Proceedings of the 2020 28th Conference on Nuclear Engineering
Joint With the ASME 2020 Power Conference
ICONE28-POWER2020
August 2-6, 2020, Anaheim, California, USA

ICONE28-POWER2020-16557

PREDICTION OF WIND SPEED, POTENTIAL WIND POWER, AND THE ASSOCIATED UNCERTAINTIES FOR OFFSHORE WIND FARM USING DEEP LEARNING

Do-Eun Choe[1], Gary Talor[2], Changkyu Kim[3]

[1]Assistant Professor, College of Engineering, Prairie View A&M University, TX dochoe@pvamu.edu
[2]Student, College of Engineering, Prairie View A&M University, TX
[2]Graduate student, Department of Material Science & Engineering, Texas A&M University, TX

ABSTRACT

Floating offshore wind turbines hold great potential for future solutions to the growing demand for renewable energy production. Thereafter, the prediction of the offshore wind power generation became critical in locating and designing wind farms and turbines. The purpose of this research is to improve the prediction of the offshore wind power generation by the prediction of local wind speed using a Deep Learning technique. In this paper, the future local wind speed is predicted based on the historical weather data collected from National Oceanic and Atmospheric Administration. Then, the prediction of the wind power generation is performed using the traditional methods using the future wind speed data predicted using Deep Learning. The network layers are designed using both Long Short-Term Memory (LSTM) and Bi-directional LSTM (BLSTM), known to be effective on capturing long-term time-dependency. The selected networks are fine-tuned, trained using a part of the weather data, and tested using the other part of the data. To evaluate the performance of the networks, a parameter study has been performed to find the relationships among: length of the training data, prediction accuracy, and length of the future prediction that is reliable given desired prediction accuracy and the training size.

Keywords: Offshore Wind Energy, Energy, Neural Network, Long-Short-Term Memory

1. INTRODUCTION

Most of current offshore wind turbines (OWT) installed up to date are bottom-mounted substructures. This fixed type offshore wind turbines are installed in shallow water less than 25 m depth. In opposition, a massive utilization of wind resource potential in U.S and many other countries has been suggested by installing the OWT in deeper water (Jonkman, 2009; Musial & Ram, 2010). Floating offshore wind turbines (FOWT) are considered as one of the most promising alternatives of current OWT by its greater wind energy potential. The FOWT installed in deeper water can be exposed to stronger and steadier wind fields which helps efficient wind energy generation. Moreover, wind farms in deeper waters are in general less sensitive to space availability, noise restriction, visual pollution, and regulatory problems. However, the inevitable presence of inherent uncertainties in the system inhibits accurate prediction of system design. The multi-faceted uncertainties in FOWT system have contributed to increases the cost of FOWT. As a part of the efforts to minimizing the uncertainties involved in the FOWT system, the paper is focusing on the reliable prediction of the wind energy, which may be critical on locating and siting the wind farms. Seemingly-minor reduction of an uncertainty may contribute to the large amount of the total cost saving. There have been a few research published in the prediction of wind speed using traditional Machine Learning (non-deep-learning), such as Support Vector Machine (SVM), while Deep Learning (DL) technique has hardly been used (Mohandes et al., 2004; Salcedo-Sanz et al., 2011, Zhou et al., 2011, Salcedo-Sanz et al., 2014). By using the future predictability of Deep Learning technique, it is probable to apply this method to build accurate wind models that helps with deciding the optimized locations with less uncertainties compared to previously proposed models.

On the other hand, a recent report (Grossman et al., 2013; Figure 1) shows that there is a rapid growing gap between the amount of data collected and the amount of analyses/usages of the information in science and engineering application. The motivation of this study is to find more effective methods to analyze the cumulated data and knowledges. The propose study in this paper presents a method using Deep Learning (DL) techniques for the wind power prediction.

