

# Short Term Prediction of Wind Speed Based on Long-Short Term Memory Networks

**Umar T. Salman**

Electrical Engineering Department  
KFUPM, Dhahran  
Saudi Arabia

**Shafiqur Rehman**

Center for Engineering Research  
The Research Institute, KFUPM, Dhahran  
Saudi Arabia

**Basit Alawode**

Electrical Engineering Department  
KFUPM, Dhahran  
Saudi Arabia

**Luai M. Alhems.**

Center for Engineering Research  
The Research Institute, KFUPM, Dhahran  
Saudi Arabia

*Power utilities, developers, and investors are pushing towards larger penetrations of wind and solar energy-based power generation in their existing energy mix. This study, specifically, looks towards wind power deployment in Saudi Arabia. For profitable deployment of wind power, accurate knowledge of wind speed both in spatial and time domains is critical. The wind speed is the most fluctuating and intermittent parameter in nature compared to all the meteorological variables. This uncertain nature of wind speed makes wind power more difficult to predict ahead of time. Wind speed is dependent on meteorological factors such as pressure, temperature, and relative humidity and can be predicted using these meteorological parameters. The forecasting of wind speed is critical for grid management, cost of energy, and quality power supply. This study proposes a short-term, multi-dimensional prediction of wind speed based on Long-Short Term Memory Networks (LSTM). Five models are developed by training the networks with measured hourly mean wind speed values from 1980 to 2019 including exogenous inputs (temperature and pressure). The study found that LSTM is a powerful tool for a short-term prediction of wind speed. However, the accuracy of LSTM may be compromised with the inclusion of exogenous features in the training sets and the duration of prediction ahead.*

**Keywords :** ANN, errors, forecasting, LSTM, wind speed, wind power.

## 1. INTRODUCTION

Wind amongst other renewable sources, is becoming more popular for both grids connected large applications and isolated grids for small loads. The grid connectivity issues and power grid managements control are getting advanced with time. Saudi Arabia, for example, is targeting a new wind power capacity of 2.0GW by 2030 [1] and is planning to achieve 20% of wind power penetration in its total capacity by the end of 2030 [2]. However, the challenge associated with wind power is its uncertain nature [3, 4]. Wind power is mainly affected by wind speed and weather factors such as wind direction, temperature, atmospheric pressure, and relative humidity [5]. The geographical and topographical conditions at the wind farm sites have an influence on the output of the wind energy [6, 7]. An accurate information of the wind speed availability, that drive the wind turbines, is crucial for microgrid-siting and later for profitable and proper operation and maintenance of the system [8, 9]. The integration of wind power in microgrids where the effects of wind power fluctuations significantly affect the microgrid operation and other distributed generation were studied in [7–10].

In micro-grids, the integration of large-scale wind power may result in the fluctuations of the output power

and can cause disturbances to the power system and may lead to grid failures [11, 12]. Consequently, power quality, voltage, and frequency may be seriously affected [5, 11].

There are physical parameters of the wind turbines such as pitch, rotor diameter, blade length, and wind-farm layout design that greatly affect the power output from the turbine. A study of wind farm layout design is presented in [13], where particle swarm optimization (PSO) algorithm is used to solve wind farm optimization problem. Besides, the authors in [14] studied the optimal location of wind turbines and their performance in Iraq considering the costs and maximum possible capacities from the wind turbines in different sites. Furthermore, the steady-state deformation and stress that occur in wind turbine blades were investigated in [15]. A review of studies involving the design, optimization and techniques of different wind turbine blades were reported in [16].

The ability to accurately forecast wind speed can reduce the adverse effects of wind power fluctuations on the power system. Reliable prediction of wind speed ahead of time can allow the power management system enough time to manage the fluctuations in power through other controls. It is reported that wind speed and ultimately wind power forecasting methods are divided into statistical learning methods, physical models, modern machine learning techniques, and hybrid methods. The statistical approaches are simple and use historical data to predict wind speed in future time domain [2]. These are based on statistical time series or machine learning algorithms [17] which may

---

Received: May 2021, Accepted: June 2021

Correspondence to: Umar T. Salman,  
Electrical Engineering Department, KFUPM,  
Dhahran, Saudi Arabia

E-mail: salmantaiwo@gmail.com

doi:10.5937/fme2103643S

© Faculty of Mechanical Engineering, Belgrade. All rights reserved

FME Transactions (2021) 49, 643-652 643

include the artificial neural network (ANN) and support vector machine continuous association rule mining algorithm (CARMA) methods [18–20].

The physical approaches use the physical description (solving differential equation of mass, energy, and velocity) of the wind to model the onsite conditions and it is good for long term prediction [10]. However, the physical methods are complex and make use of many inputs resulting in a high computational cost [5]. The hybrid or combination methods on the other hand, combines the predictive ability of various methods to improve the accuracy of the model. Owing to its memory retaining capability and its performance on sequence models, especially in natural language processing (NLP), recurrent neural networks (RNN) and its variants have become the de-facto models in renewable energy forecasting. The Long-Short Term Memory (LSTM) networks particularly have shown tremendous performance in their use for wind speed energy prediction and there is no shortage of work that has been done in this area. However, the vast majority of these applications have been based on certain environments. The major reason for this is the fact that environments play a vital role in how renewable energy systems behave. This in turn have effect on their prediction. As such, each environment is unique and thus often require unique strategy for their predictions. Hence, the need for our research.

