Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

A NOVEL APPROACH FOR BUILDING WIND POWER OUTPUT PREDICTION BASED ON DEEP LEARNING RNN AND CNN MODEL

Aisha Rasheed[1], Abbas Raza[2], Sehar shakir[*3]

DOI: https://doi.org/10.5281/zenodo.20320518

Keywords
Forecast, Wind power, RNN, CNN, LSTM, GRUs

Article History
Received: 22 March 2026
Accepted: 01 May 2026
Published: 21 May 2026

Corresponding Author: *
Sehar shakir

Abstract
Wind power has come out as a quickly growing source of renewable energy now a days. Wind is intermittent in nature due to wind speed alterations. Accurate prediction of wind power is essential for efficient operation of wind power system, in return providing power network management and control. Modern wind turbines have Supervisory Control and Data Acquisition (SCADA) systems are installed for wind power forecast. In this research, LSTM, GRU and CNN-LSTM based deep learning models are constructed. To evaluate the prediction performance of neural network-based models, a novel comparative analysis is performed utilizing the lookback parameter. In the predictive model, wind speed, nacelle orientation, yaw error, blade pitch angle, and ambient temperatures were considered as input features, while wind power was assessed as an output feature. Input features were directed in models based on wind turbine physical process. The deep learning models have been given training, testing and validation against SCADA measurements. The study of the results reveals GRU gives minimum MAPE value of 0.074 having slight difference with MAPE values of Hybrid (CNN-LSTM)'s MAPE of 0.0751 and Bi-LSTM giving MAPE of 0.078. The simulation result showed that proposed GRU model gives suitable MAE values of validation and MAPE values of predictions in comparison to Hybrid (CNN-LSTM) and Bi-LSTM retaining more accuracy with less computational cost.

INTRODUCTION

The role of renewable energy is becoming significant for environmental protection and sustainable development of economy. The exploding pollution of exhaustible human resources is causing greenhouse effect [1]. The continuous depletion of fossil fuels including oil, natural gas and coal, and overconsumption of energy in this developed global economy, has turned the world's perspective towards clean and viable energy sources for electricity production [2]. Wind energy has become one of the most promising options because of its abundant availability and year-round harvesting feature. The energy from wind turbine has become one of the rapidly growing technologies for developed as well as developing nations [3]. According to the National Renewable Energy Laboratory (NREL) report, Pakistan has the potential of 346 GW for wind power generation. Wind energy is inherently intermittent, and its characteristic of instability rest bounds its integration in national power grid [4]. However, accurate prediction of wind power and wind speed is an effective solution to overcome this difficulty and reduce the uncertainty of wind energy [5],[6]. The integration of wind power plants in power grid demands at least two to three days of advanced estimated power. The issue faced by wind energy for prediction of power output is due to incorrect or less accurate weather forecast. Three approaches have been used for wind energy and power forecast including physical models, statistical models, and hybrid models. These models reveal the relationship between output power and input parameters including wind speed and other variables. Physical methods are mainly centered on numerical weather prediction (NWP) approach which provides improved performance in the case of the long-time horizon. These models obtain microscale information of metrological data (wind direction, humidity, temperature, air density, and atmospheric pressure, etc.) and perform correlation analysis to combine the technical parameters which are used for the prediction of power output [7]. For medium- and long-term forecast, physical models are suitable, but excessive empirical and complex parameters cannot give high accuracy predictions in real-time. These models require plenty of time in computation and cost additional resources as well. The second type of model is the statistical model. These models usually consider time series of wind speed and wind power as the basis. These models develop the relationship by mapping input characteristics and output characteristics through historically spanned data [8]. These models are easy to construct using parameters approach checking and procedures of pattern recognition. Statistical models are suitable for short and medium-term forecasts, but they are unable to tackle the non-linearities with higher accuracy. Third and well-known approach for handling the non-linearities because of their fault tolerance and strong robustness is hybrid models [9]. These models have good adaptability and pertinence to wind power output prediction.

Nomenclature

x_i	Predicted value	$\hat{x}_i$	True value
σ	Sigmoid function	x_{min}	Min. value of input
x_p	Initial value of input	i_t	Input gate
f_t	Forget gate	o_t	Output gate
z_t	Update gate	r_t	Reset gate
i_t	Input state	h_t	Hidden state
$w_i, w_f, w_o,$	Weights of input, forget and output states	b_i, b_f, b_o	Biases of input, forget and output states
$w_z, w_r,$	Weights of update and reset states	$b_z, b_r,$	Biases of update and reset states
$U_z, U_r,$	Hidden weights of update and reset states		

ANN	Artificial Neural Network	BPNN	Back Propagation Neural Network
ANFIS	adaptive Neuro-Fuzzy Inference System	ICA	Imperialistic Competitive Algorithm
ARIMA	Autoregressive integrated moving-average	RMSE	Root Mean Square Error
LSTM	Long Short-Term Memory	MAE	Mean Absolute Error
PCC	Pearson Correlation Coefficient	MLP	Multi-layer Perceptron
RNN	Recurrent Neural Network	GRU	Gated Recurrent Unit
NWP	Numerical Weather Prediction	FFBP	Fed-forward back propagation

RBFNN	Radial Basis Function Neural Network	DGF	Double Gaussian function
CNN	Convolutional Neural Network	MSE	Mean Square Error
MAPE	Mean Absolute Percentage Error	AIC	Akaike's Information Criterion
SVM	Support Vector Machine	RF	Random Forest
NARX	Non-linear autoregressive exogenous	IF	Isolation Forest
WNN	Wavelet neural network	RKF	Recurrent Kalman Filter
FS	Fourier Transform	PSO	Particle Swarm Optimization

By applying Deep Learning hybrid model to the desired wind farm, we aim to make a wind energy forecast model with a leading time of few hours to days. The problem focuses on implementation of hybrid model that minimizes the excessive shortfalls in energy production and optimizes the profit level for wind farms. Wind power is predicted in an accurate manner by ANN based multiple meteorological measurements. The proposed ANN models differently use past trends in frequently recorded data of wind speed, pressure, and temperature change to predict the future values of wind power. Figure 1 illustrates the classification of wind power forecast techniques based on time series.

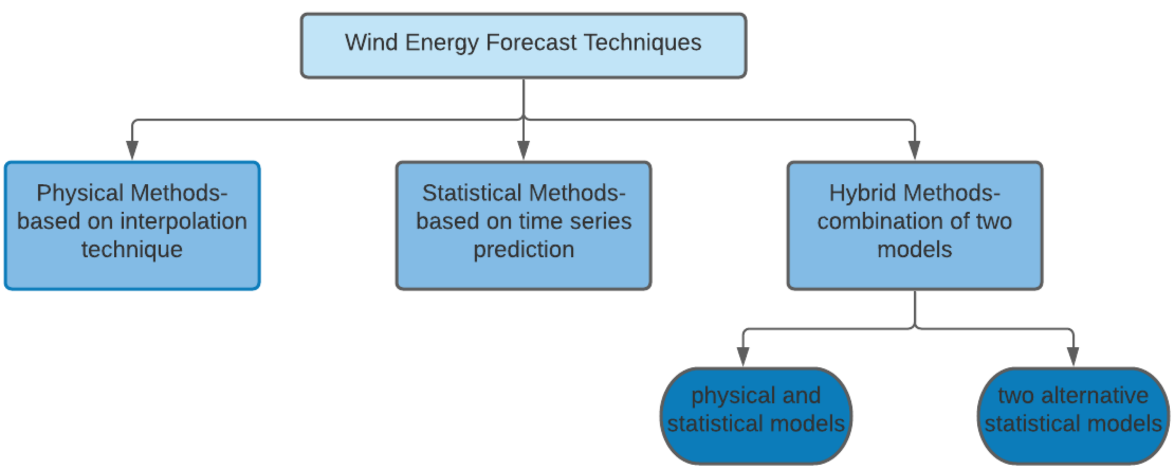

Figure 1: Classification of wind power forecast techniques [41]

Maximilian Du proposed an improved approach based on long-short term memory (LSTM) for better short-term wind predictions [13]. The LSTM is one of the updated version of Recurrent Neural Network (RNN) which carries temporal information by using a continuous cell state. The model is modified using an adaptive compression algorithm, ReLu units, and cell regulation for getting an effective increase in accuracy and decrease in naïve characters. Research from

Warsaw University of Technology, Koszykowa proposed an ensemble method for a two-day ahead forecast of a small wind turbine [14]. After performing statistical analysis on time-series energy generated data, input data is used to verify the data impact on forecast quality. Different combinations of predictors have been used for the development of ensemble methods based on ANN type multi-layer perceptron (MLP). Hyperparameters tuning in several complex models provides an optimal choice for future predictions of wind power. The results drawn from the research demonstrate the merits of using the ensembled model which is based on a hybrid method. The research concludes that LSTM based deep learning model is the most accurate one then comes MLP and then a simple machine learning model.

The hybrid models represent a nonlinear complex relationship in terms of classification or forecasting and help in extraction of dependence factor between different input variables in the training process. Jordan Nielson provided an approach based on ANNs in which atmospheric inputs were used for improvement of wind turbine power output prediction [10]. ANN used in this paper is fed forward back propagation (FFBP). The role of metrics is evaluated in the study that relate to atmospheric boundary-layer including turbulence intensity, wind density, wind speed, wind shear, Richardson as input parameters for ANN. Hyperparameter adjustments provided the difference in accuracy and performance. Stefan Balluff, Jorg Bendfeld worked on wind speed and energy prediction of an offshore wind farm [11]. They used a recurrent neural network in their estimation model. Neural networks use input plus, bias values, and weights to determine output using a simplified system of number processing. The number of inputs, hidden layers, and outputs varies from a few to many numbers depending upon the estimation model requirement, but the important thing to consider in RNN is the link between forward time-based information and hidden layers to calculate the output from one layer to another. Wei Sun, Yuwei Wang provided a hybrid model to forecast wind speed [12]. They used empirical model decomposition which deals with non-linear and non-stationary signals to reduce the random fluctuations. A backpropagation neural network is used with two hidden layers to enhance the accuracy and strengthen the sensitivity of wind speed prediction.

