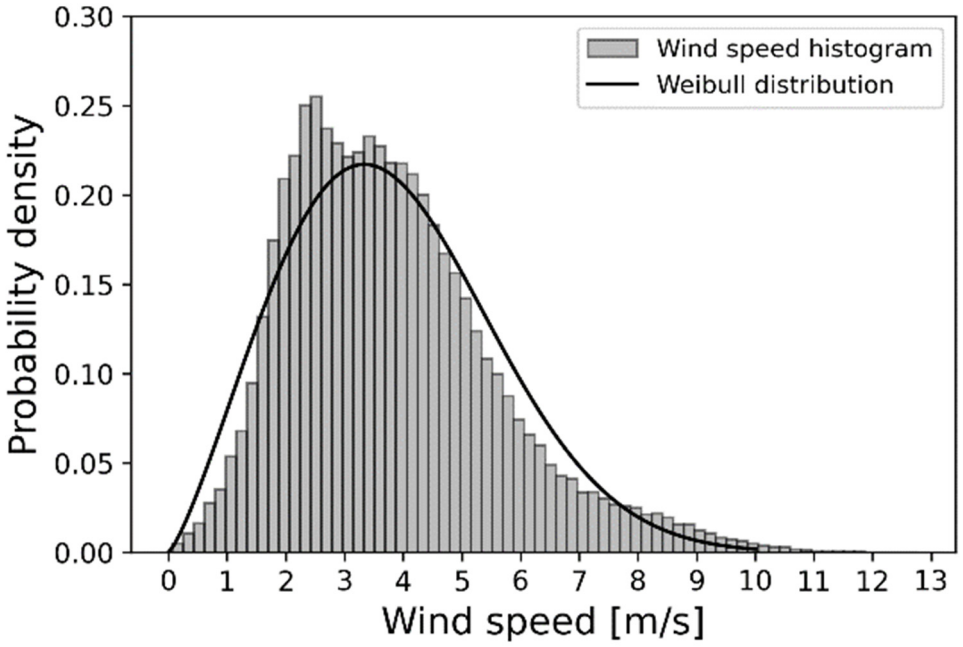


Figure 8. Histogram and Weibull distribution of wind speeds.

Figure 9 represents a wind rose generated using wind direction and speed data. The measured data exhibit wind directions ranging from east to north, with the dominant wind direction being ENE (east and northeast).

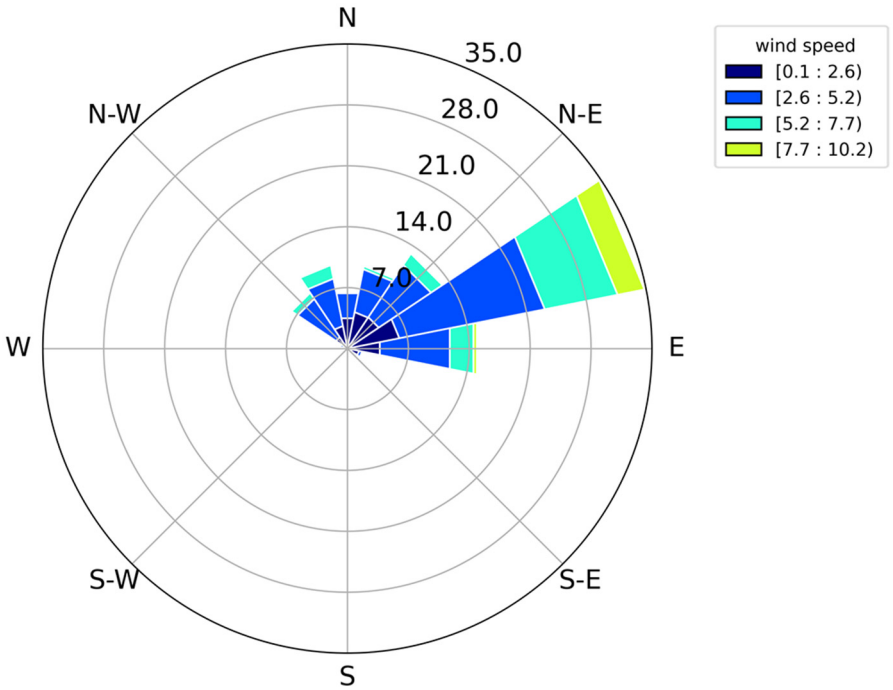


Figure 9. Wind rose plot at the site of the wind turbine from SCADA data.