In the presence of extensive dataset, a hybrid of Ensemble Empirical Model Decomposition (EEMD), Genetic Algorithm (GA), and LSTM was proposed in [21] as a way to achieve short-term wind speed prediction. The EEMD was utilized to decompose the sequence wind speed data for feature extraction and the GA-LSTM algorithm was applied to the extracted features to predict future wind speed. The feature extraction step culminates in a 38.48% increase in prediction accuracy as opposed to the non EEMD model. The data utilized was extracted from the wind speed data of various sites across the United States. A combination of variational mode decomposition (VMD), singular spectrum analysis (SSA), LSTM, and extreme learning machines (ELM) was adopted in [22] for wind speed forecasting. The SSA step was utilized as a form of feature extraction step to boost the prediction accuracy. Several experiments were performed in which their multi-step algorithm was compared with several other algorithms and achieved better accuracy. The data utilized here is from the samples obtained from one observation site of a wind farm in China between May to December, 2015. Again, the ELM and LSTM was combined with differential evolution algorithm for wind speed forecasting in [23]. In their work, data was gathered from a wind farm in Inner Mongolia, China. Two forecasting models were developed in which one was for short-term prediction of 10 minutes and the other for slightly longer prediction of one-hour ahead. Autoencoders with LSTM [24], two-stage LSTM [25] are some other methods that have been employed in wind speed forecasting. A comprehensive review of the various LSTM and support vector machines (SVM) models utilized for wind speed forecasting is equally presented in [26].

In this study, a short-term prediction of wind speed based on LSTM is proposed. The study presents five different Cases to investigate the predictive ability of LSTM in the short-term prediction of wind speed. Besides, four Scenarios of time step ahead predictions are further proposed to investigate the performance of LSTM network in the presence of exogenous inputs. The strategy adopted in this paper unveils unique characteristics of LSTM in short term prediction of wind speed when the training features are selected in different combination.

## 2. DATA PREPARATION

Predication of wind speed requires the use of weather parameters temperature, atmospheric pressure, and relative humidity. Although, highly correlated parameter has the strongest impact on the wind data, it usually not a good practice to include such data in ANN forecast in the presence of other parameters [27]. However, in this study, wind direction data is not considered because of its even extremely high fluctuating nature. Furthermore, it does not add to the wind speed intensity as well. The hourly mean data is obtained from a meteorological station located in Dhahran, Eastern Province, Saudi Arabia and covers a period of 40 years from 1980 to 2019 resulting in 342,624 data points out of which 100,000 have been used for training.

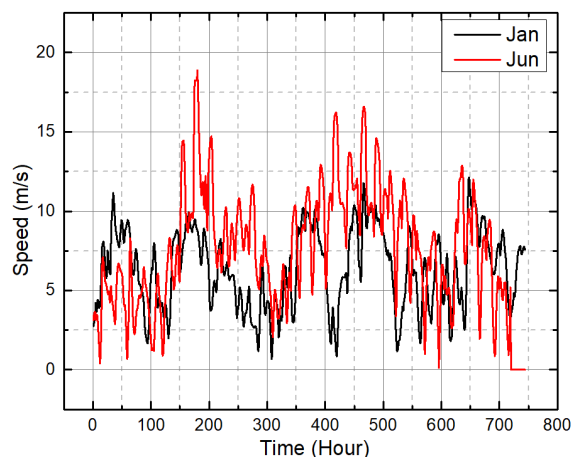


Figure 1. Wind speed measured at 50m height, Dhahran 2018.

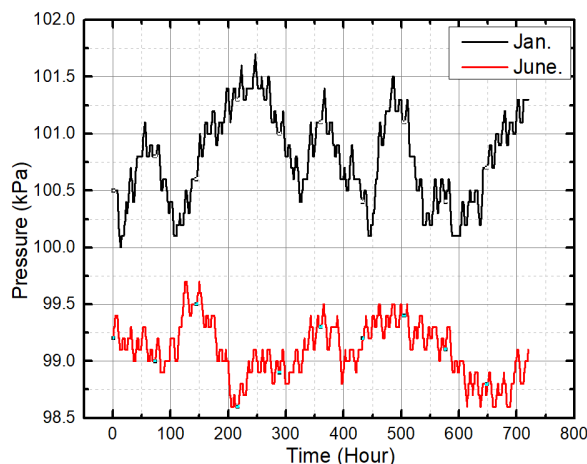


Figure 2. Atmospheric pressure measured near the ground; Dhahran 2018.

The erratic nature of the weather parameters, hourly variation with time (during January and June months for 2018), is shown in Figs. 1, 2 and 3. Higher magnitudes of wind speed values are observed in June compared to January (Fig. 1). This indicates the wind speed seasonal variability which must be addressed by the models. Higher values of atmospheric pressure measured are observed during winter (January) compared to summertime (June) as shown in Fig. 2. The effect of height (2m and 10m) and the season on ambient temperature values is clearly visible in Figs. 3 (a & b). Two temperature measurements are considered at heights of 10m and 2m because vertical temperature difference causes the movement of the air masses [28, 29].

## 2.1 Data Processing

The original data include the wind speed, temperature at two heights, wind direction and pressure on hourly averaged basis. All the column data is scaled to a range of [0, 1] to obtain the best network performance. Since the inputs to the network is the past time data series, so to obtain better prediction accuracy, previous 60 hours measured values are considered. Based on the input data and the models, the future predictions are made under four Scenarios which are 1, 2, 5, and 7 hours ahead for all the cases presented here.