A research methodology designs deep learning model based on hybrid approach using both LSTM and CNN networks [14]. Time slide Window (TSW) is proposed to use in for mining the data features of provided active power data. The configured model uses regression criteria on real-time dataset of wind farm. Ying-Yi Hong proposed a forecast for 24h ahead wind power output utilizing the hybrid approach of CNN with Radial Basis Function Neural Network (RBFNN) [15]. Double Gaussian Function (DGF) is used as an activation function in this model. Convolution, pooling, and kernel operations are used in the CNN architecture for extracting characteristics of wind power while dealing with uncertain characteristics. For performing backend operations, Keras and TensorFlow are configured. The results observed by R-square metric revealing the outperformance of proposed approach for the prediction of 24h wind output power.

Table 1 and Table 2 illustrate the brief summaries of accuracies, results, and observations on previously researched wind power predictions.

Table 1: Summary of accuracies on previously researched wind power predictions

Author	Year/Journal	Algorithm applied/ Technique used	Activation functions	Error/Accuracy (%)
Lorenzo Donadio [15]	2021	Integration of NWP and ANNs	tanh and linear function	NMSE- 8.76 RMSE- 13.03
Elcin Tan [16]	2021	(ANN) approach for prediction of both wind speed from WRF model and wind power	logarithmic, sigmoid	nRMSE- 13
Cem Özen [17]	2021	LSTM- for short term multidimensional predictions	Tanh	MSE- 3.3094 MAE- 1.4451
Navid Shirzadi [18]	2021	ML techniques including SVM and RF and DL techniques including RNN, LSTM and NARX	Tanh	MAPE- 4
Rui Yin [19]	2020	LSTM- based on climate model	tanh, sigmoid	MAE- 10.91
Zi 'Lin [20]	2020	Isolate Forest (IF) and neural networks based integrated approach	ReLu	MSE- 0.003
Khaled Khelil [21]	2020	Integration of ANN with discrete wavelet transform	tansig, logsig	RMSE- 13 MAPE- 19
Hamed H.H. Aly [22]	2020	RKF, FS, WNN and ANN	sin(x), cos(x)	nRMSE- 3.6153 MAPE- 0.04446
MOHAMMED BOU-RABEE [23]	2020	ANN merged with PSO	linear and logsig	MAPE- 3
Jaseena K U [24]	2019	RNN-LSTM	Sigmoid tanh	RMSE- 1.653 MAE- 1.332 MAPE- 16.419
Maximilian Du [25]	2019	LSTM	sigmoid, tanh, ReLu	NMAE- 0.438
Ying-Yi Hong [26]	2019	CNN- RBFNN	gaussian function	NMSE- 2.6144 MAPE- 9.2625
Yongsheng Wang [27]	2019	Deep learning of grouped time series (LW-CLSTM)	tanh, sigmoid	MAE- 1.36

Wei Sun [28]	2019	PSO- least squares SVM	radial-basis function	MAE- 5.91 RMSE- 7.29
Jordan Nielson [29]	2019	ANN	sigmoid, tanh, ReLu	MAE- 4

Summary of features and observations on previously researched wind power prediction models

Table 2: Summary of features and observations on previously research wind power prediction models

Year/ Journal	Author	Deep learning Model Summary	Observation & finding
2021 [15]	Lorenzo Donadio	NWP- ANN	Provides the choice of parameters, features, target variables
2021 [18]	Navid Shirzadi	LSTM- NARX	Predicts the tendency for both off-peak and peak values, with min. error and max. accuracy.
2020 [20]	Zi Lin	NN with Pearson product-moment correlation	Blade pitch angles are signified among all the features on the predictive model
2020 [27]	Yongsheng Wang	CNN-LSTM	LSTM effective in extracting time-series characteristics and training time is reduced in CNN
2019 [26]	Ying-Yi Hong	CNN cascaded with RBFNN	Volatile wind power features in time-series are extracted in CNN
2019 [29]	Nielson, Jordan	FFBP- ANN	Lowest MAE with input features of turbulence intensity, Richardson Number, density and wind shear
2019 [40]	A. Di Piazza	ANN-NARX	Short term time horizon predictions are well evaluated by NARX
2018 [36]	Aamer Bilal Asghar	ANFIS)	Less computational cost for estimation of fast estimation of optimized rotor speed and wind speed.
2015 [37]	T.R. Ayodele	Feed Forward ANN-BP	Wind speed prediction can be made effective using temperature, wind direction and time of the day as inputs.

2014 [38]	M. Jabbari Ghadi	ICA-ANN	NWP models enhance the accuracy of wind power prediction
2013 [39]	Nawal Cheggaga	MLP-Levenberg-Marquardt learning algorithm	Significant improvement of robust forecast and accuracy.

Now a days, multivariate forecasting of wind power is an active area of research and investigation of dependent variables for output power is still challenging as wind speed is mainly focused leading to potential uncertainties for predictions. The research work utilizes specifically pitch angles of three blades in prediction of wind power using. Accurate predictions from the short-term wind forecast have expandable applications to meet the demands of the wind farm. The novel contribution of this paper is utilization of influential factors specifically including both pitch angle of each blade and nacelle orientation for wind power output prediction. Compared with baseline evaluation, multivariate deep learning configurations including Bi-LSTM, GRU and Hybrid CNN-LSTM are designed using most reliable input features to get the proposed improved accuracy in power predictions with minimum computational cost. The proposed models in this paper have fast computational speed and non-linear mapping ability for perfect prediction in the short term.

The remainder of this research paper is briefly described below; **Section 2** covers the features of the wind farm SCADA dataset utilized in the study, obtained from three-year high frequency monitoring wind turbine. **Section 3** focuses on basic methodology for constructing deep learning models. Starting with the data pre-processing and features extraction, deep learning configuration is discussed, following the cell structures of LSTM, GRU and CNN. A detailed comparative analysis of three models is performed in **Section 4. Section 5** discusses the observations and conclusions drawn from the above trained models of SCADA measurements.

2. Turbine characteristics of SCADA dataset

The investigated data is metrological data recorded from Sapphire Wind Power Company Limited (SWPCL). The data provides with the major properties of three bladed Onshore upwind turbine with 1600kW rated power.

The specification and dimensions of GE-Energy wind turbine can be found in the table as followed [35]

Table 3: Major properties of wind turbine SWPCL

Wind class	IEC Iib (TCIII+)
Model	1.6xle (33 wind turbines)
Rotor diameter	82.5 m
Hub height (maximum)	100
Hub height (minimum)	80
Swept area	5,346 m^2
Rotor speed (maximum)	18 radian/min
Rotor speed (minimum)	9 radian/min
Wind speed	3-25 m/s
Rated wind speed	11 m/s
Capacity of wind farm	49.5 - 52.8 MW
Power control	Electric drive pitch control
Generator	Double fed Asyn

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

To evaluate the accuracy of proposed method predictions, practical metrological readings of wind turbine have been used for conducting the case study. The wind farm has installed capacity of 50MW. The historical data is collected from July. 1st, 2018 to July. 1st, 2021. Multiple time sequence data is included in this dataset, i.e. blade positions of three blades, nacelle orientation, temperature, wind speed and power output etc.

3 Methodology

3.1 Exploratory Data Analysis

To analyze and investigate input data for summarizing the main characteristics, exploratory data analysis (EDA) is used. It normally employs data visualization methods. EDA reveals the better relationship between dataset variables to ensure the validation of results. Necessary python libraries are imported including pandas, numpy and matlplotlib. After interpreting the total number of rows, columns and datatypes, statistical summary is used to understand the numerical fields. Histograms of all the numeric fields were plotted to find the shape, center, and spread of data observations. The histogram of power output with respect to frequency distribution in demonstrated in Figure 2.

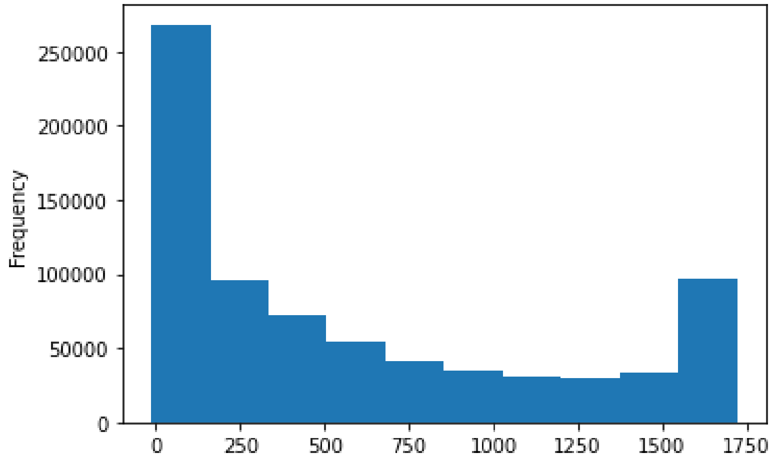

Figure 1: Histogram of Output Power by SCADA measurements

Research activity for the proposed approach is presented in the flow chart. Seven essential steps followed are demonstrated in Figure 3 below.