The power performance of a wind turbine is considered one of the most crucial factors among various indicators that represent the characteristics of the turbine. The operation of a wind turbine is controlled based on the wind speed, ranging from stop state to generation state (lager than cut-in speed, 3 m/s). The control of the wind turbine through its operation state also influences the static and dynamic structural behaviors. To investigate the range of wind speed and electric power applied in this paper, the power curve is indicated in

Figure 10. The power curve of the turbine shows the relationship between wind speed and turbine power. As observed therein, the measurement represents data below rated wind speed of 11.3 m/s, and most of the data locate around cut-in speed. Firstly, computations of the average for each power during 1 min are conducted, and then the average is referred to as the Mean values. The maximum and minimum for each measurement are also calculated and then named as Max values and Min values, respectively. Finally, Figure 10 indicates Mean, Max, and Min values normalized by the maximum Mean values. It is seen that the wind speed range is about 1 m/s to 10 m/s and most of the data collapse under a wind speed of 6 m/s.

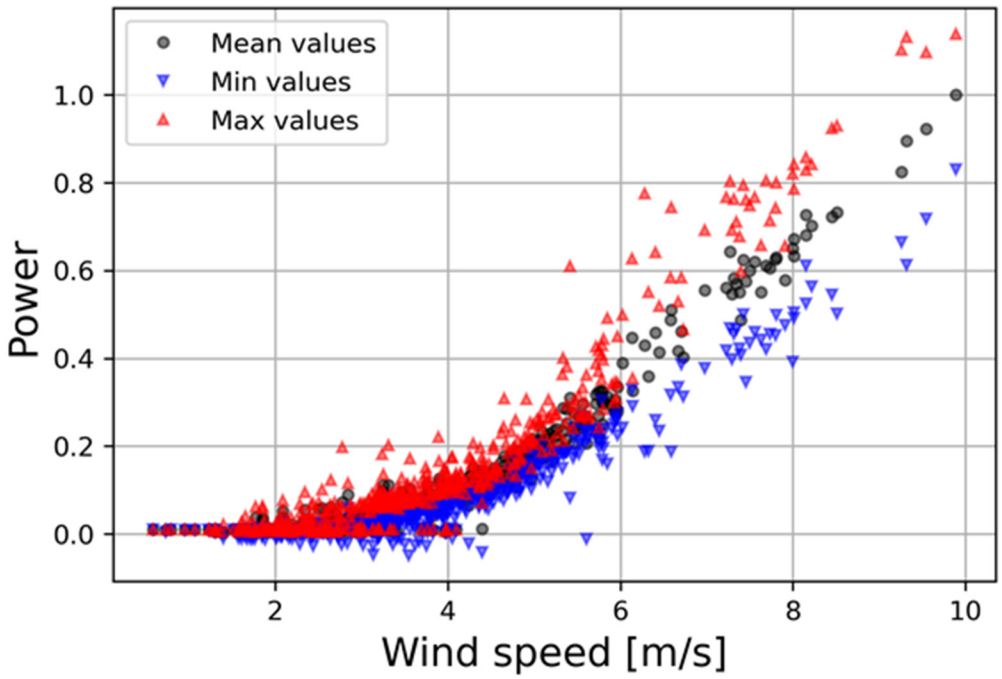


Figure 10. Normalized power curve statistics for the training period.

Figure 11 represents the capture matrix of wind speed within the training range. It consists of wind speed and turbulent intensity binned with a step of 1 m/s and 2%, respectively. The collected wind speed data are all averaged over 10 min. According to the capture matrix, the most frequently collected wind data correspond to a speed of 4 m/s and a turbulent intensity value in the range of 0.15 < v < 0.17. Additionally, it can be observed that lower wind speeds are associated with higher turbulent intensity values, while higher wind speeds are associated with lower turbulent intensity values.

3.3. Neural Network Construction

We conducted feature selection for variables associated with the tower bending strain in wind turbine SCADA data using the Pearson correlation coefficient. The Pearson correlation coefficient is an indicator that quantifies the correlation between two sets of variables. It ranges between −1 and +1, where values close to −1 indicate a strong negative correlation and values close to +1 indicate a strong positive correlation. The following equation expresses the correlation coefficient $Corr(x, y)$ between two variable sets, x, and y.

$$Corr(x,y) = \sum_{i=1}^{M} \left[\left(x_i - \bar{x}\right) \cdot \left(y_i - \bar{y}\right) \right] / \sqrt{\left(x_i - \bar{x}\right)^2 \left(y_i - \bar{y}\right)^2} \quad (12)$$

where, x_i and y_i represent the i-th data points of variables x and y, respectively, while $\bar{x}$ and $\bar{y}$ denote the means of the sample populations x and y, respectively. Figure 12 shows the correlation coefficients calculated using Equation (12), visualized as a heatmap.