As an instance, Case 2 utilizes wind speed data measured at 50m, and temperatures at 10m, and at 2m to predict the future wind speed. To demonstrate the accuracy and predictability of the models, several batches, each of 60 input data points, are selected randomly from the total of 100,000 data points and next hour data as output. Although other values of input data points such as 30, 45 and 75 were also tried but 60 gives the best trade-off between accuracy, and training time. This gives an input shape of (99940, 60, 3). The output is the corresponding future wind speed of shape (99940, 1). From this data, 5000 randomly selected data points are used as the testing data and the remaining as training sets, respectively.

## 2.2 Network Training

The network was trained based on backpropagation through time (BPTT) using the Adam optimizer. The mean squared error (MSE) metric is used for the loss function. The goal of the optimizer is to minimize the MSE between the predicted future and the measured values. Furthermore, we desired that the network be retrained several times if the value of mean absolute error (MAE) exceeds a predefined some threshold value (0.5 m/s). If this is not achieved after several trials, then the best of the trained results will be retained. During training, a batch of 128 data points are utilized and the network is trained for 40 epochs. For all the cases, training took on for an average of 28 minutes on an intel Core i7, 8th Generation CPU with 16 GB of RAM running on a windows 10 OS.

## 3. METHODOLOGY

Each parameter (wind speed, temperature, and pressure) of the training sample is modeled as time series used as input to the models. The time series samples are made of

$n$  observations  $[x_1, x_2, \dots, x_n]$  which are used to predict the wind speed in future time domain. Five models are developed; each model is trained for one-hour ahead prediction and thereafter-tested 1 hour, 2 hours, 5 hours and 7 hours ahead. This implies that the input time series is a function of the past historical hourly mean values.

The training input data for each sample is modeled as a function  $F(\cdot)$ , as represented in Equation (1):

$$x^i(t) = F(x^i(t), x^i, t, \dots, x^n(t-1)) \quad (1)$$

where  $i$  is the number of input parameters being considered, and  $t$  is the number of sample observations.

## 3.1 Network Training

ANN has network architecture consisting of neurons, connecting strength, nodes properties, and updating rules [30, 31]. The neurons have natural ability to store and figure out experimental knowledge which can be used to validate future occurrences [32]. Some of the unique attributes of the ANN are its capabilities to process information with very high speed, mapping, tolerance for faults, robustness and generalization. Thus, an ANN is excellent in performance when it comes to system identification, system modeling, optimization, and prediction [33]. The ANN has been used to solve complex nonlinear engineering problems in real-world [32–35]. The ANN model is made of parametric components like weights ( $w_{ij}$ ), connecting synapses or links, bias ( $b_j$ ), and activation function  $f^*$ . These parameters relate the input  $x_i$  to the output  $y_j$  as shown in Fig. 4 and expressed in the following equations. Each input,  $x_i$  is multiplied by the weight,  $w_{ij}$  to give the summation output  $s_j$ , Equation (2a). The output of the summation applied to the activation function to give the final output signal  $y_j$ , equation (2b) which is the desired limit or range of the amplitude. Example of activation functions listed in [33] are linear, sigmoid, Gaussian, and Gaussian complements which chosen based on specific problem.

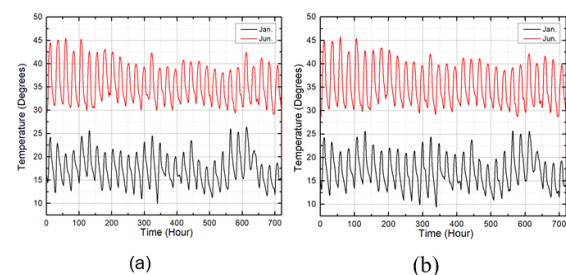


Figure 3. Temperature measured at height (a) 10m and (b) 2m, Dhahran 2018.

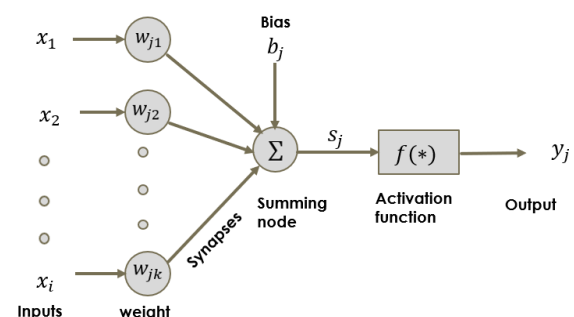


Figure 4. The model of a non-linear neuron.

where  $j$  and  $k$  are the numbers neurons and synapses respectively.

$$S_j = \sum_{k=1}^i w_{jk} \quad (2)$$

$$y_j = f \left( \sum_{k=1}^i w_{jk} + b_j \right) \quad (3)$$

As suitable for this problem, a type of neural network proposed in this study is a recurrent neural network called long-short term memory (LSTM) networks. A typical structure of a LSTM is shown in Fig. 5.

### 3.2 Long-short Term Memory Networks

The LSTM NN was initially proposed in [36] and it is a special type of recurrent network. The LSTM is based on the principle that the status of a current cell can be affected by the status of the previous cells. This is in fact, a recurrent neural network.

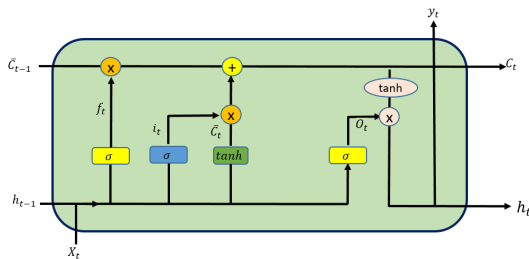


Figure 5. A typical structure of an LSTM Network.