The paper presents the framework of wind power prediction using a DL technique. The wind speed prediction is performed

using Long Short-Term Memory (LSTM) neural networks. The framework includes design of the network, training and prediction of the wind speed using the trained networks, and calculation of the predicted wind power. The wind energy generation is presented in terms of the wind power density in this paper, which is the wind energy generation per blade sweep area in order to provide the generalized quantities instead of a specific wind turbine.

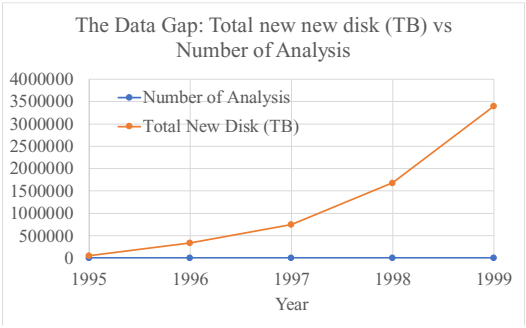

FIGURE 1. DATA-ANALYSIS GAP IN SCIENCE AND ENGINEERING APPLICATION (GROSSMAN ET AL., 2013)

2. LSTM Neural Network

In this paper, the wind speed prediction is performed using deep layered networks of LSTM and Bidirectional LSTM. LSTM neural network is one of Recurrent Neural Network (RNN), which is well-known to capture long-term dependencies and model sequential data. The bi-directional LSTM, which operates both forward and backward process of LSTM, enables the capture of the past and future contexts. Deep bi-directional LSTM is stacked LSTM of multiple layers as elaborated in later this section. In this research, we tested several alternative layers using (1) basic LSTM and multiple layers of LSTM, (2) bi-directional LSTM (BiLSTM) and multiple layers of BiLSTM, and (3) Multiple layers of both LSTM and BiLSTM.

The principal differentiator of LSTMs that a few additional gates are used at each time step to control the passing of information along the sequences which can capture long-term dependencies in higher accuracy detailed in the following equations and **Figure 2**. The input gate $\boldsymbol{i}^t$, the recurrent gate called "forget gate" $\boldsymbol{f}^t$, and the output gate $\boldsymbol{o}^t$, and a memory cell, $\boldsymbol{c}^t$, are represented as follows (Gers et al. 1999 & Zhao etl al., 2017):

$$\boldsymbol{i}^t = \sigma(\boldsymbol{W}^i\boldsymbol{x}^t + \boldsymbol{V}^i\boldsymbol{h}^{t-1} + \boldsymbol{b}^i),$$
$$\boldsymbol{f}^t = \sigma(\boldsymbol{W}^f\boldsymbol{x}^t + \boldsymbol{V}^f\boldsymbol{h}^{t-1} + \boldsymbol{b}^f),$$
$$\boldsymbol{o}^t = \sigma(\boldsymbol{W}^o\boldsymbol{x}^t + \boldsymbol{V}^o\boldsymbol{h}^{t-1} + \boldsymbol{b}^o),$$
$$\boldsymbol{c}^t = \boldsymbol{f}^t \odot \boldsymbol{c}^{t-1} + \boldsymbol{i}^t \odot \tanh(\boldsymbol{W}^c\boldsymbol{x}^t + \boldsymbol{V}^c\boldsymbol{h}^{t-1} + \boldsymbol{b}^c),$$
$$\boldsymbol{h}^t = \boldsymbol{o}^t \odot \tanh(\boldsymbol{c}^t).$$

where $\boldsymbol{W} \in \mathbb{R}^{d\times k}$, $\boldsymbol{V} \in \mathbb{R}^{d\times d}$ $\boldsymbol{b} \in \mathbb{R}^d$, are the model parameters, k associated with W is dimensionality of hidden vectors, and $\odot$ represents an element-wise product. Figure 2 (a) shows the basic structures of LSTM regression.