TI Range	0 m/s	1 m/s	2 m/s	3 m/s	4 m/s	5 m/s	6 m/s	7 m/s	8 m/s	9 m/s
0 < v <= 0.05	0	0	0	0	0	0	0	0	0	0
0.05 < v <= 0.07	0	0	0	0	0	0	0	0	0	0
0.07 < v <= 0.09	0	0	0	0	0	0	0	0	0	1
0.09 < v <= 0.11	0	0	0	0	0	0	0	2	4	2
0.11 < v <= 0.13	0	0	0	0	1	2	0	11	3	1
0.13 < v <= 0.15	0	0	0	1	13	21	6	6	0	0
0.15 < v <= 0.17	0	0	0	8	47	25	3	0	0	0
0.17 < v <= 0.19	0	0	1	24	20	11	2	0	0	0
0.19 < v <= 0.21	0	0	6	29	6	2	2	0	0	0
0.21 < v <= 0.23	0	0	15	18	3	0	0	0	0	0
0.23 < v <= 0.25	0	2	15	11	1	0	0	0	0	0
0.25 < v <= 0.27	0	3	6	5	0	0	0	0	0	0
0.27 < v <= 0.29	0	2	5	2	1	0	0	0	0	0
Total	0	7	48	98	92	61	13	19	7	4

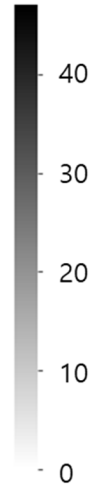


Figure 11. Capture matrix for the training period.

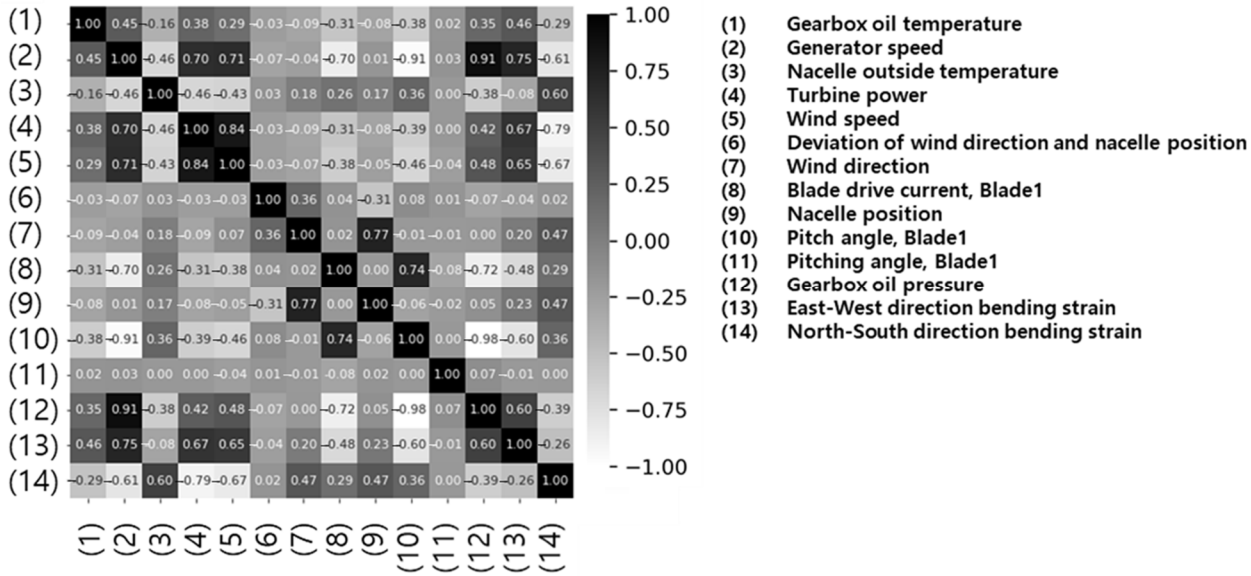


Figure 12. Heatmap of the pairwise Pearson correlation coefficient between each pair of the measured 12 SCADA and 2 bending strain data.