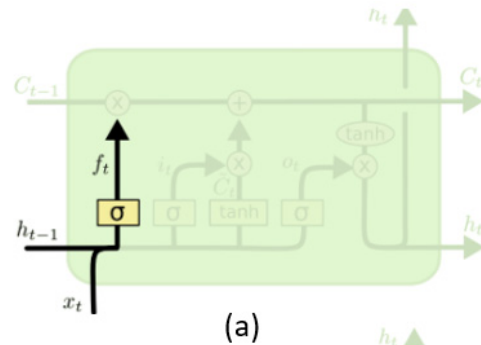
This description is depicted in Fig. 5. An LSTM cell has an output gate, input gate, and a forget gate layer also called the sigmoid layer. The function of the input gate is to control the amount of data that goes into a cell. The forget gate regulates the amount of the values left in a cell while the output gate together with values in the cell defines the output of an LSTM.

The LSTM can extract relevant information from streams of data, remember them, and use them to predict future values [37]. The tanh network ensures that values stay between -1 and 1, thus regulate the NN output. The cell state acts as hardware. It can carry relative information through the sequence chain. It is essentially the memory of the network. Information from the earlier time steps carried to the previous time steps, this helps to reduce the short-term memory [38].

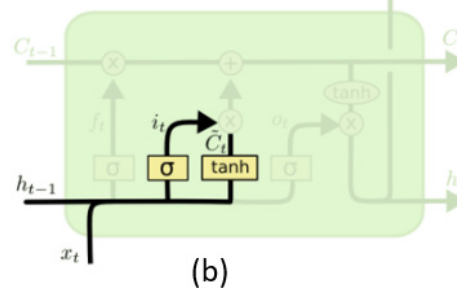
The gates are types of NN that determine the information that are allowed in the cell state, they learn information that would be kept or forgetting during training. The activation function in the gates is sigmoid like the tanh function. The sigmoid squeezed information between 0 and 1. This helps to decide what data to remember or to forget. When a value is multiplied by zero, it is forgotten and when it is multiplied by 1, it is remembered. So, the gates help to regulate information flow within a cell.

The forget gate (Fig. 6(a)) decides which information is kept or discarded. Information from the previously

hidden state,  $h_{t-1}$  is combined with the one in the current state and then fed into the forget gate. The output of the forget gate  $f_t$  is passed into the sigmoid (3) and the information in the hidden state is passed on for further processing. Information is further processed in the next stage which is made of a sigmoid layer, known as the input gate and a tanh layer. The input gate and a tanh layer, shown in Figure 6(b) determines what information is stored in the cell state. While the tanh layer creates a vector of new candidate to be added to the state, the sigmoid layer decides the value to be updated. The output from the sigmoid output decides which information to be kept from the tanh activation output. Next, the cell state is calculated using the information from the gates and the hidden state using Equation (1). Because the hidden state has information about the previous inputs and it helps in prediction. The tanh output is multiplied with the sigmoid output (Fig. 7) to determine the information in the hidden state ( $h_t$ ) which is now the output. The new cell and the hidden states then carried over to the next time step.



(a)



(b)

Figure 6. A section of the of an LSTM showing (a) forget gate, (b) hidden gate [37].

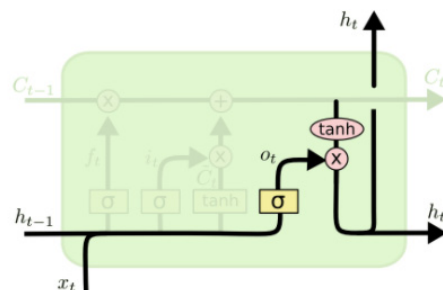


Figure 7. A section of the of an LSTM showing how the combination of the sigmoid and tanh output [37].

The steps involved in calculating the inputs, forget, and the output of the gates are described as follows:

Step 1: We calculate the inputs of the three gates and the candidate cell using Equations (3) to (6).

$$net_f = W_f \cdot [h_{t-1}, x_t] + b_f \quad (3)$$

$$net_i = W_i \cdot [h_{t-1}, x_t] + b_i \quad (4)$$

$$net_o = W_o \cdot [h_{t-1}, x_t] + b_o \quad (5)$$

$$net_{\bar{c}} = W_{\bar{c}} \cdot [h_{t-1}, x_t] + b_{\bar{c}} \quad (6)$$

where  $net_{\bar{c}}, net_f, net_i, net_o$  are the candidate cell, forget gate, input gate, and output gate respectively.

Step 2: We compute the three gate units using equation (7-10).

$$f_t = \sigma(net_f) \quad (7)$$

$$i_t = \sigma(net_i) \quad (8)$$

$$O_t = \sigma(net_o) \quad (9)$$

$$\bar{C}_t = \tanh(net_{\bar{c}}) \quad (10)$$

$$C_t = f_t \times C_{t-1} + i_t \times \bar{C}_t \quad (11)$$

where  $C_t$  is the new cell state,  $f_t$  is forget gate,  $i_t$  is the input gate,  $O_t$  is the output gate and  $\bar{C}_t$  is candidate all at time t. The activation function  $\sigma(\cdot)$  and  $\tanh(\cdot)$  are defined in equation (12)

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (12)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (13)$$

Step 3: Finally, we calculate the output,  $h_t$  of the LSTM using equation (14).

$$h_t = O_t \times \tanh(C_t) \quad (14)$$

#### 4. PROBLEM FORMULATION

The training samples consist of training inputs and targets. The inputs are the hourly wind speed (HWS), hourly temperatures (at 10m and 2m) and hourly pressure during 1980 to 2019. The training and the testing features are further described in the next subsections.