The bi-directional LSTM uses both forward and backward processed information as shown in Figure 2 (b), which can be represented as follows:

$$\overrightarrow{\boldsymbol{i}}t = \sigma\left(\overrightarrow{\boldsymbol{W}}i\ \overrightarrow{\boldsymbol{x}}t + \overrightarrow{\boldsymbol{V}}i\overrightarrow{\boldsymbol{h}}t-1 + \overrightarrow{\boldsymbol{b}}i\right),$$
$$\overrightarrow{\boldsymbol{f}}t = \sigma\left(\overrightarrow{\boldsymbol{W}}f\ \overrightarrow{\boldsymbol{x}}t + \overrightarrow{\boldsymbol{V}}f\overrightarrow{\boldsymbol{h}}t-1 + \overrightarrow{\boldsymbol{b}}f\right),$$
$$\overrightarrow{\boldsymbol{o}}t = \sigma\left(\overrightarrow{\boldsymbol{W}}o\ \overrightarrow{\boldsymbol{x}}t + \overrightarrow{\boldsymbol{V}}o\overrightarrow{\boldsymbol{h}}t-1 + \overrightarrow{\boldsymbol{b}}o\right),$$
$$\overrightarrow{\boldsymbol{c}}t = \overrightarrow{\boldsymbol{f}}t \odot \overrightarrow{\boldsymbol{c}}t-1 + \overrightarrow{\boldsymbol{i}}t \odot \tanh\left(\overrightarrow{\boldsymbol{W}}c\ \overrightarrow{\boldsymbol{x}}t + \overrightarrow{\boldsymbol{V}}c\overrightarrow{\boldsymbol{h}}t-1 + \overrightarrow{\boldsymbol{b}}c\right),$$
$$\overrightarrow{\boldsymbol{h}}t = \overrightarrow{\boldsymbol{o}}t \odot \tanh\left(\overrightarrow{\boldsymbol{c}}t\right).$$
$$\overleftarrow{\boldsymbol{i}}t = \sigma\left(\overleftarrow{\boldsymbol{W}}i\ \overleftarrow{\boldsymbol{x}}t + \overleftarrow{\boldsymbol{V}}i\overleftarrow{\boldsymbol{h}}t+1 + \overleftarrow{\boldsymbol{b}}i\right),$$
$$\overleftarrow{\boldsymbol{f}}t = \sigma\left(\overleftarrow{\boldsymbol{W}}f\ \overleftarrow{\boldsymbol{x}}t + \overleftarrow{\boldsymbol{V}}f\overleftarrow{\boldsymbol{h}}t+1 + \overleftarrow{\boldsymbol{b}}f\right),$$
$$\overleftarrow{\boldsymbol{o}}t = \sigma\left(\overleftarrow{\boldsymbol{W}}o\ \overleftarrow{\boldsymbol{x}}t + \overleftarrow{\boldsymbol{V}}o\overleftarrow{\boldsymbol{h}}t+1 + \overleftarrow{\boldsymbol{b}}o\right),$$
$$\overleftarrow{\boldsymbol{c}}t = \overleftarrow{\boldsymbol{f}}t \odot \overleftarrow{\boldsymbol{c}}t+1 + \overleftarrow{\boldsymbol{i}}t \odot \tanh\left(\overleftarrow{\boldsymbol{W}}c\ \overleftarrow{\boldsymbol{x}}t + \overleftarrow{\boldsymbol{V}}c\overleftarrow{\boldsymbol{h}}t+1 + \overleftarrow{\boldsymbol{b}}c\right),$$
$$\overleftarrow{\boldsymbol{h}}t = \overleftarrow{\boldsymbol{o}}t \odot \tanh\left(\overleftarrow{\boldsymbol{c}}t\right).$$