Generally, a correlation coefficient of $|Corr(x, y)| \geq 0.1$ is considered to indicate a weak correlation. Therefore, in this study, if the coefficient of either the east–west or north–south bending strain exceeded 0.1, it was selected as a feature for training the model. Additionally, the features (6) deviation of wind direction and nacelle position and (11) pitching angle were included, despite having lower correlation coefficients, due to their close association with operational domain and mis-control of the wind turbine. Therefore, we selected a total of 18 features, including 12 features of SCADA data and 2 3-axis accelerometer data, to train the bending strain prediction model for wind turbine towers. In this study, the accelerations of 1500 kHz and bending strains of 500 Hz were down-sampled to 50 Hz for the prediction. Down-sampling was performed using a simple decimation method, as depicted in Figure 6b. However, down-sampling through simple decimation can introduce undesired noise signals, including aliasing. A low-pass filter (LPF) was applied to the raw data prior to resampling to mitigate the generation of unwanted signals, such as aliasing.

The gain $G(\omega)$ of an nth-order Butterworth low-pass filter is expressed in terms of the transfer function $H(j\omega)$ as follows:

$$G(\omega) = |H(j\omega)| = \frac{G_0}{\sqrt{1 + \left(\frac{j\omega}{j\omega_c}\right)^{2n}}} \tag{13}$$

where, n is the order of filter, ω_c is the cutoff frequency and G_0 is the gain at zero frequency.

The input and output data in the training have different units, as indicated in Table 1. When data values have different ranges due to their respective units, it can lead to biased learning. Therefore, in this study, we normalized the input and output data to reduce the impact of dimension and increase cohesion between the data. We used a min–max scaler for normalization. The normalization (z_i) and denormalization (x_i) equations are as follows:

$$z_i = (x_i - \min(x))/(\max(x) - \min(x)) \tag{14}$$

$$x_i = z_i(\max(x) - \min(x)) + \min(x) \tag{15}$$

where, z_i represents the normalized data, x represents the actual values, and x_i represents the i-th actual value of the data.

To evaluate the performance of the model identification, we used the following metrics [27]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{16}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - y\prime)^2} \tag{18}$$

The parameters manually selected in the machine learning model are referred to as hyperparameters. In this study, hyperparameter optimization was conducted for the machine learning model, specifically focusing on the cell size, batch size, and epochs. Consequently, the optimization for the machine learning models used in this study was configured as shown in Table 2.

Table 2. Structure and hyperparameters of selected machine learning models.

Model	**ANN**		**RNN**		**LSTM**		**GRU**	
Sensor Direction	**E-W**	**N-S**	**E-W**	**N-S**	**N-S**	**E-W**	**E-W**	**N-S**
No. of perceptron (Dense) [28]	18-256-256-1		-		-		-	
Cell size	-		1024		1024		1024	
Batch size	30,000		30,000		30,000		30,000	
Epochs	300		600	500	100		300	

The structure of the ANN model was adopted from the research by Wang, Z. et al. [28]. Furthermore, to minimize the loss function during model training, an optimizer is necessary, hence, in this study, we employed the adaptive moment estimation algorithm, commonly known as Adam, as the optimizer. The loss function was set to mean square error (MSE). Additionally, the learning rate was set to 0.01.

4. Results and Discussion

Table 3 summarizes the prediction errors of the ANN, RNN, LSTM, and GRU machine learning models for predicting the bending strain of wind turbine towers. The RNN model exhibited the most significant error in prediction, while the LSTM model demonstrated the smallest prediction error. The GRU model also showed high accuracy in its predictions, but the $\mathbf{R}^2$ value was relatively lower than the LSTM model. This indicates that the GRU model has relatively lower accuracy in predicting local variations compared to the LSTM model. Among the machine learning algorithms, the LSTM algorithm exhibited the best performance in predicting the bending strain of the towers. Therefore, the following analysis focuses on the results predicted by the LSTM algorithm for the prediction.

Table 3. Results of prediction errors.

	East–West Direction Bending			North–South Direction Bending		
Model	**RMSE (mV/V)**	**MAPE (%)**	$\mathbf{R^2}$	**RMSE (mV/V)**	**MAPE (%)**	$\mathbf{R^2}$
ANN	0.0182	5.7865	0.7823	0.0096	2.4427	0.9124
RNN	0.0437	14.6669	−0.2599	0.0179	3.9957	0.7276
LSTM	0.0101	3.5445	0.9330	0.0066	1.5027	0.9634
GRU	0.0120	3.7803	0.9051	0.0122	3.2543	0.8741

In Figure 13, we illustrate the time history of the predicted and measured bending strains of the wind turbine tower, along with the absolute percentage error (APE) for each sample corresponding to 50 Hz. Figure 13a depicts the strain in the east–west direction, while Figure 13b represents the strain in the north–south direction. The solid blue line represents the measured values, the solid red line represents the predicted values, and the solid black line represents the APE time history. Both results match well with the measured values.