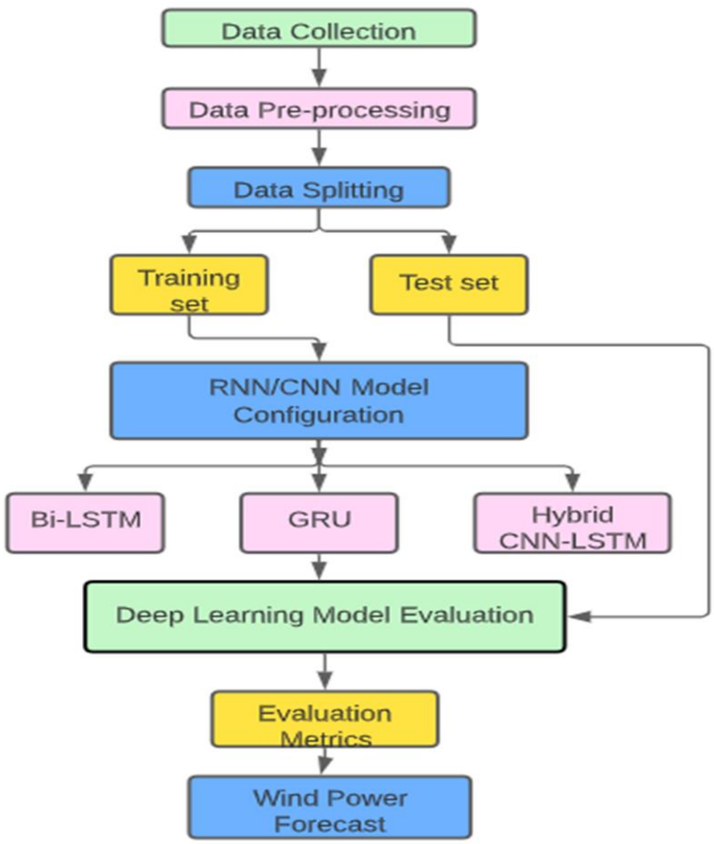

Figure 2: Research Activity Flow-chart

3.2 Data Preprocessing

Data pre-processing is used for identification and correction of inaccurate records from a dataset along with recognizing irrelevant or unreliable dataset variables. Missing data is handled using linear interpolation and null values are filled with the median values. Before final processing the dataset into model for training, categorical variables are converted into numeric values. To figure out if the give time series is stationarity, ADF Test is used [34]. Outliers were detected from normal findings and removed from the dataset, representing error detections, and measuring variability.

The histogram for each feature is plotted in the figure illustrated below. The rated speed of wind is 11 m/s while the median and mean of the dataset recorded is 6.46 m/s and 6.59 m/s. It can be illustrated that, at most of the operating time values of generated powers are nearly to the rated values of power (1600 kW). The histogram in Figure 4 for nacelle position inclined with the bimodal distribution illustrating the roughly classification of wind directions in the two main clusters. Also the temperature ambient followed the normal distribution.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

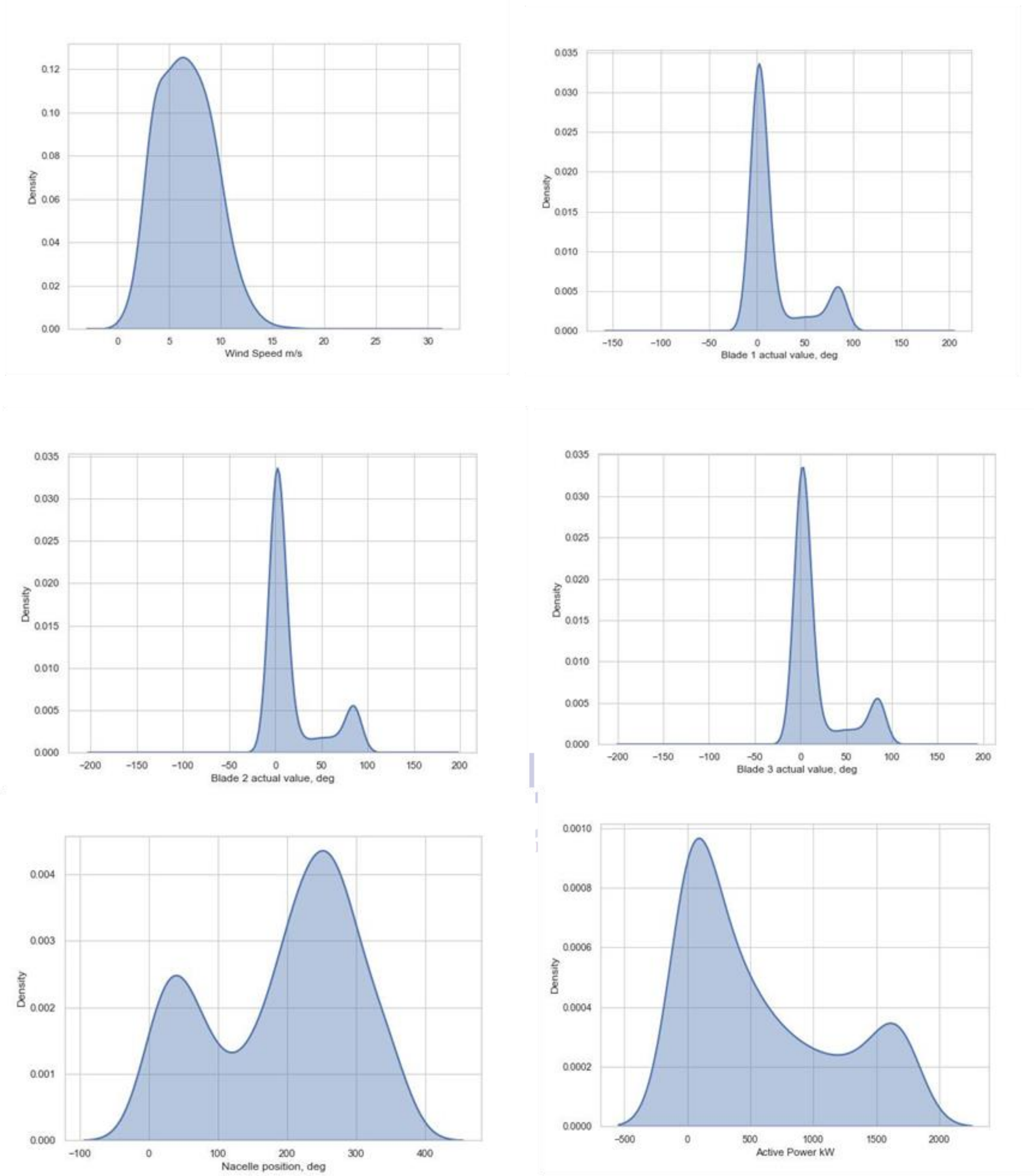

Figure 3: Histograms of selected input and features for obvious outlier detections

3.3 Correlations and Heatmap

The parameters having greater influences on power generation from wind are essential to discover from SCADA dataset. Correlation is a critical underlying factor for multivariate forecasting. It shows how variables of a dataset are related to each other and how they move concerning each other. The value of correlation moves from -1 to +1. A positive value of correlation indicates that the variables are moving in the same direction while negative correlation indicates the opposite relationship. The heatmap in the figure below illustrates the matrix values graphically with different shades of color for different correlation values. The correlation coefficients are calculated using Pearson product-moment as displayed in Figure 5 of heatmap. Linear relationship follows the most common

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

measurement of the strengths using heatmap between any two variables. By dividing the covariance of the targeted variable by the product of standard deviations, this correlation can be find out.

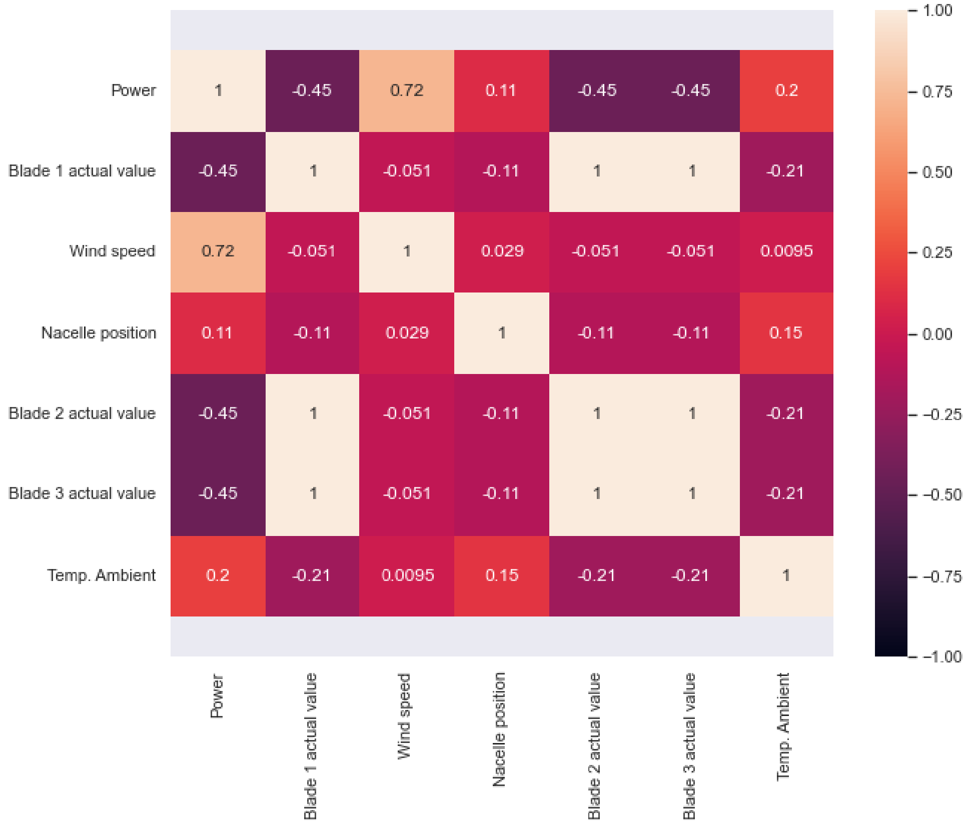

Figure 4: Input features against output power in heatmap correlation

The numerical values of variables are much different in scale when the time sequence composed of different variables is used to forecast. Hence, the value of output and input variables are normalized to reduce the effects of different scale values. Using Min-Max, input features were scaled down into a range of 0 and 1 scaler before flowing them deep learning layer. The correlation used for scaling can be expressed in Equation (1):

$$X_{Scaled} = \frac{x_p - x_{min}}{max(x) - min\,(x)} \tag{1}$$

where x_p is the first value of input and x_{min} is the minimum value.

3.4 Baseline Evaluation

To measure the performance of model, training a baseline model is a good idea as it is used as a reference. Evaluation of baseline model is repeatable and established as a persistent model. The working principle of baseline model is simple, as it forecasts the next timestamp value as the previous one. The model works on a simple principle, next time step is forecasted as the previous one, in other saying t+1 value is the same as t. In multi-step forecasting, forecast values of t+1, t+2, t+3 can be adapted same as t, absolute forecast horizon will remain the same. The Figure 6 below illustrates the difference between labels and prediction to judging hyper-parameter optimization in algorithm.