##### 4.1 Training Parameters

The description of the training parameters and their combinations to achieve the different Cases is depicted in Fig. 8. Furthermore, the results of each cases compared and the effect of each exogenous parameter in the prediction of wind speed are investigated. The models are realized in each case and are tested to predict wind speed output in four Scenarios, which include predictions 1 hour, 2 hours, 5 hours and 7 hours ahead of time.

###### i. Case 1

In this case, the training for the LSTM model is made of only one feature set of historical HWS data measured at height 50m for training and target.

###### ii. Case 2

This is the first case with exogenous inputs. The input includes HWS in addition to temperatures measured at 10m and 2m. The target  $y_t$  is the HWS used in Case 1.

###### iii. Case 3

This case utilizes the wind speed and the atmospheric pressure as input. The target sample corresponds to the hourly measured wind speed data.

###### iv. Case 4

In this case, the LSTM model utilizes wind speed, temperatures at 10m and 2m, and atmospheric pressure as input. It includes all features being considered in this study and the target is HWS.

###### v. Case 5

The input in this case include the temperature at two levels and pressure only. The target is the measured HWS.

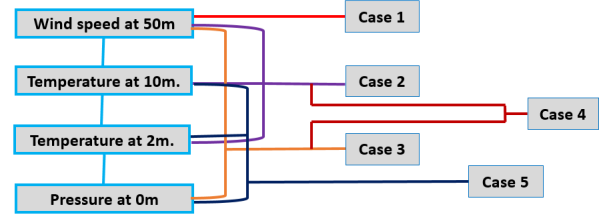


Figure 8. Description of the training parameters and input features

#### 4.2 Forecast Error Matrix

The forecast error for wind speed defined as the difference between the measured and the predicted wind speed values. The metrics for the evaluation of the forecast error include mean absolute error (MAE), root-mean-squared error (RMSE), mean squared error (MSE), and symmetric mean absolute error. In this paper, only MAE and MSE are used as the error metrics and defined in equations (15,16).

$$MAE = \frac{\sum_{t=1}^N e_t^2}{N} \quad (15)$$

$$MAE = \frac{\sum_{t=1}^N e_t}{N} \quad (16)$$

where  $e_t$  is the absolute difference between the forecasted and the measured values of the wind speed for the testing period. The MSE computed to measure the level of deviation between the forecasted and the measured wind speed recorded. The forecast errors are desired to be less than 0.5m/s, otherwise the training is repeated until a better model is achieved and then the final testing is done. Throughout the rest of this article, measured wind speed is used interchangeably with real speed or actual wind speed and predicted value is interchangeably used as forecasted value.

#### 5. RESULTS

The results of all the cases are presented in this section. The training MSE and prediction scatter plots for all cases are shown in Fig. 9 to Fig. 11. The MSE values decrease as the Epochs number increases and becomes constant (~0.0005) after 40 Epochs. Fig. 9 shows the

training MSE plot recorded in Case 1 while Fig. 10 depicts the scatter plots between the actual and the predicted wind speed values. The scatter plots show that both values are closely matched confirming the effectiveness of the model. Also, Fig. 11 shows the training MSE of Cases 2 to 4. It is evident that the MSE values for Cases 2, 3 and 4 are approximately equal but are far different from that for Case 5. Moreover, the MSE becomes constant at the 15th epoch for the first four Cases. However, the MSE in Case 5 attained the 40<sup>th</sup> epoch before approaching a constant value and are larger than the values in the previous Cases (Cases 1-4).

Also, Fig. 12 (a-d) show the correlation scatter diagrams of each model in all Cases. The scatter plots tell at a glance the performance of each model. Accuracy of the forecast could be interpreted from the correlation and regression of each plot. Furthermore, correlation coefficients of all Cases and Scenarios are given in Tables 1 to 5. Also, Table 6 summarizes all training parameters, network algorithms and assumptions for all Cases and Scenarios discussed in this article. Each Case is further analyzed separately in Sections 5.1 to 5.6.

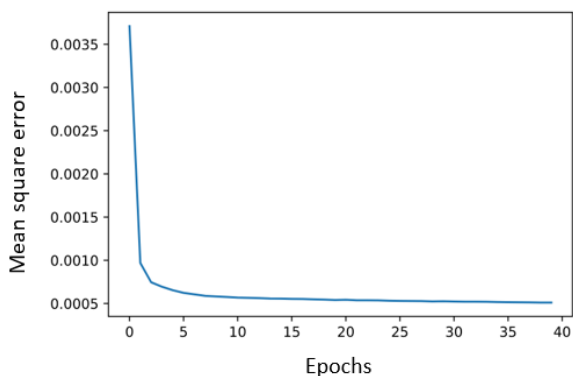


Figure 9. Training MSE, Case 1

#### i. Case 1

In Case 1, the model is trained with hourly mean wind speed only, made of 100, 000 data points for one hour ahead predictions. The training performance of this case is shown in Figs. 9 and 10. The predictions are obtained for four different scenarios including 1, 2, 5 and 7 hours ahead of time. The error metrics obtained from the testing of this model is presented Table 1, where MSE has higher magnitudes compared to MAE values. However, these values tend to decrease as the prediction duration increases from 1 hour to 7 hours ahead.