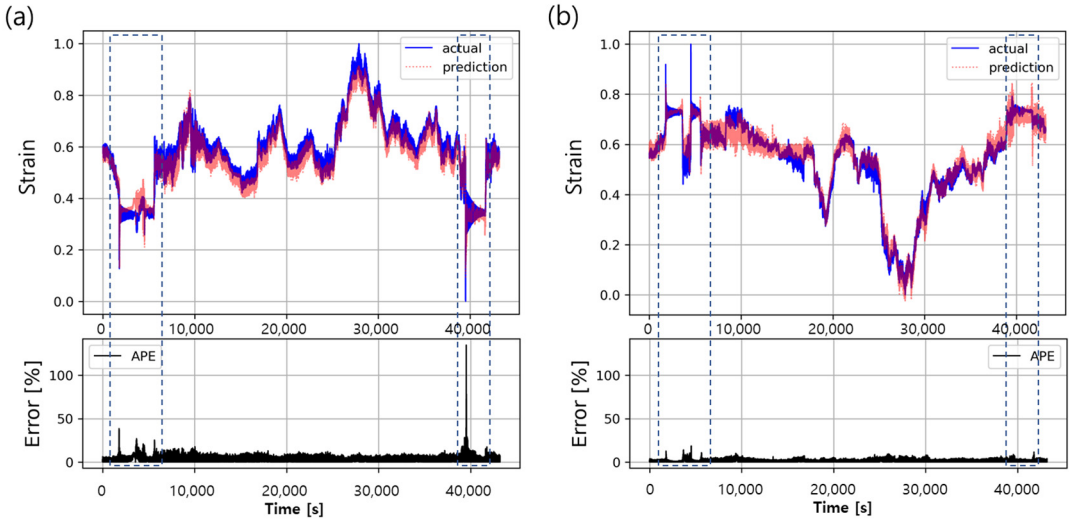


Figure 13. Prediction results of normalized bending strain compared to the measurement results with APE histories; (**a**) east–west direction, (**b**) north–south direction.

Peaks of APE are located between 2000 and 6000 s and between 39,000 and 42,500 s as shown in Figure 13. These peaks occur when there are rapid decreases or increases in the strain. To investigate this in detail, Figure 14 displays the rotational speed of the generator, wind speed, and blade pitch angle. It is observed that APE peaks locate during the period when the wind speed drops below the cut-in speed and enters a stop state (rotor speed is zero), which means that the blade pitch control and rotor is in braking. The inertia

response of the turbine in the stop state varying with wind speed is different from that in the generation state. Braking operation in the stop state can cause the impulsive loads on the wind turbine. Those change in the response and impulsive load is considered as the main reason for the higher APE in this period.

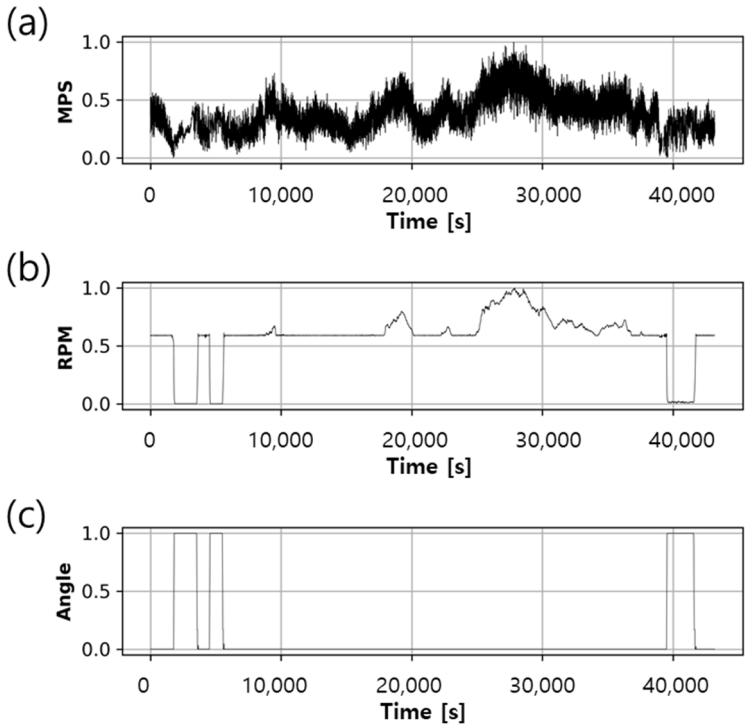


Figure 14. Time series normalized data plot of (**a**) wind speed, (**b**) generator rotational speed and (**c**) blade pitch angle of test data.