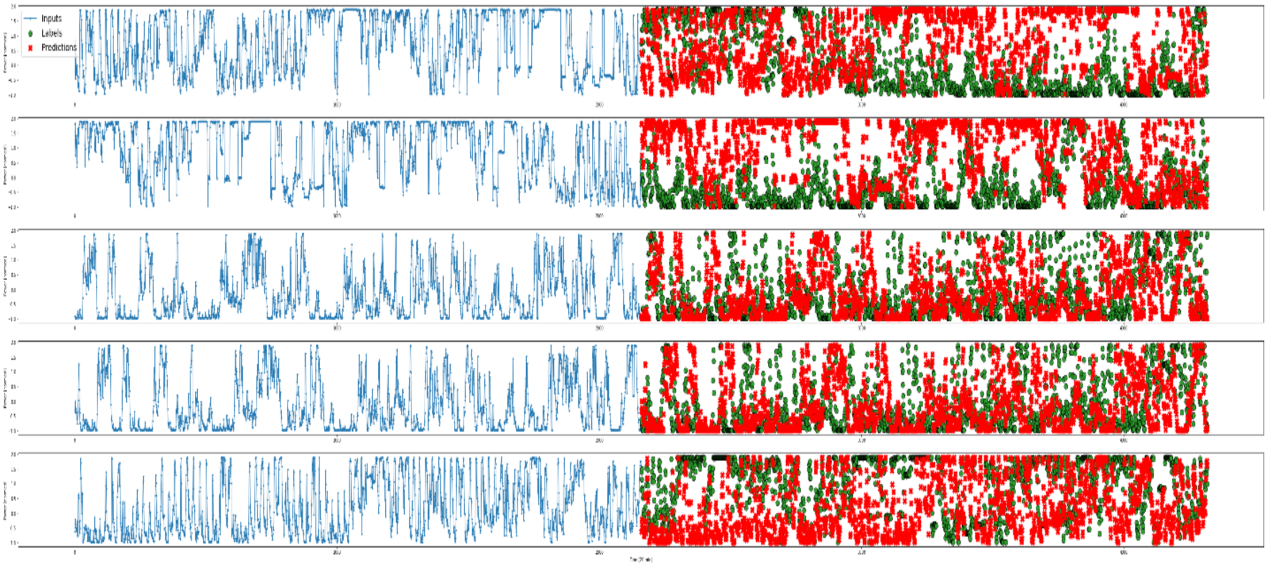

Figure 6: Multistep baseline evaluation model

3.5 Deep-learning configuration

In this paper, a platform named as TensorFlow supported and developed by Google, is used for creating deep learning setup. Major programing language used for employment of this configuration is Python3. Large-scaled datasets with specific attributes can be handled smoothly by TensorFlow including multidimensional arrays. Tensors are also called as multi-dimensional arrays. The accuracy in predictions of wind power is increased by graphical flow of tensors from one layer to the other in structure of neural network. Several deep learning structures were tested, and evaluations were carried out with different number of layers, activations, and neuron numbers in every layer. The designed neural network comprised of the five input features including nacelle orientation according to wind direction, wind speed, angles of three blades and temperature ambient whereas the output variable is wind turbine active power. The cost function of MSE and RMSE were two metrics used for measurement of outputting tensors accuracy and calculating the mean value of tensor elements along several capacities of the tensor after execution of training phase.

3.6 Cell structures of LSTM, GRU and Hybrid model

For handling problems of sequential time-series, RNN is purposely implemented with the advantage of looking backs on historical information while CNN is utilized for feature engineering and creating informative representations in time-series. Although, the problems of gradient explosion and gradient disappearance limits its application in case of long-term dependencies as parameter solving algorithm is backpropagated in time and single in structure [31].

3.6.1 LSTM architecture

Based on RNN, concept of three gates named forget gate, input gate and output gate with memory cell, are introduced into LSTM, for controlling information flow across LSTM cells and determining the input, retention, and output of information, in particular [32].

The Figure 7 below represents the simple structure of LSTM block.

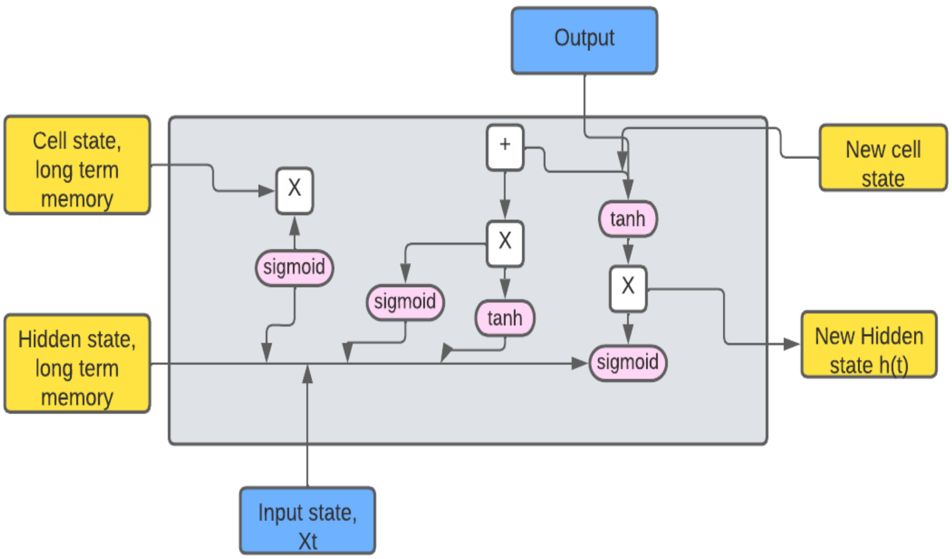

Figure 7: Internal structure of LSTM cell [42]

Input gate tells about new information to be stored in cell state. Forget gate identifies the information to be throw away and output gate provides activation for final input of block. Equations 2, 3 and 4 below are used for input, forget and output gate of LSTM cell.

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i \quad (2)$$

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f \quad (3)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o \quad (4)$$

3.6.2 GRU architecture

Similarly, special gates including update and reset gates are utilized in GRU, for reducing gradient dispersion, and enable the competency of long-term memory with less computation loss. Input gate and forget gate of LSTM are replaced by the update gate in GRU, to find the degree of retention from the previous information in the current forecasting [33]. If the time-series is decomposed into different non-stationary components in pre-processing, it gives more accurate predictions as compared to original data predictions. The Figure 8 below represents the simple structure of GRU block.

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

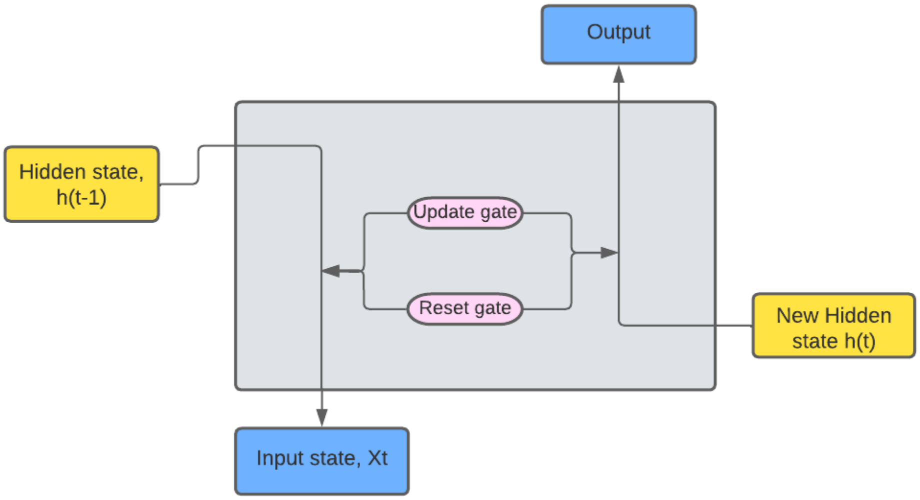

Figure 8: Internal structure of GRU cell [42]

The update gate tells the model how much of the past information needs for passing along to the future from previous time span. The information is filtered using update gate and risk of vanishing gradient is reduced. Reset model allows the model how much information needs to be forget [30]. Equation 5 and Equation 6 are used for update and reset gates of GRU. Basically, the model uses this gate to decide how much of the past information needs to be forget.

$$z_t = \sigma(w_z x_t + h_{t-1} U_z + b_z) \quad (5)$$

$$r_t = \sigma(w_r x_t + h_{t-1} U_r + b_r) \quad (6)$$

3.6.3 Hybrid CNN-LSTM architecture

The CNN-LSTM hybrid architecture uses CNN and LSTM both cell structures. CNN layers are used for feature extraction of spatial input data, then it is combined with LSTM for supporting the sequence of the predictions. LSTM cell works as a temporal extractor for input features and time-lag features. The Figure 9 below demonstrates the architecture of hybrid CNN-LSTM model.

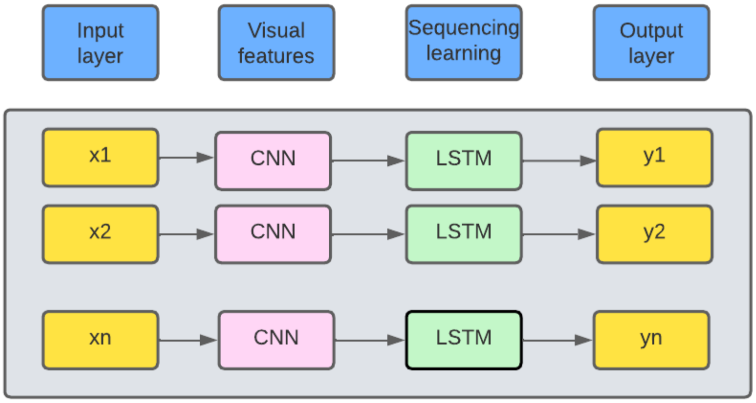

Figure 9: Hybrid CNN-LSTM Structure [43]

The architecture flow of all three deep learning configurations is given below in Table 4.