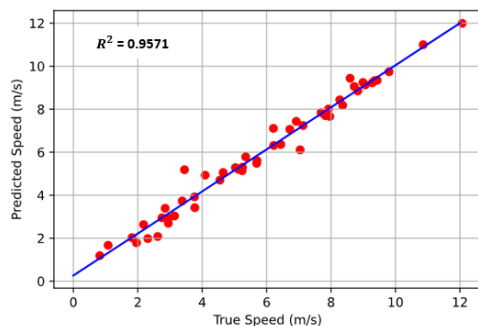


Figure 10. Correlation plots between measured and predicted wind speeds, Case 1

Table 1. Errors and correlation coefficients in case 1 for all scenarios.

Hours ahead	MSE	MAE	Correlation Coefficient
1	0.2140	0.2992	0.9571
2	0.2013	0.2967	0.9704
5	0.1689	0.2850	0.9752
7	0.2042	0.2972	0.9650

Fig. 13(a) compares the measured and the predicted wind speed values for one hour ahead. The predicted values are in close agreement with the measured values with few exceptions. Overall, the predicted values follow the trend of the measured wind speed values which is a strong justification of the accuracy of the model. The same model used to predict wind speed values 2, 5, and 7 hours' head (Figs. 13(b), 13(c), and 13(d)). It is observed from these Figures that the model is capable of predicting the wind speed values accurately up to 7 hours ahead of time. In all of these cases, except few exceptions, the predicted values followed the trends of measured wind speed values. This demonstrates robustness of the LSTM model in short term predictions. Furthermore, the results demonstrate that model realized by training input data sample for an hour ahead using LSTM method could be used to predict wind speed values up to 7 hours ahead without compromising the accuracy as long as the input is correctly fed into the model.

#### ii. Case 2

In this case, the model is trained with a combination of wind speed and exogenous (temperatures at 10m and 2m) inputs. The resulting MSE and MAE values are summarized in Table 2. Comparing to Case 1, MSE and MAE values are lower in this Case 2. This can be attributed to the effect of temperature on wind speed variations because the uneven heating of the earth surface with time causes the wind flow. Again, like in Case 1, MSE values are lower compared to MAE magnitudes and are seen to be decreasing with increasing prediction duration from 1 hour to 7 hours. The MSE value decreases from 0.1961 to 0.1824 (~7%) corresponding to prediction duration of 1 and 7 hours.

Table 2. Errors and correlation coefficients in Case 2 for all scenarios

Hours ahead	MSE	MAE	Correlation Coefficient
1	0.1961	0.2798	0.9650
2	0.1849	0.2755	0.9480
5	0.1614	0.2738	0.9717
7	0.1824	0.2751	0.9471

One hour ahead prediction of wind speed compared with the measured values in Fig. 14(a). The predicted values have an excellent match with the measured ones and more importantly follow the trend of measured values with time. Furthermore, these results seem to be a bit better than those in Case 1 (Fig. 13(a)). Also, Figs. 14(c-d) which are obtained using the same model to predict the wind speeds at 2, 5 and 7 hours ahead show that the model performed a little better than that in Case

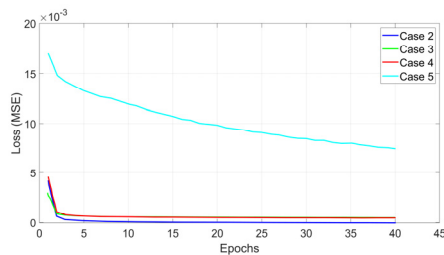
1. Moreover, the error metrics, summarized in Table 2, also indicated that forecast with exogenous parameters, temperatures at 10m and 2m in this, as inputs can perform better in LSTM networks. Additionally, the observation shows that the model realized for an hour ahead prediction can be used predict wind speed for of up to 7 hours ahead accurately.

### iii. Case 3

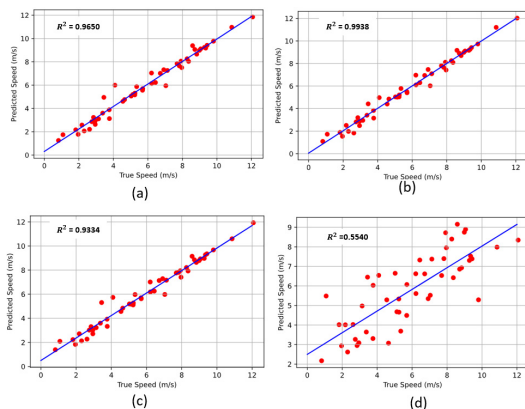
In this Case, the model is trained with wind speed and pressure measured 50m and near ground surface as the inputs. Table 3 summarizes the magnitudes of MSE and MAE obtained in this Case. The values recorded are also below 0.5m/s and thus indicative of high performance. Fig. 11(c) shows that the MSE plot assume similar pattern with those obtained for Case 1 and Case 2. The comparison between the predicted and measured wind speed values (Fig. 15(a)) show very close agreement with each other. Furthermore, a close look at Figs. 13(a) and 14(a) and comparison with Fig. 15(a) confirms better performance in Case 3 compared to Case 1 and Case 2. The predicted values, for other Scenarios for 2, 5, and 7 hours ahead predictions, are also in close agreement with the measured wind speed values (Figs. 15(b), 15(c), and 15(d)). It is observed across all of these plots that this model performed comparatively good like Cases 1 and 2.