Figure 15 illustrates the APE of the prediction according to the wind speed bin. It can be observed that most of the APE peaks are located in the wind speed range below 2 m/s. This range corresponds to the stop state with higher turbulent intensity. It is seen that braking rotor and higher intensity can induce higher peaks of APE.

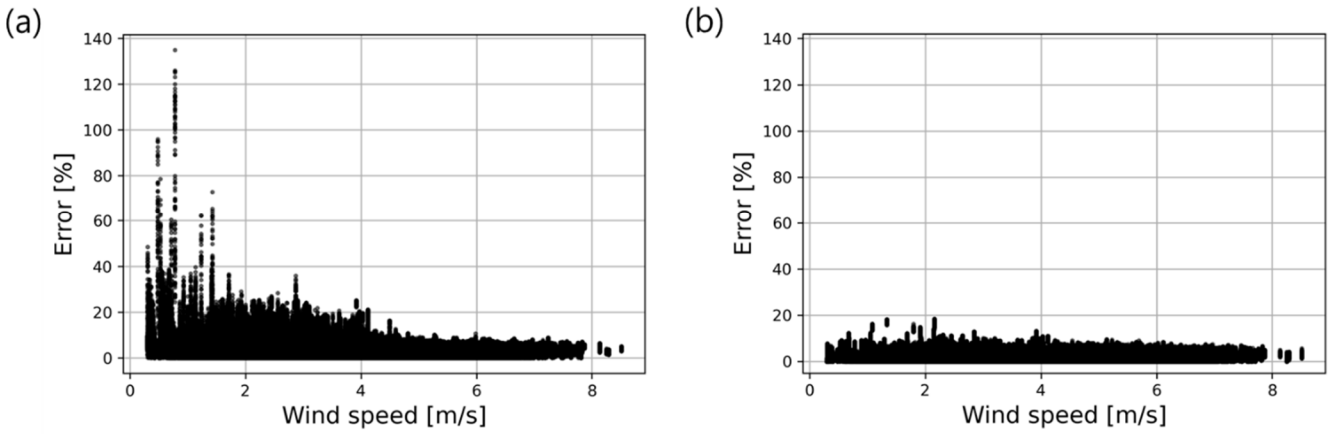


Figure 15. APE between prediction and measurement results plotted against wind speed; (**a**) east–west direction and (**b**) north–south direction.

APE values in the wind speed range below 2 to 4 m/s are generally higher than those in the wind speed range above 4 m/s. At this range, the turbine is in both stages of run-up or generation state, where the generator is in free-torque at the run-up state.

Similar to the range below 2 m/s, the free-torque causes a change in dynamic response of the wind turbine according to wind speed, and then it can induce higher APE than the generation state. While the stop state includes the braking operation, both the run-up and the generation states exclude the braking, resulting in a lower peak of APE than in the stop state.

Figure 16 shows the box plots for the APE binned with a step of 1 m/s based on the wind speed shown in Figure 15. The median and the inter-quartile range are positioned within a very low range of errors. According to Figure 15, the error appears to decrease as the wind speed increases. However, the pattern looks different when observed through statistical distribution, as shown in Figure 16. According to Figure 16, as the wind speed exceeds 6 m/s, both the median and the inter-quartile range (IQR) of the box plot show a slight increase. It indicates that the error increases slightly for wind speeds exceeding 6 m/s.