Table 4: Architecture flow of Bi-LSTM, GRU and Hybrid CNN-LSTM model

Architecture	**Bi-LSTM**	**GRU**	**Hybrid CNN-LSTM**
Model	Sequential	Sequential	Sequential
Input layer (type)	Bidirectional LSTM	GRU	Conv(1D)
Hidden layers (type)	Bidirectional LSTM	Dropout	LSTM
	Dense	GRU	LSTM
		Dropout	Dense
		Dense	Dense
			Dense
			Lambda
			LSTM
			Dense

4 Results and discussions

Number of factors can be used for comparison of different wind power estimation methods including input features, accuracy measurement and computational time. The two important factors are accuracy and computational time which are considered normally. Several metrics are used for evaluating the performance of wind power forecast. Measured wind power deviation from the actual power can be assessed easily using different statistical approaches. The reflection of past studies shows the selective evaluation criteria including root mean square, mean square error, absolute percentage error and normalized error. Deep learning models are evaluated based on two metrics including mean square error (MSE), mean absolute error (MAE). Following Equation 7 and Equation 8 are the defined equations for finding the value of metrics MAE and MSE respectively.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |x_i - \hat{x}_i| \tag{7}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \tag{8}$$

While the training database was kept as it is, the deep learning configuration was altered at each model trial to investigate the difference in accuracies with several lookback features, number of epochs and batch size. Three predictive models are examined through identical testing and validation datasets. The learning rate is set as 0.2 in these models of neural networks.

4.1 Performance Assessment and Comparative Analysis

The paper has performed an analysis on lookback period used with deep learning models utilized in the research, and comparison is being done in prediction of wind power performance. A lookback period (also called as lag) is defined as how many previous timestamps are used in prediction of subsequent timestamp [45]. For instance, if lookback period is set to 5, this means timestamps at *t-4*, *t-3*, *t-2*, *t-1*, and *t* are utilized to predict the value at *t+1*. The selection of lag period is an arbitrary process, for finding the best predictive performance on the validation set. However, another method is utilization of ACF and PACF and find significant non-zero autocorrelation as illustrated in the Figure 10.

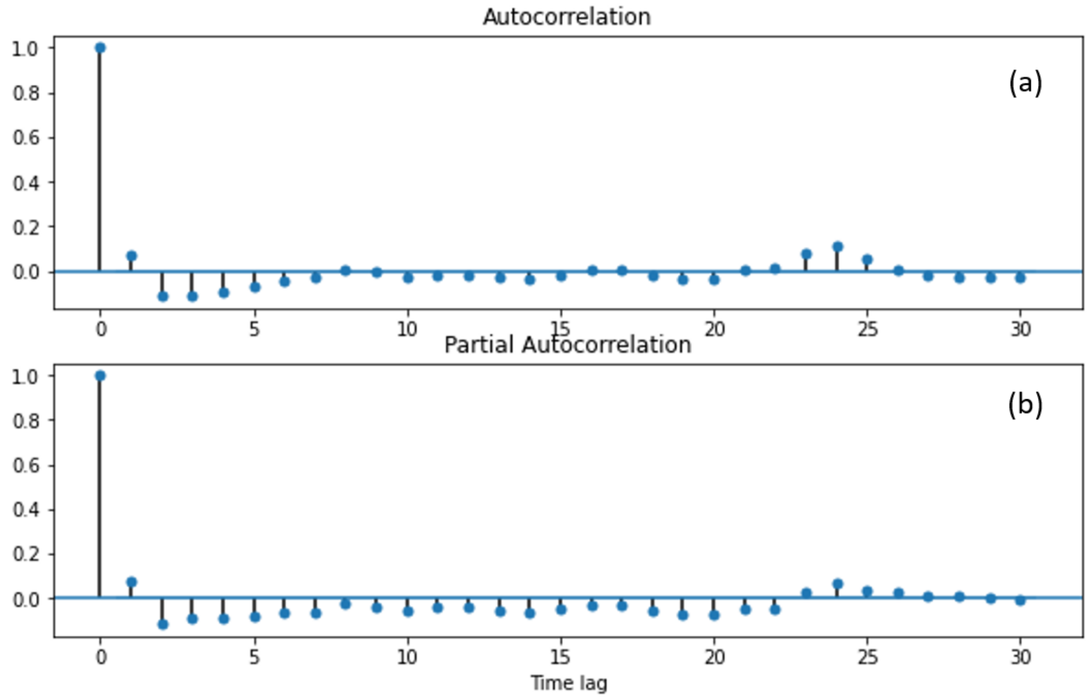

Figure 10: (a) ACF and (b) PACF Graphs for coefficients

For detection of dataset patterns and checking randomness, autocorrelation analysis is used. ACF and PACF plots help in determining the parameters of Moving Average and Auto-regressive models use ACF and PACF plots to find the values of parameters. ACF provides us the values of autocorrelation of any series with its lagged values. But instead of finding correlations of present with lags like ACF, partial auto-correlation function finds correlation of the residuals (which remains after removing the effects which are previously explained by the earlier lag(s)) with the next lag value hence it is 'partial' as we remove already found variations before we find the next correlation [44].

As Figure 10 illustrates the several correlations, that are non-zero hence the time-series is not random. High degree auto correlation has been observed between adjacent lag = 1 as PACF plot shows. Therefore lookback period =1 is utilized first to model the process.

4.1.2 Bi-LSTM Configuration

For Bi-LSTM Model, comparison of prediction with different lookback periods and batch sized in training and validation gives minimum mean square error of 0.0056 and 0.0129. The minimum values are observed with minimum batch size and lookback period equal to 3.

The corresponding training and validation loss plots of Bi-LSTM model are illustrated in Figure 11(a), 11(b) and 11(c) with several lookback periods for interpreting the stability of the model configured.

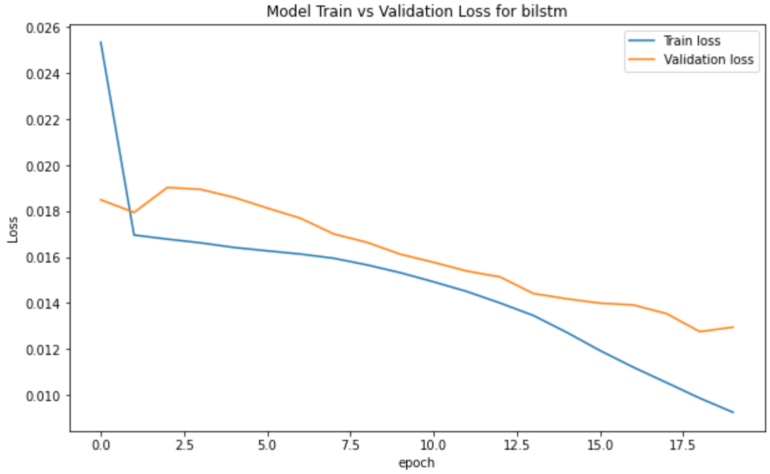

Figure 11(a): Training and Validation loss plot of Bi-LSTM model with lookback period = 1

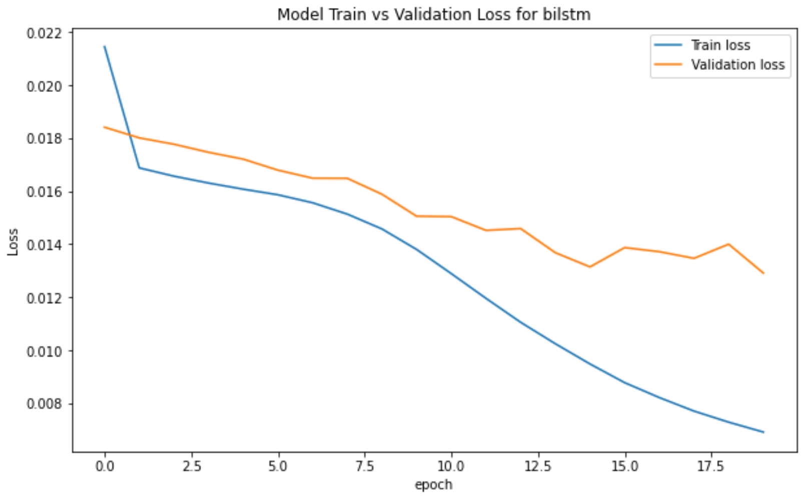

Figure 11(b): Training and Validation loss plot of Bi-LSTM model with lookback period = 3

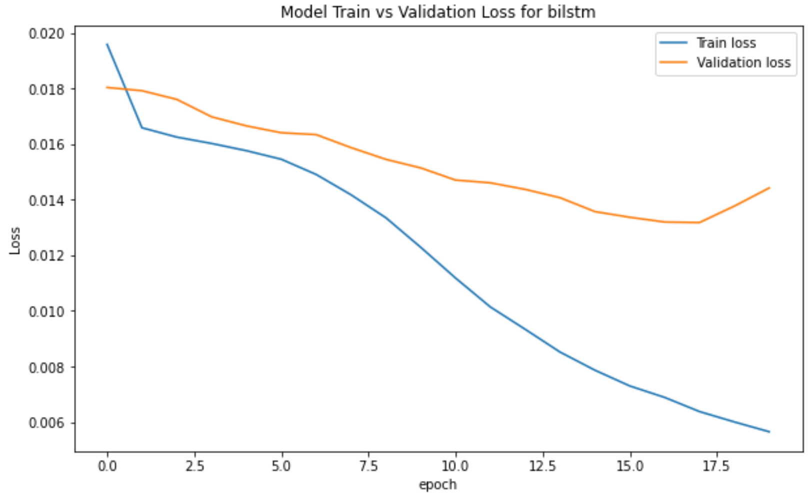

Figure 11(c): Training and Validation loss plot of Bi-LSTM model with lookback period = 5

4.1.2 GRU Configuration

For GRU model, the highest accuracy of training comes up with MSE of 0.0071 whereas highest accuracy in validation comes up with MSE of 0.016 with lookback period equal to 1. GRU model is more stable in minimum lookback period as it can work well with minimum history of data points, taking less computational time as compared to Bi-LSTM and Hybrid CNN-LSTM model. The Figures 12(a), 12(b) and 12(c) illustrate the training and validation loss with modifications in lookbacks period.