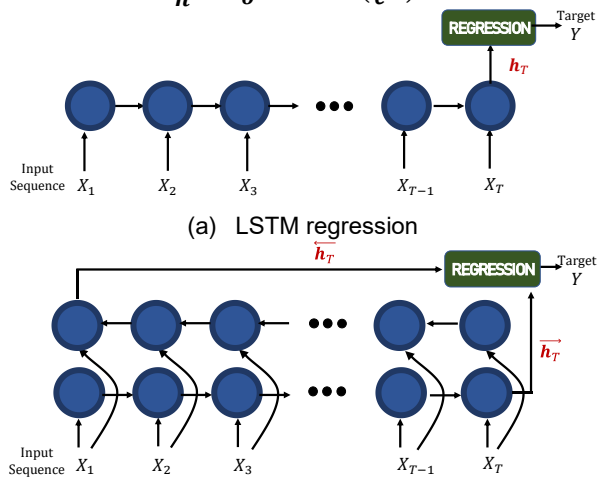

FIGURE 2 LSTM REGRESSION VS BI-DIRECTIONAL LSTM REGRESSION

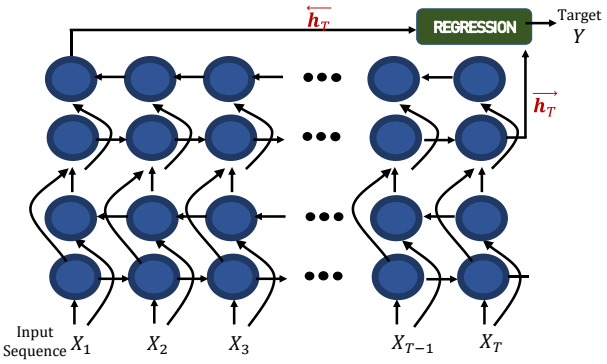

FIGURE 3. DEEP BI-DIRECTIONAL LSTM LAYERS

Figure 3 shows the deep LSTM layers, which is stacked bi-directional LSTMs. The hidden output of one LSTM layer is propagated through time as in a single i-directional LSTM, and at the same time, the output of the one LSTM layer is used as the input data to the next bi-directional LSTM. Every hidden layer receives an input sequence which includes the output sequences of both forward and backward layers from the previous layer. Figure 4 shows sample configurations of layering the LSTM and BiLSTM networks during the full analyses using MATLAB.

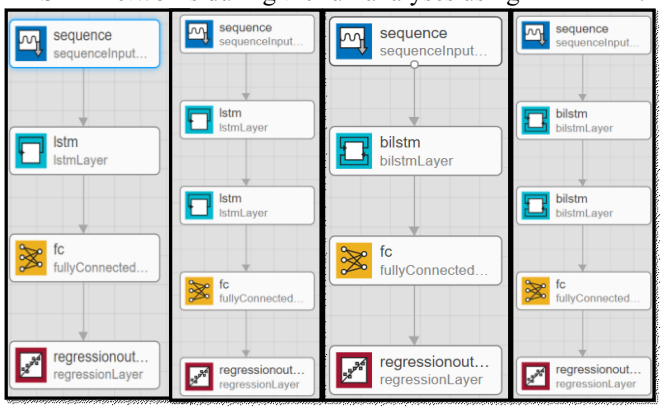

FIGURE 4. ALTERNATIVE LAYERS OF LSTM AND BI-DIRECITONAL LSTM

3. Wind Speed Predicted by LSTM layers

The data used for the example analysis of this research was collected from the National Oceanic and Atmospheric Administration (NOAA) from one specific buoy that gathered information for an extended amount of time. The coordinates of this buoy where 54.97, -9.04, which is roughly 16 miles off the coast of the island of Arranmore of Ireland. The wind data time range that was used was from May 2009 to May 2019, encompassing 10 full years of wind information and was measured from around 100 meters from sea level.