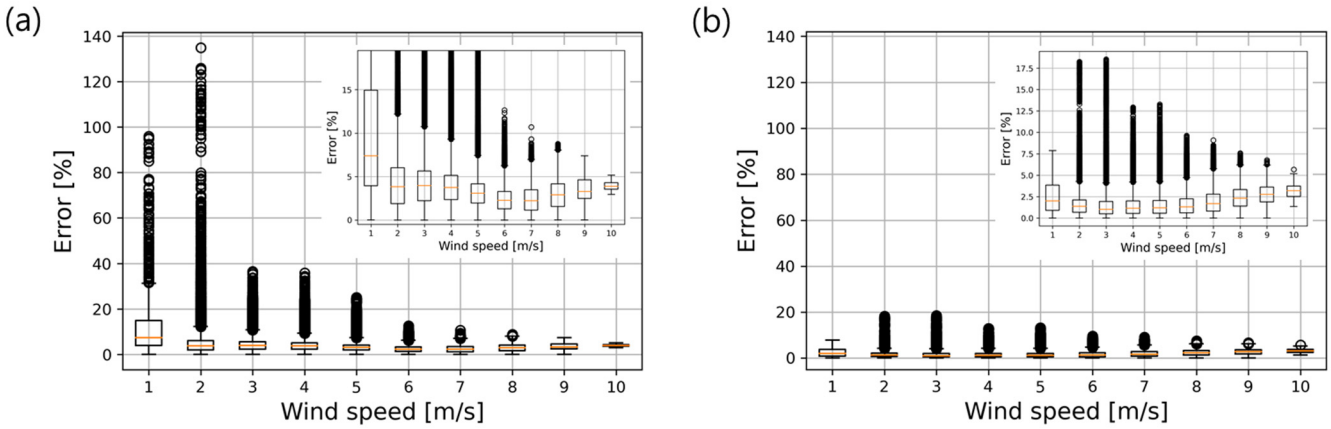


Figure 16. Box plot of APE between prediction and measurement for each wind speed bin and its localization in range 0 to 20% error; (**a**) east–west direction and (**b**) north–south direction.

5. Conclusions and Future Works

In this study, we predicted and analyzed the bending strain of wind turbine towers with a temporal resolution of 50 Hz using data obtained from the SCADA and additional measurements. Firstly, we acquired the necessary data for training from the SCADA and additional systems, ensuring their synchronization upon acquisition. Secondly, resampling is applied to the measured data after low-pass filtering. Thirdly, machine learning techniques to predict the offshore wind turbine tower's bending strain are developed, where machine learning algorithms included the ANN, RNN, LSTM, and GRU models.

In the analysis of the prediction results, the bending strain is compared with measurement and accuracy is computed as APE. The results demonstrated that the LSTM algorithm outperformed the others in terms of prediction accuracy for the wind turbine tower's bending strain.

However, it should be noted that higher prediction errors were observed in low wind speed regions. This can be attributed to factors, such as high turbulent intensity during strong winds and the impact of dynamic response change resulting from blade pitch control and free-torque of a generator during the stop state and the run-up. Moreover, in the stop state, the braking operation is considered as the main reason for the higher peak of APE in the prediction of the strain. Despite this, our research presented the possibility of estimating the remaining lifespan of wind turbines based on existing controller data and commonly required CMS data without the additional long-term load measurement systems.

This approach allows for the utilization of a dataset with a higher temporal resolution, offering various indicators for effective wind turbine management, such as DEL, spectral response, and peak detection. Thus, the methodology and prediction procedure are key innovations in our study. While many machine learning studies focus on accuracy in time series prediction, it is essential to consider the operational conditions of wind turbines,

including stop, run-up, and generation states. Moreover, it is observed that lower prediction accuracy during stop and run-up states can be attributed to factors like higher turbulent intensity, rotor braking, and blade pitching in lower wind speed ranges. It is important to note that the direct application of machine learning models to dynamic or static responses can lead to incorrect results in specific wind speed ranges, despite satisfactory accuracy in overall time series analysis. These findings provide valuable guidelines for integrating machine learning into wind turbine analysis.

Author Contributions: G.-S.L. provided the basic idea for this study and supervised the research; S.L. performed the utilized machine learning algorithms and performed the analysis of the results; S.K. developed a measurement system to acquire data of the wind turbine. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20228520020040) and funded by the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20203020020010).

Acknowledgments: This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE)(20228520020040, Development of O&M scheduling technology for large-scale offshore wind farms optimized in domestic environments) and the Korea Institute of Energy Technology Evaluation and Planning(KETEP), and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20203020020010, Technology Enhancement of Testing and Certification for IECRE).

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

Nomenclature

b_0	Bias at the output layer
w_j	Connection weight between the j_{th} of the hidden layer and the single output perceptron
b_{kj}	Bias at the k_{th} perceptron of the hidden layer
i_t	Input gate
W_r	Weight between the input and h_{t-1} in the reset gate
W_z	Weight between the input and h_{t-1} in the update gate
b_r	Bias between the input and h_{t-1} in the reset gate
b_z	Bias between the input and h_{t-1} in the update gate
h_{t-1}	Standard GRU unit output at time $t-1$
x_t	Input at time t
y_i	Actual *ith* value of y
$\hat{y}_i$	Predicted *ith* value of y
$y\prime$	Mean value of y