**Table 3. Errors and correlation coefficients in Case 3 for all scenarios.**

Hours ahead	MSE	MAE	Correlation Coefficient
1	0.2065	0.2938	0.9938
2	0.2003	0.2940	0.9662
5	0.1649	0.2804	0.9859
7	0.1791	0.2797	0.9749



**Figure 11. Training MSE plots for Cases 2, 3, 4 and 5.**



**Figure 12. Correlation plots between measured and predicted wind speeds for 1 hour ahead prediction for Cases 2-5. (a) Case 2 (b) Case 3, (c) Case 4, (d) Case 5**

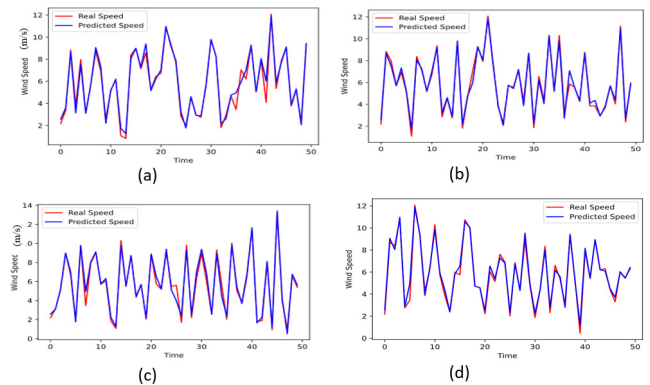
### iv. Case 4

Under this case, the model is trained with combination of all features in Cases 1, 2, and 3 (wind speed, temperatures at 2m and 10m, and atmospheric pressure) and is tested for all Scenarios, as in previous cases. Table 4 provides the model performance in terms of MSE and MAE values. It is evident from the data in Table 4 that for Scenario 1, (one hour ahead predictions) and Scenario 2, (two hours ahead predictions), this model (Case 4) performed better than the Cases 1 and 3. For 5 hours ahead predictions (Scenario 3), the model in Case 4 performed equally well compared to models for Cases 1 to 3. However, MAE values remained almost in the same range as in Cases 1, 2, and 3 for all Scenarios.

**Table 4. Errors and correlation coefficients in Case 4 for all scenarios.**

Hours ahead	MSE	MAE	Correlation Coefficient
1	0.2005	0.2975	0.9334
2	0.1921	0.2946	0.9211
5	0.1659	0.2869	0.9415
7	0.1900	0.2919	0.9214

To further strengthen the predictability of this model, the predicted values under all scenarios are compared with measured hourly mean wind speed values in Figs. 16(a-d). The patterns recorded in all Scenarios showed that the predicted values have an outstanding match with measured values. It is further noted that the predicted values follow the changing trends of measured wind speed values with time for all Scenarios. As a visual observation, trend following nature in the present Case 4 seems to be even better than those in Case 1 (Fig. 13), Case 2 (Fig. 14), and Case 3 (Fig. 15). Hence it can be said that addition of temperature and pressure as input parameters along with the wind speed enhance the model predictability and can be used if measurements are available.



**Figure 13. Wind speed prediction in Case 1. (a) 1 hour ahead (b) 2 hours ahead, (c) 5 hours ahead, (d) 7 hours ahead.**

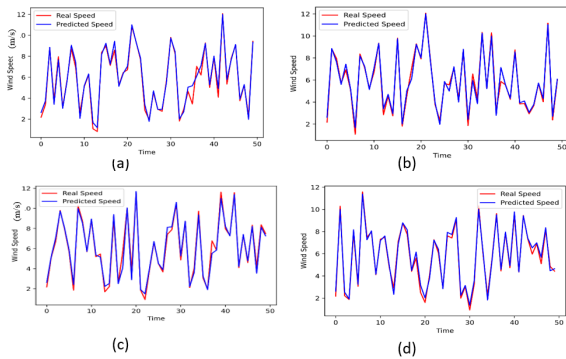
### v. Case 5

This a special case, for the sake of completion, where exogenous features, excluding the wind speed, are used as input to train the model while the target remains the wind speed. The error metrics, summarized in Table 5 for this Case 5, show outrageously high values of MSE and MAE compared to all the previous Cases discussed earlier. Also, the plots of the measured wind speed

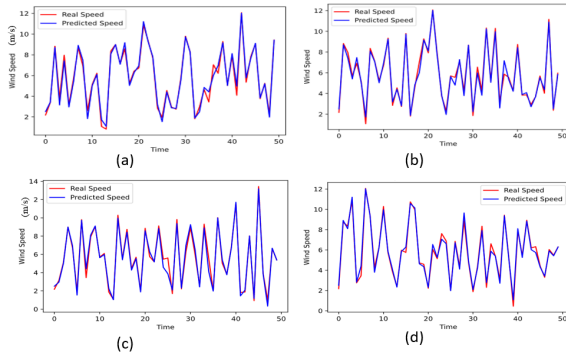
versus predicted show a poor match but still follows the trend of the measured values (Figs. 17(a-d)). It is observed that, this model's predictive accuracy is poor relative to other models presented in sub-sections 5.1 to 5.4. Therefore, for a short-term wind speed prediction using an LSTM, the model should not be trained based on exogenous parameters only as input.