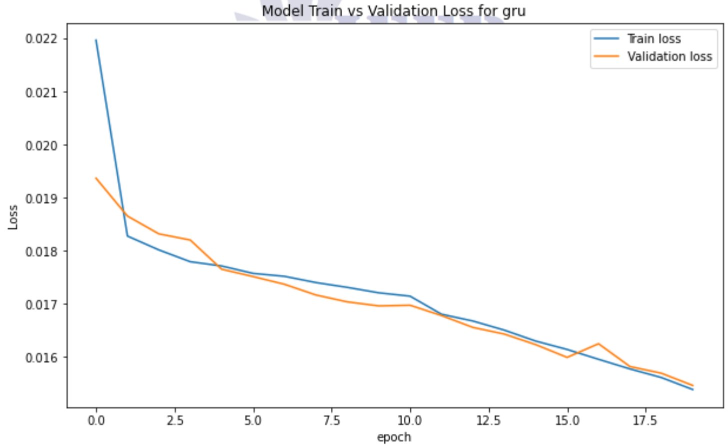

Figure 12(a): Training and Validation loss plot of GRU model with lookback period = 1

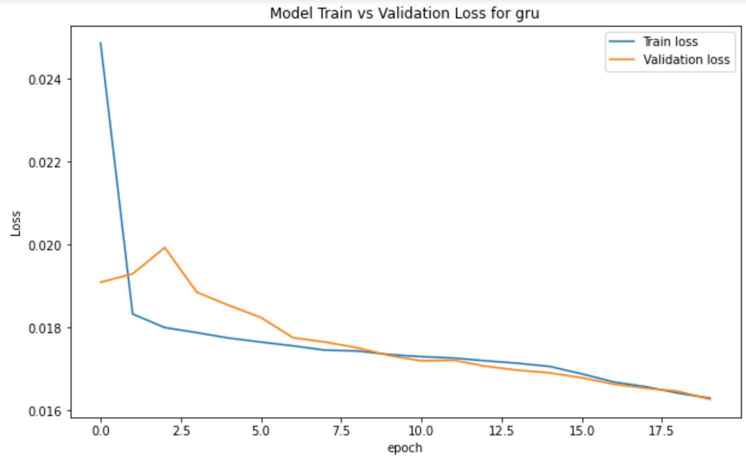

Figure 12(b): Training and Validation loss plot of GRU model with lookback period = 3

Figure 12(c): Training and Validation loss plot of GRU model with lookback period = 5

4.1.3 Hybrid (CNN-LSTM) Configuration

Hybrid Model gives more accuracy with more computational cost, and lookback period equal to 5, with MSE of 0.0087 in training loop and MSE of 0.0129 in validation loop. Hybrid CNN-LSTM model gives stable results with more history of data points as compared to Bi-LSTM model and GRU model. Figures 13(a), 13(b) and 13(c) demonstrate the difference in training and validation loss of Hybrid model with respect to lookback period.

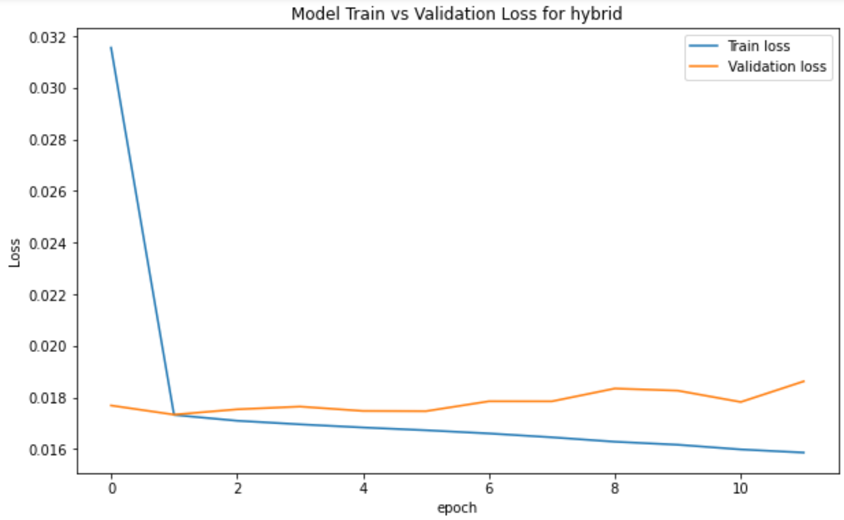

Figure 13(a): Training and Validation loss plot of Hybrid CNN-LSTM model with lookback period = 1

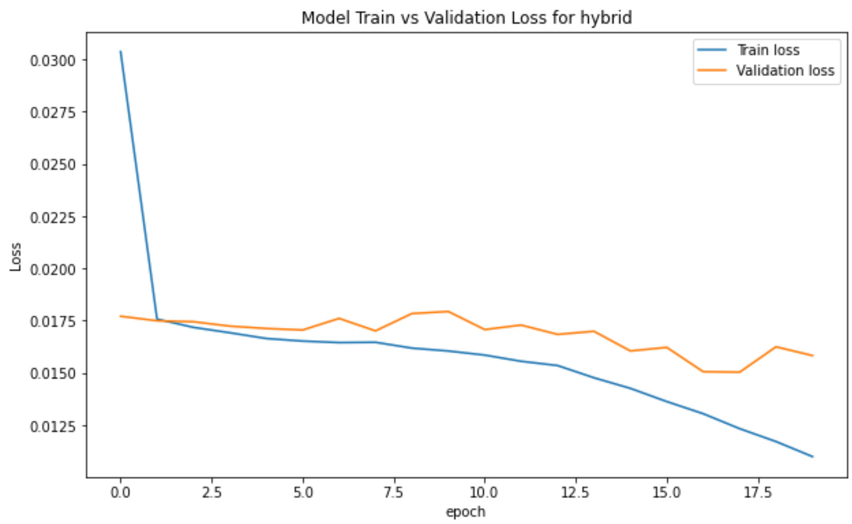

Figure 13(b): Training and Validation loss plot of Hybrid CNN-LSTM model with lookback period = 3

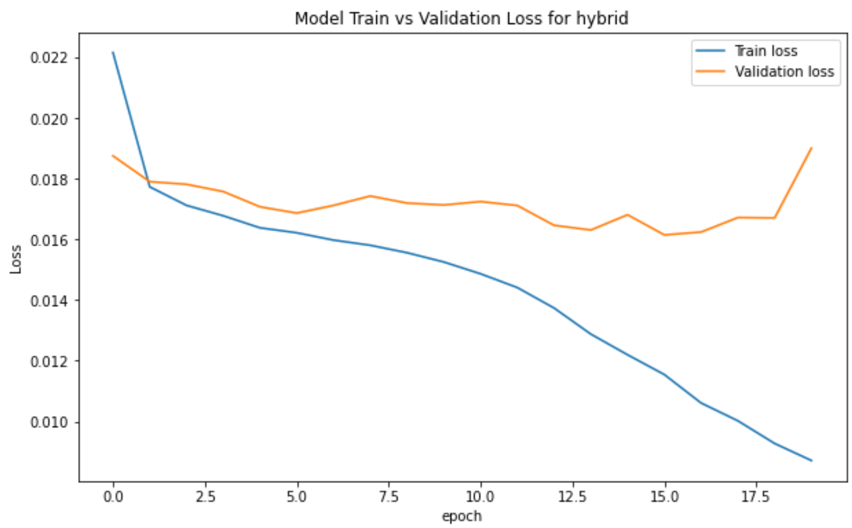

Figure 13(c): Training and Validation loss plot of Hybrid CNN-LSTM model with lookback period = 5

4.2 Accuracy of forecasting

The parameters of each model are tested using several runs to optimize the accuracy and performance for the individual algorithms based on different lookback periods, number of epochs and batch size for training the model and computational cost evaluation as well. For assessment of accuracy for all three algorithms, the output sensitivity is observed, and the suitable parameters are applied to the models.

Table 5(a), 5(b) and 5(c) below display the training and validation loss of three deep learning configuration with three different batch-sizes and for each batch size, lookback period is changed for finding suitable wind prediction performance.

Table 5(a): Training and Validation loss with different batch size and lookback period = 1

Deep-learning model	Batch-size	Training (MSE)	Validation (MSE)	Computational time
Bi-LSTM	32	0.0058	0.0132	284
GRU	32	0.0145	0.0155	145
Hybrid (CNN-LSTM)	32	0.011	0.0189	245
Bi-LSTM	64	**0.0071**	**0.0124**	**186**
GRU	64	0.0153	0.0154	75
Hybrid (CNN-LSTM)	64	0.0132	0.0191	140
Bi-LSTM	128	0.0092	0.013	170
GRU	128	0.0163	0.0163	90
Hybrid (CNN-LSTM)	128	0.0148	0.0197	150

Table 5(b): Training and Validation loss with different batch size and lookback period = 3

Deep-learning model	Batch-size	Training (MSE)	Validation (MSE)	Computational time
Bi-LSTM	32	**0.0056**	**0.0129**	**280**
GRU	32	0.0148	0.0159	140
Hybrid (CNN-LSTM)	32	0.0142	0.0162	240
Bi-LSTM	64	0.0071	0.0134	186
GRU	64	0.0071	0.0129	76
Hybrid (CNN-LSTM)	64	0.0143	0.0171	135
Bi-LSTM	128	0.0086	0.0129	200
GRU	128	0.0165	0.0165	100
Hybrid (CNN-LSTM)	128	0.0143	0.0171	165

Table 5(c): Training and Validation loss with different batch size and lookback period = 5

Deep-learning model	Batch-size	Training (MSE)	Validation (MSE)	Computational time
Bi-LSTM	32	0.0056	0.0144	265
GRU	32	0.0148	0.0152	140
Hybrid (CNN-LSTM)	32	**0.0087**	**0.0124**	**235**
Bi-LSTM	64	0.0069	0.0129	219
GRU	64	0.0154	0.0155	77
Hybrid (CNN-LSTM)	64	0.011	0.0158	150
Bi-LSTM	128	0.0089	0.0135	195
GRU	128	0.0163	0.0163	100
Hybrid (CNN-LSTM)	128	0.0128	0.0201	160

All models are validated in different scenarios; final MAE of Bi-LSTM, GRU and Hybrid CNN-LSTM model is illustrated in Figure 14. Bi-LSTM model gives more accurate results when compared with SCADA measurements in short-term prediction from two to three hours. Figure 15, Figure 16, and Figure 17 give the comparison of each model's performance with true values.