For the prediction of the wind speed, several alternative LSTM and bi-LSTM layers are designed as shown in Figure 4. After studies of various alternative layers, it is found that the benefit of layering multiple network is not significant. Therefore, the analysis is performed with a single LSTM. To train the designed networks, the series of observations $\boldsymbol{x} = [x_1, x_2, x_3, x_4, \dots x_n]$ representing the wind speed data over 10 years in every hour interval are used. The set is separated by $\boldsymbol{x} = [\boldsymbol{x_{TR}}, \boldsymbol{x_{TS}}]$, where $\boldsymbol{x_{TR}}$, represents training data and $\boldsymbol{x_{TS}}$ represents the data used to test the performance of the model. The input training set is $\boldsymbol{x_{TR}} = [x_{TR-1}, x_{TR-2}, x_{TR-3}, \dots x_{TR-n-1}]$ and the output training set is represented by $\boldsymbol{y_{TR}} = [x_{TR-2}, x_{TR-3}, x_{TR-4}, \dots x_{TR-n}]$. The analyses performed in only hourly level. However, the future analyses are further expendable in terms of daily and yearly level.

Once the models are trained using the sets of data, the trained model is tested for its prediction in two different strategies, P-A & P-B: P-A predicts $\boldsymbol{y_{TS,A}}$ by the inputs of, $\boldsymbol{x_{TS,A}} = [x_{TS-1}, x_{TS-2}, x_{TS-3}, \dots x_{TS-n-1}]$ with continuous training while testing and P-B predicts $\boldsymbol{y_{TS,B}}$ the inputs, $\boldsymbol{x_{TS,B}} = [x_{TS-1}, y_{TS,A-1}, y_{TS,A-2}, \dots y_{TS,A-n}]$. For the parameter study, for each strategies, the analyses were performed for the training size 3600 hrs (150 days or approximately 5 month) to 8640 hrs (360 days, approximately 12 month) with the increments of every 10 days. The prediction period was performed with varying ranges from 216 hrs (9 days) to 720 hrs (30 days or 1 month) with the increments of 72 hrs (3 days).

4. Results: Predicted Wind Speed

In this section, the results of the wind speed at the buoy location are presented and analyzed. For the parameter study, total of 352 analyses are performed for both P-A & P-B prediction strategies (22 increments of training length×8 increments of prediction length × 2 strategies = 352). Parameter studies are presented for given size of the training data (N_T) in order to achieve the accuracy, measured in Root Mean Sqaure Error (RMSE). The length of the prediction, N_P.was added as the third parameter to observe the reliable prediction length given required prediction accuracy. The analyses are performed in hourly prediction level. Future analyses are planned for daily and yearly prediction levels

Figure 5 and Figure 6 show sample analyses of the hourly wind speeds (m/s). Both are the results for the same test length of 504 hrs (21 days), while Figure 5 shows the results using 3600 hrs (5 month) of training length and Figure 6 shows those using 1440 hrs (2 month) of training length. The prediction strategy P-A is to train while testing. The two predictions show different performances in the begging of the prediction. The smaller training size (Figure 6) shows larger errors in the rage of prediction during the first 24 hrs period, while it gradually converged to the observation curve. Similar trend is shown in the larger-training-sized model Figure 5, which is very minor compare to model with smaller training size (Figure 6). Overall errors are ranged similar. The P-A predictions show convergence after 24 hrs of prediction.

Figure 7 (P-A) and Figure 8(P-B) show the results from the 352 sets of daily wind speed (m/s) predictions. The P-A prediction shows the errors decrease as the prediction period increases. This is due to the network improvement while predicting the futures. P-B prediction assumed there is no further information given after the training is completed. Therefore, as the prediction period increases, the errors (RMSE) increase together. In addition, the errors (RMSE) decrease as the training size increases.

5. Results: Predicted Wind Power Generations

The predicted wind power generation in this paper is represented by the wind energy density, which is the energy generation per a blade sweep area A in order to provide generalize results rather than those for a specific wind turbine. The wind energy density is calculated based on the DL- predicted wind speed using traditional wind power calculation method detailed in this section.