References

1. IRENA. *Renewable Capacity Statistics 2023*; International Renewable Energy Agency: Abu Dhabi, United Arab Emirates, 2023.
2. Seyr, H.; Muskulus, M. Safety indicators for the marine operations in the installation and operating phase of an offshore wind farm. *Energy Proc.* **2016**, *94*, 72–81. [CrossRef]
3. Lian, J.J.; Cai, O.; Dong, X.F.; Jiang, Q.; Zhao, Y. Health monitoring and safety evaluation of the offshore wind turbine structure: A review and discussion of future development. *Sustainability* **2019**, *11*, 494. [CrossRef]
4. Liu, R.W.; Yuan, W.Q.; Chen, X.Q.; Lu, Y.X. An enhanced CNN-enabled learning method for promoting ship detection in maritime surveillance system. *Ocean. Eng.* **2021**, *235*, 109435. [CrossRef]
5. Choi, J.; Son, E.; Lee, G.; Kang, M.; Lee, J.; Hwang, S. Development of fatigue load prediction algorithm for wind turbines. *J. Korean Sol. Energy* **2021**, *41*, 93–105. [CrossRef]
6. Noppe, N.; Iliopoulos, A.; Weijtjens, W.; Devriendt, C. Full load estimation of an offshore wind turbine based on SCADA and accelerometer data. *J. Phys. Conf. Ser.* **2016**, *753*, 72025. [CrossRef]
7. Santos, F.; D'Antuono, P.; Robbelein, K.; Noppe, N.; Weijtjens, W. Long-term fatigue estimation on offshore wind turbines interface loads through loss function physics-guided learning of neural networks. *Renew. Energ.* **2023**, *205*, 461–474. [CrossRef]
8. Luis, V.; Martin, K. Analysing wind turbine fatigue load prediction: The impact of wind farm flow condition. *Renew. Energ.* **2017**, *107*, 352–360.

9. Cosack, N.; Kühn, M. An approach for fatigue load monitoring without load measurement devices. *Sci. Proc. EWEA* **2009**, *1*, 69–73.
10. Santos, F.; Noppe, N.; Weijtjens, W.; Devrientd, C. Data-driven farm-wide fatigue estimation on jacket-foundation OWTs for multiple SHM setups. *Wind Energy Sci.* **2022**, *7*, 299–321. [CrossRef]
11. Li, G.; Wang, C.; Zhang, D.; Yang, G. An Improved Feature Selection Method Based on Random Forest Algorithm for Wind Turbine Condition Monitoring. *Sensors* **2021**, *21*, 5654. [CrossRef]
12. Cai, J.; Luo, J.; Wang, S.; Yang, S. Feature selection in machine learning: A new perspective. *Neurocomputing* **2018**, *300*, 70–79. [CrossRef]
13. Xiao, X.; Roh, B.; Hamilton, C. Porosity management and control in powder bed fusion process through process-quality interactions. *CIRP J. Manuf. Sci. Technol.* **2022**, *38*, 120–128. [CrossRef]
14. Xiao, X.; Waddell, C.; Hamilton, C.; Xiao, H. Quality Prediction and Control in Wire Arc Additive Manufacturing via Novel Machine Learning Framework. *Micromachines* **2022**, *13*, 137. [CrossRef] [PubMed]
15. Hinton, G.E. Learning multiple layers of representation. *Trends Cogn. Sci.* **2007**, *11*, 428–434. [CrossRef]
16. Dahl, G.E.; Sainath, T.N.; Hinton, G.E. Improving DNNs for LVCSR using rectified linear units and dropout. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013.
17. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386. [CrossRef]
18. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
19. Rosenblatt, F. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. *Arch. Gen. Psychiatry* **1962**, *7*, 218–219.
20. Samarasinghe, S. *Neural Network for Applied Science and Engineering*; Auerbach Pub.: New York, NY, USA, 2006.
21. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
22. Gers, F.A.; Schraudoph, N.N.; Schmidhuber, J. Learning precise timing with LSTM recurrent networks. *J. Mach. Learn. Res.* **2022**, *3*, 115–143.
23. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural. Comput.* **1997**, *9*, 1735–1780. [CrossRef]
24. Britz, D.; Goldie, A.; Luong, M.T.; Le, Q. Massive exploration of neural machine translation architectures. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processin, Copenhagen, Denmark, 9–11 September 2017; pp. 1442–1451.
25. Chollet, F. Keras. 2015. Available online: https://keras.io (accessed on 19 May 2023).
26. Torres, J.L.; García, A.; Prieto, E.; Francisco, A. Characterization of wind speed data according to wind direction. *Sol. Energy* **1999**, *66*, 57–64. [CrossRef]
27. Kusiak, A.; Verma, A. Analyzing bearing faults in wind turbines: A data-mining approach. *Renew. Energ.* **2012**, *48*, 110–116. [CrossRef]
28. Wang, Z.; Qiao, D.; Tang, G.; Wang, B.; Yan, J.; Ou, J. An identification method of floating wind turbine tower responses using deep learning technology in the monitoring system. *Ocean Eng.* **2022**, *261*, 112105. [CrossRef]

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