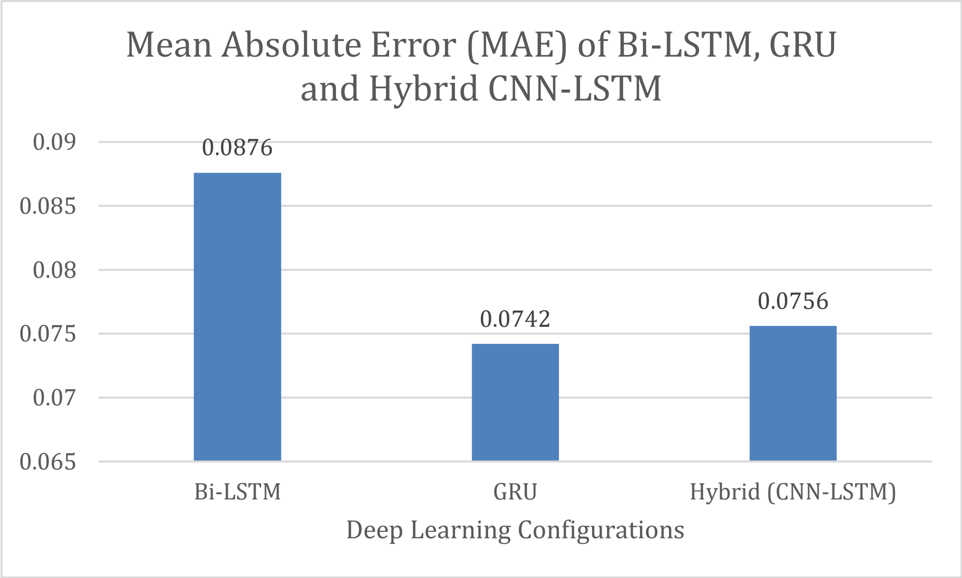

Figure 14: Variations of final MSEs in different deep learning scenarios

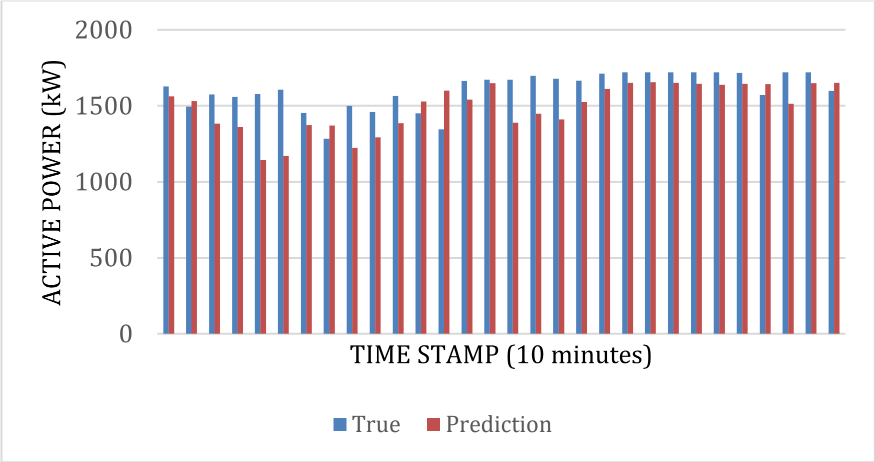

Figure 15: Comparison of active power b/w actual SCADA measurements and Bi-LSTM model predictions

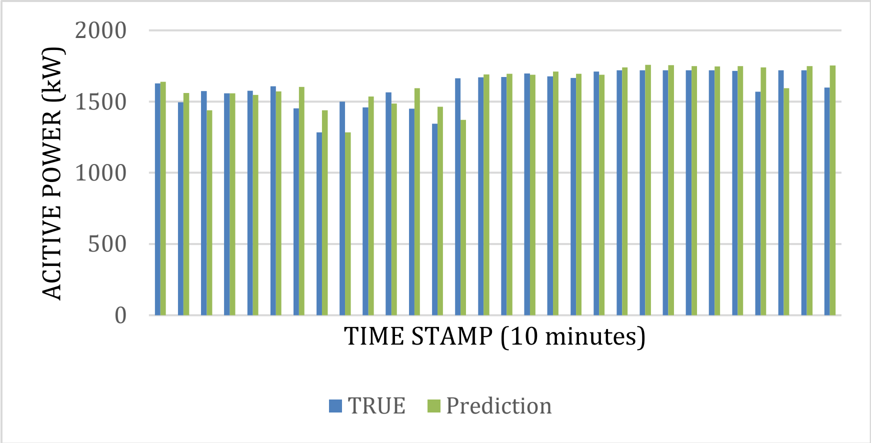

Figure 16: Comparison of active power b/w actual SCADA measurements and GRU model predictions

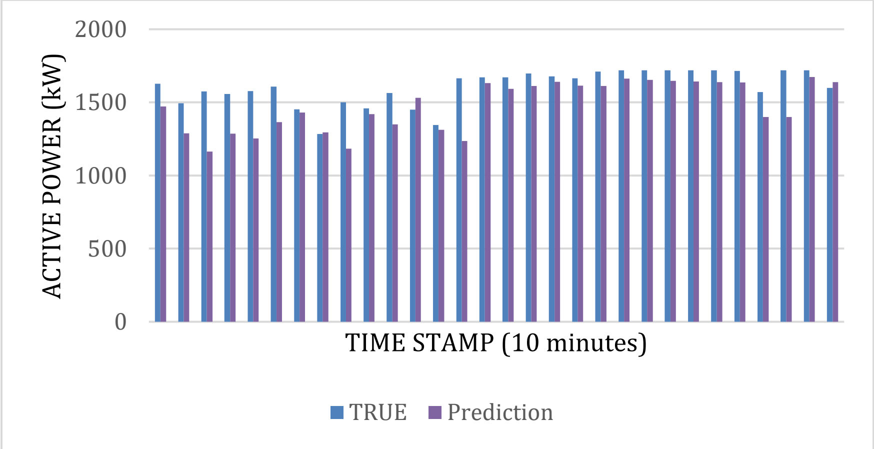

Figure 17: Comparison of active power b/w actual SCADA measurements and Hybrid (CNN-LSTM) model predictions

If we look at Figure 18 showing the predictions of three deep-learning configurations with true values, it is observed that:

GRU gives minimum difference with true value at almost every point and minimum historical data is sufficient for GRU for learning the pattern.

Bi-LSTM takes more lag and computational time to prove its accuracy in prediction when compared with GRU structure.

Hybrid (CNN-LSTM) gives is more suitable for feature extraction but takes more computational time for sequential dataset and needs more historic data as compared to GRU model and Bi-LSTM. As it needs maximum lookback period for proving its accuracy.

-

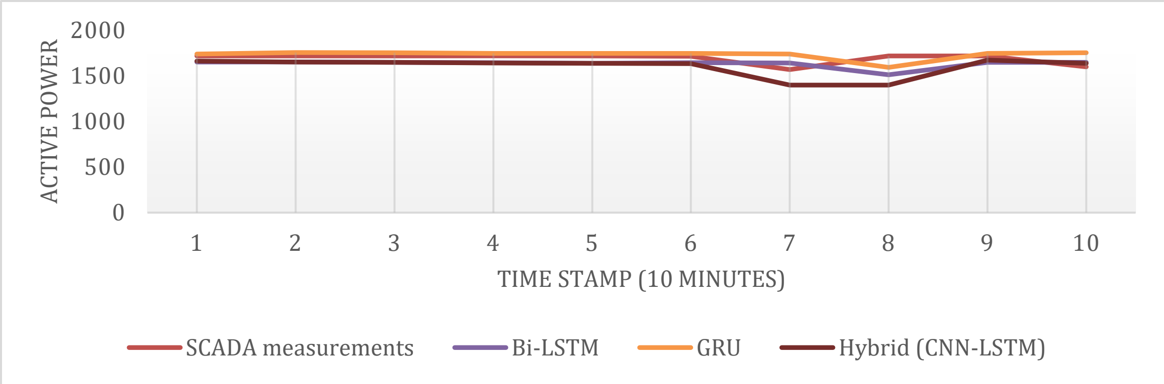

Figure 18: Comparison of SCADA measurements with Bi-LSTM, GRU and Hybrid (CNN-LSTM) Models

5 Conclusion

The paper presents neural networks variants including RNN (LSTM & GRU) and Hybrid CNN-LSTM to predict the active power of wind turbine based on wind speed, wind density, ambient temperature, nacelle orientation and pitch angles of blade I, II and III. The models are trained with technical approach adjusting number of parameters including number of epochs, learning rate, batch size and number of lags. GRU gives minimum MAPE value of 0.074 having slight difference with MAPE values of Hybrid (CNN-LSTM)'s MAPE of 0.0751 and Bi-LSTM's MAPE of 0.078. Bi-LSTM and GRU networks can filter out redundant information in modeling non-linear complex relationships by taking less time while CNN when cascaded with LSTM also gives improved results than Bi-LSTM because feature extraction of CNN is deep, containing more layers to discriminate the characteristics at the cost of learning cost. Bi-LSTM and Hybrid model predict runoff with almost identical accuracy. The GRU model has less parameters with simpler structure, smallest lag and requires less time for model training. It can be considered for preference in for short term runoff prediction. The experiments performed showed the suitable MAE values of validation and MAPE values of predictions for GRU model in comparison to Hybrid CNN-LSTM and Bi-LSTM with minor accuracy difference and less computational cost. The outcomes of the research work may lead to the conclusion that potential of these deep learning models can be fully utilized to extract regularities and dependencies of the input data as reliably as possibly. Future work for development could be performed by restructuring the selected model and analysis of internal pattern for deep information.