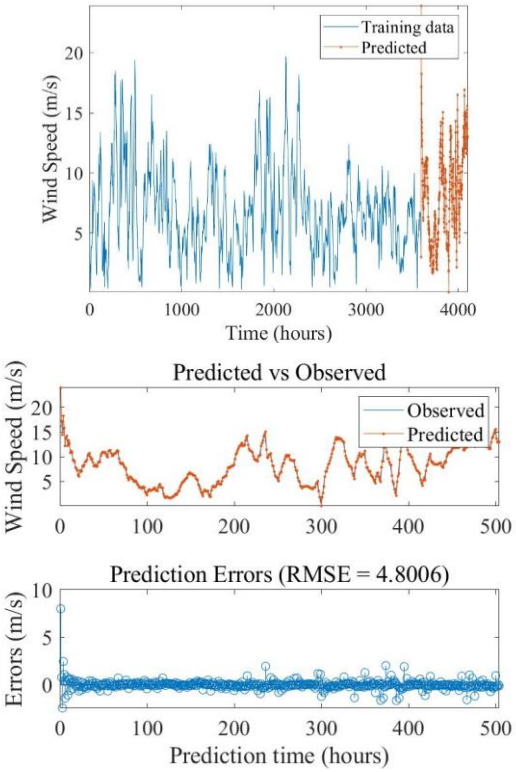

FIGURE 5. SAMPLE RESULTS OF HOURLY PREDICTION: P-A, TRAINING 3600HRS, TESTING 504 HRS

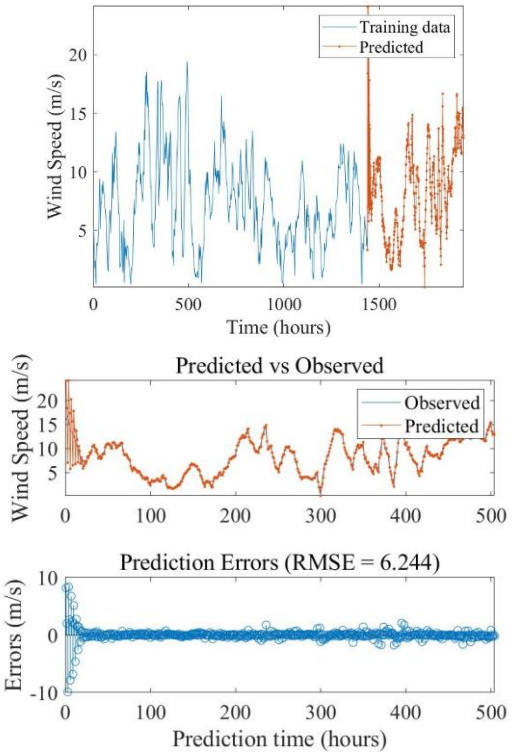

FIGURE 6. SAMPLE RESULTS OF HOURLY PREDICTION: P-A, TRAINING 1400HRS, TESTING 504 HRS

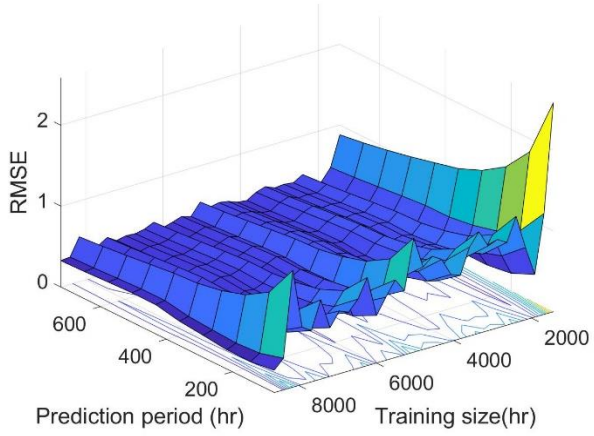

FIGURE 7. PARAMETER STUDY: TRAINING SIZE, PREDICTION ERROR (RMSE), & PREDICTION LENGTH FOR P-A

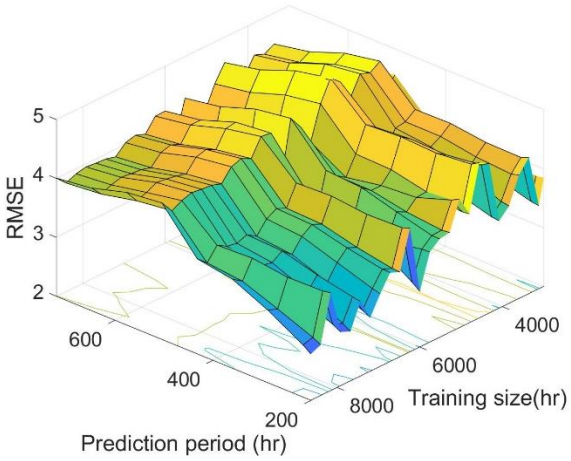

FIGURE 8. PARAMETER STUDY: TRAINING SIZE, PREDICTION ERROR (RMSE), & PREDICTION LENGTH FOR P-B

Wind energy density, the energy generated per a blade sweep area A is estimated based on the Weibull distribution of wind speed, which is one of commonly used methods for wind power calculation. The probability densify function and cumulative distribution function are written as following:

$$f(V) = \frac{k}{c}\left(\frac{V}{c}\right)^{k-1} e^{-(\frac{V}{c})^k}$$

$$F(V) = 1 - e^{-(\frac{V}{c})^k}$$

where k is the shape parameter, c is the scale parameter, and V is wind speed. The shape parameter k can be calculated using maximum likelihood method (Jiang et al. 2016). However, for this research a simplified expression, which is derived from a modified maximum likelihood method by Christofferson and Gillett (1987), is used to avoid the complexity and the iterations

of the calculation. The parameters *k* and *c* are expressed as (Perez et al., 2007):

$$k = \frac{\pi}{\sqrt{6}}\left[\frac{N(N-1)}{N(\sum_{i=1}^{N} ln^2 V_i) - (\sum_{i=1}^{N} \ln V_i)^2}\right]^{0.5}$$

$$c = \left(\frac{1}{N}\sum_{i=1}^{N} V_i^k\right)^{1/k}$$

where N represents the number of the total observation, V is wind speed at each data point. Based on the distribution, the wind energy density and can be written as (Zhao et al., 2017):

$$\frac{E}{A} = \frac{1}{2}\rho c^3 \Gamma\left(\frac{k+3}{k}\right) T$$

where E denote the wind energy generated by the wind turbine, A denote a blade sweep area, ρ is fluid density, Γ demotes the Gamma function, and T is the time period or duration of the measure expressed in hour.

Figure 9 shows sample prediction of the wind power density calculation. Left of the figure shows the testing size 288 and the right side of the figure shows the testing size of 504 hrs, using the same size of the training data. Figure 10 shows displays the different size of the training data with the same testing size of 720 hrs. The beginning of the prediction shows the only differences. The results are obtained based on P-A prediction data.

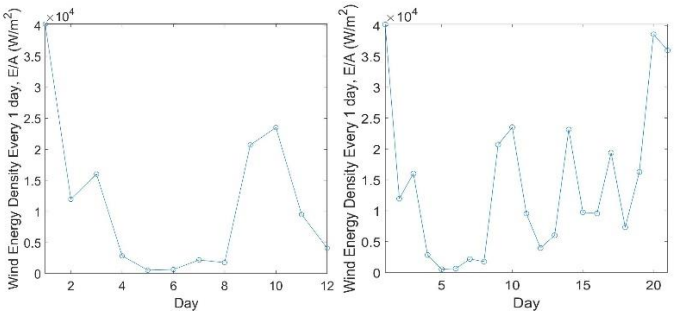

FIGURE 9. WIND ENERGY DENSITY PREDICTION WITH TRAINING LENGTH OF 1440 HRS FOR PREDICTIONS OF 288 HRS (LEFT) & 504 HRS (RIGHT)