REFERENCES

Bou-Rabee, M., Lodi, K. A., Ali, M., Ansari, M. F., Tariq, M., & Sulaiman, S. A. (2020). One-month-ahead wind speed forecasting using hybrid AI model for coastal locations. IEEE Access, 8, 198482–198493. https://doi.org/10.1109/ACCESS.2020.3028259

Aly, H. H. H. (2020). A novel deep learning intelligent clustered hybrid models for wind speed and power forecasting. Energy, 213. https://doi.org/10.1016/j.energy.2020.118773

Asghar, A. B., & Liu, X. (2018). Adaptive neuro-fuzzy algorithm to estimate effective wind speed and optimal rotor speed for variable-speed wind turbine. Neurocomputing, 272, 495–504. https://doi.org/10.1016/j.neucom.2017.07.022

Lin, Z., & Liu, X. (2020). Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network. Energy, 201. https://doi.org/10.1016/j.energy.2020.117693

Lin, Z., Liu, X., & Collu, M. (2020). Wind power prediction based on high-frequency SCADA data along with isolation forest and deep learning neural networks. International Journal of Electrical Power and Energy Systems, 118. https://doi.org/10.1016/j.ijepes.2020.105835

Sun, W., & Gao, Q. (2019). Short-term wind speed prediction based on variational mode decomposition and linear-nonlinear combination optimization model. Energies, 12(12). https://doi.org/10.3390/en12122322

Wang, Y.S., Gao, J., Xu, Z.W., Li, L.X. (2019). A short-term output power prediction model of wind power based on deep learning of grouped time series. European Journal of Electrical Engineering, Vol. 22, No. 1, pp. 29-38

Sun, W., & Gao, Q. (2019). Short-term wind speed prediction based on variational mode decomposition and linear-nonlinear combination optimization model. Energies, 12(12).

Catalão, J. P. S., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renewable Energy, 36(4), 1245–1251.

Nielson, J., Bhaganagar, K., Meka, R., & Alaeddini, A. (2020). Using atmospheric inputs for artificial neural networks to IMPROVE wind turbine Power prediction. Energy, 190, 116273.

Balluff, S., Bendfeld, J., & Krauter, S. (2015). Short term wind and energy prediction for offshore wind farms using neural networks. 2015 International Conference on Renewable Energy Research and Applications (ICRERA).

Sun, W., & Wang, Y. (2018). Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. Energy Conversion and Management, 157, 1–12.

Du, M. (2019). Improving lstm neural networks for better short-term wind power predictions. 2019 IEEE 2nd International Conference on Renewable Energy and Power Engineering (REPE). https://doi.org/10.1109/repe48501.2019.9025143

Piotrowski, P., Kopyt, M., Baczyński, D., Robak, S., & Gulczyński, T. (2021). Hybrid and ensemble methods of two days ahead forecasts of electric energy production in a small wind turbine. Energies, 14(5), 1225.

Donadio, Lorenzo, Jiannong Fang, and Fernando Porté-Agel. 2021. "Numerical Weather Prediction and Artificial Neural Network Coupling for Wind Energy Forecast" Energies 14, no. 2: 338. https://doi.org/10.3390/en14020338

Verma, M., Ghritlahre, H.K. Forecasting of Wind Speed by Using Three Different Techniques of Prediction Models. Ann. Data. Sci. (2021). https://doi.org/10.1007/s40745-021-00333-0

Salman, U. T., Rehman, S., Alawode, B., & Alhems, L. M. (2021). Short Term Prediction of Wind Speed Based on Long-Short Term Memory Networks. FME Transactions, 49(3), 643–652. https://doi.org/10.5937/fme2103643S

Shirzadi, Navid, Ameer Nizami, Mohammadali Khazen, and Mazdak Nik-Bakht. 2021. "Medium-Term Regional Electricity Load Forecasting through Machine Learning and Deep Learning" Designs 5, no. 2: 27. https://doi.org/10.3390/designs5020027

Yin, R., Li, D., Wang, Y., & Chen, W. (2020). Forecasting method of monthly wind power generation based on climate model and long short-term memory neural network. Global Energy Interconnection, 3(6), 571–576. https://doi.org/10.1016/j.gloei.2021.01.003

Lin, Z., & Liu, X. (2020). Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network. Energy, 201. https://doi.org/10.1016/j.energy.2020.117693

Berrezzek, F., Khelil, K., Bouadjila, T. (2019). Efficient wind speed forecasting using discrete wavelet transform and artificial neural networks. Revue d'Intelligence Artificielle, Vol. 33, No. 6, pp. 447-452. https://doi.org/10.18280/ria.330607

Aly, H. H. H. (2020). A novel deep learning intelligent clustered hybrid models for wind speed and power forecasting. Energy, 213. https://doi.org/10.1016/j.energy.2020.118773

Bou-Rabee, M., Lodi, K. A., Ali, M., Ansari, M. F., Tariq, M., & Sulaiman, S. A. (2020). One-month-ahead wind speed forecasting using hybrid AI model for coastal locations. IEEE Access, 8, 198482–198493. https://doi.org/10.1109/ACCESS.2020.3028259

Jaseena, K. U., & Kovoor, B. C. (2019). Deep learning based multi-step short term wind speed forecasts with LSTM. In ACM International Conference Proceeding Series. Association for Computing Machinery. https://doi.org/10.1145/3368691.3368698

Du, M. (2019). Improving LSTM Neural Networks for Better Short-Term Wind Power Predictions. In 2019 IEEE 2nd International Conference on Renewable Energy and Power Engineering, REPE 2019 (pp. 105–109). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/REPE48501.2019.9025143

Hong, Y. Y., & Rioflorido, C. L. P. P. (2019). A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. Applied Energy, 250, 530–539. https://doi.org/10.1016/j.apenergy.2019.05.044

Wang, Y., Gao, J., Xu, Z., & Li, L. (2020). A short-term output power prediction model of wind power based on deep learning of grouped time series. European Journal of Electrical Engineering, 22(1), 29–38. https://doi.org/10.18280/ejee.220104

Wang, Y., Gao, J., Xu, Z., & Li, L. (2020). A short-term output power prediction model of wind power based on deep learning of grouped time series. European Journal of Electrical Engineering, 22(1), 29–38. https://doi.org/10.18280/ejee.220104

Nielson, J., Bhaganagar, K., Meka, R., & Alaeddini, A. (2020). Using atmospheric inputs for Artificial Neural Networks to improve wind turbine power prediction. Energy, 190. https://doi.org/10.1016/j.energy.2019.116273

10.2. Gated Recurrent Units (GRU) — Dive into Deep Learning 1.0.0-Alpha1.Post0 Documentation". D2l.Ai, 2022, https://d2l.ai/chapter_recurrent-modern/gru.html.

huai Gao, Yuefei Huang, Shuo Zhang, Jingcheng Han, Guangqian Wang, Meixin Zhang, Qingsheng Lin, Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation, Journal of Hydrology, Volume 589, 2020, https://doi.org/10.1016/j.jhydrol.2020.125188.

Mateus, Balduíno & Mendes, Mateus & Farinha, José & Assis, Rui & Cardoso, A.J.M.. (2021). Comparing LSTM and GRU Models to Predict the Condition of a Pulp Paper Press. Energies. 10.3390/en14216958.

Yang, Shudong & Yu, Xueying & Zhou, Ying. (2020). LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example. 98-101. 10.1109/IWECAI50956.2020.00027.

A. Zhu, X. Li, Z. Mo and R. Wu, "Wind power prediction based on a convolutional neural network," 2017 International Conference on Circuits, Devices and Systems (ICCDS), 2017, pp. 131-135, doi: 10.1109/ICCDS.2017.8120465.

GE General Electric GE 1.6 - 100 - 1,60 MW - Wind turbine (wind-turbine-models.com)

Aamer Bilal Asghar, Xiaodong Liu, Adaptive neuro-fuzzy algorithm to estimate effective wind speed and optimal rotor speed for variable-speed wind turbine, Neurocomputing, Volume 272, 2018, Pages 495-504, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2017.07.022.

Aguinaga, Alvaro & Luo, Xianwu & HIDALGO DIAZ, Victor & Cando, Edgar & Llulluna, Fredy. (2017). A Feed-Forward Backpropagation Neural Network Method for Remaining Useful Life Prediction of Francis Turbines. 10.11159/icmie17.126.

M. Jabbari Ghadi, S. Hakimi Gilani, H. Afrakhte, A. Baghramian, A novel heuristic method for wind farm power prediction: A case study, International Journal of Electrical Power & Energy Systems, Volume 63, 2014, Pages 962-970, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2014.07.008.

Cheggaga N. New Neural Networks Strategy Used to Improve Wind Speed Forecasting. Wind Engineering. 2013;37(4):369-379. https://doi:10.1260/0309-524X.37.4.369

G. L. Tona, M. Luna, A. di Piazza and M. C. di Piazza, "Development of a Forecasting Module based on Tensorflow for Use in Energy Management Systems," IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, 2019, pp. 3063-3068, doi: 10.1109/IECON.2019.8926801.

Shouman, E. R. M. (2022, May 13). Wind power forecasting models. IntechOpen. Retrieved February 3, 2023, from https://www.intechopen.com/chapters/81755

Phi, M. (2020, June 28). Illustrated guide to LSTM's and GRU's: A step by step explanation. Medium. Retrieved January 3, 2023, from https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21.

Tasdelen, A., & Sen, B. (2021, July 8). A hybrid CNN-LSTM model for pre-miRNA Classification. Nature News. Retrieved February 3, 2023, from https://www.nature.com/articles/s41598-021-93656-0.

Monigatti, L. (2022, August 2). Interpreting ACF and PACF plots for time series forecasting. Medium. Retrieved February 14, 2023, from https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c

Brownlee, J. (2019, August 28). How to backtest machine learning models for time series forecasting. MachineLearningMastery.com. Retrieved February 14, 2023, from https://machinelearningmastery.com/backtest-machine-learning-models-time-series-forecasting/.