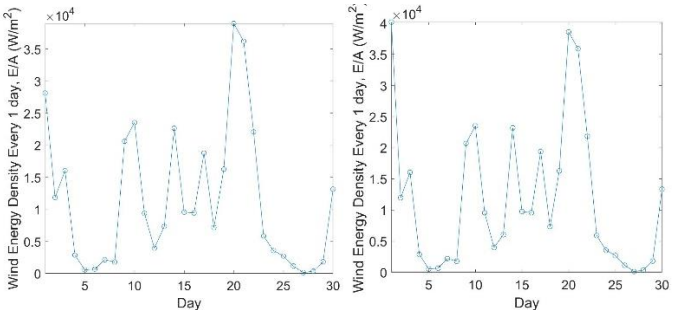

FIGURE 10. WIND ENERGY DENSITY PREDICTION WITH TRAINIG SIZES OF 1440 HRS (LEFT) AND 6000 (RIGHT) AND PREDICTION LENGTH OF 720 HRS.

6. CONCLUSION & FUTURE WORK

The paper presents a research-in progress to predict the wind power generation at a potential wind farm location using Deep Leaning techniques. Parameter study is presented by performing 352 sets of training and testing to find the relationships among the size of the training data and the prediction error for the given prediction length. Example analyses are presented with the predicted wind speed and wind power generation at a selected buoy location in hourly scale. Future research includes the investigation on additional weather parameters training and testing, improvement of the network layers, and incorporation of the physics in DL to improve the prediction.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant #1700406.

REFERENCES

Chang, T. J., Wu, Y. T., Hsu, H. Y., Chu, C. R., & Liao, C. M. (2003). Assessment of wind characteristics and wind turbine characteristics in Taiwan. Renewable energy, 28(6), 851-871.

Christofferson, R. D., & Gillette, D. A. (1987). A simple estimator of the shape factor of the two-parameter Weibull distribution. Journal of climate and applied meteorology, 26(2), 323-325.

Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.

Jiang, W., Xie, C., Zhuang, M., Shou, Y., & Tang, Y. (2016). Sensor data fusion with z-numbers and its application in fault diagnosis. Sensors, 16(9), 1509.

Jonkman, J. M. (2009). Dynamics of offshore floating wind turbines—model development and verification.

Mohandes, M. A., Halawani, T. O., Rehman, S., & Hussain, A. A. (2004). Support vector machines for wind speed prediction. Renewable energy, 29(6), 939-947.

Musial,W. & Ram,B. (2010), Large-Scale Offshore Wind Powerin the United States-Assessment of Opportunities and Barriers. Technical Report. NREL/TP-500-40745.

Pérez, I. A., Sánchez, M. L., & García, M. Á. (2007). Weibull wind speed distribution: Numerical considerations and use with sodar data. Journal of Geophysical Research: Atmospheres, 112(D20).

Salcedo-Sanz, S., Ortiz-Garcı, E. G., Pérez-Bellido, Á. M., Portilla-Figueras, A., & Prieto, L. (2011). Short term wind speed prediction based on evolutionary support vector regression algorithms. Expert Systems with Applications, 38(4), 4052-4057.

Salcedo-Sanz, S., Pastor-Sánchez, A., Prieto, L., Blanco-Aguilera, A., & García-Herrera, R. (2014). Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization–Extreme learning machine approach. Energy Conversion and Management, 87, 10-18.

United States Department of Energy. (2015). Wind Vision: A New Era for Wind Power in the United States. Executive summary March 2015. Report nr DOE/GO-102015–4557, 50 pp.

Zhao, R., Yan, R., Wang, J., & Mao, K. (2017). Learning to monitor machine health with convolutional bi-directional LSTM networks. Sensors, 17(2), 273.

Zhou, J., Shi, J., & Li, G. (2011). Fine tuning support vector machines for short-term wind speed forecasting. Energy Conversion and Management, 52(4), 1990-1998.