**Table 5. Errors and correlation coefficients in Case 5 for all scenarios.**

Hours ahead	MSE	MAE	Correlation Coefficient
1	3.5318	1.4736	0.5540
2	3.5540	1.4737	0.5436
5	3.6738	1.4900	0.4088
7	3.3094	1.4451	0.4499



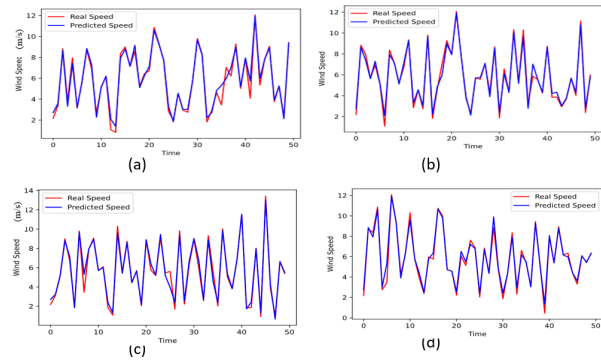
**Figure 14. Wind speed prediction in Case 2, (a) 1 hour ahead (b) 2 hours ahead, (c) 5 hours ahead, (d) 7 hours ahead.**



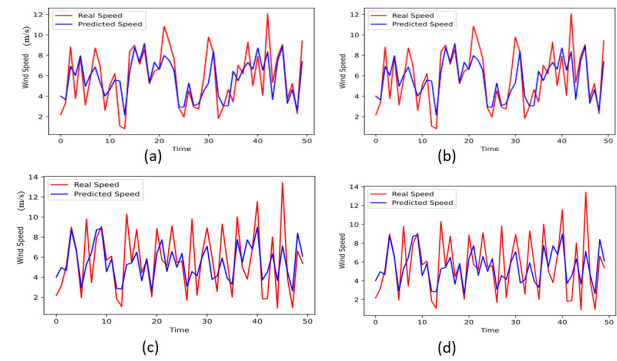
**Figure 15. Wind speed prediction in Case 3, (a) 1 hour ahead (b) 2 hours ahead, (c) 5 hours ahead, (d) 7 hours ahead.**

**Table 6. Training parameters for all cases.**

P	Case 1	Case 2	Case 3	Case 4	Case 5
Hidden layer	1	✓	✓	✓	✓
Nodes	100	✓	✓	✓	✓
Time step (hour)	60	✓	✓	✓	✓
Inputs	Wind speed	Wind speed and temperatures at 10m and 2m	Wind speed and pressure	Wind Speed, temperatures at 10m and 2m, and pressure	Temperatures at 10m and 2m and pressure
Output	Wind speed	✓	✓	✓	✓
Training Optimizer	Adam	✓	✓	✓	✓
Training metrics	MSE	✓	✓	✓	✓
Training epochs	40	✓	✓	✓	✓
Batch size	128	✓	✓	✓	✓



**Figure 16. Wind speed prediction in Case 4, (a) 1 hour ahead (b) 2 hours ahead, (c) 5 hours ahead, (d) 7 hours ahead.**



**Figure 17. Wind speed prediction in Case 5, (a) 1 hour ahead (b) 2 hours ahead, (c) 5 hours ahead, (d) 7 hours ahead.**

In this study, the LSTM method is used for short term (1 hour, 2, 5, and 7 hours ahead) prediction of wind speed. Five models are developed depending on the training input parameters (Case 1: wind speed only; Case 2: wind speed and ambient temperatures at 10m and 2m; Case 3: wind speed and atmospheric pressure; Case 4: wind speed, temperatures and pressure; and lastly Case 5: temperatures and pressure values only). The model performance is evaluated for four scenarios (Scenario 1: 1 hour ahead of time, Scenario 2: 2 hours ahead of time, Scenario 3: 5 hours ahead of time, and Scenario 4: 7 hours ahead of time) using MSE and MAE values. Furthermore, the predicted values are compared with measured wind speed values with time to confirm the trend predictability of proposed models.



## 6. CONCLUDING REMARKS

The study found that LSTM is a powerful tool (based on MSE and MAE values and trend produce-ability) for short term wind speed prediction. Furthermore, it is observed that the accuracy of LSTM improves as the number of training exogenous features increases. However, the algorithm requires that wind speed should be part of the input training parameters in order to enhance its accuracy. Moreover, the accuracy of prediction by LSTM network may be compromised if the prediction time ahead differs from the actual time step used for training the network. The accuracy of prediction deteriorates when only exogenous parameters are used as inputs for training the model.

Finally, the study recommends that LSTM may be used for short term prediction of wind speed with training input parameters as used in Cases 1 to 4. In further studies, an algorithm like the transformer can be chosen to replace the LSTM. Also, ensembling of sequence models may be considered to improve overall learning performance of the model.

## REFERENCES

- [1] Saudi Arabia to tender 2,225MW solar in 2019 and targets 2.7GW CSP in 2030 [online kernel description, <http://www.cspfocus.cn/en/market/detail1244.htm>.
- [2] M. Mehrjoo, M. J. Jozani, M. Pawlak, Wind turbine power curve modeling for reliable power prediction using monotonic regression, *Renewable Energy* 147(2020) 214 – 222.
- [3] P. Jiang, R. Li, H. Li, Multi-objective algorithm for the design of prediction intervals for wind power forecasting model, *Applied Mathematical Modelling* 67130(2019) 101–122.
- [4] P. Du, J. Wang, W. Yang, T. Niu, Multi-step ahead forecasting in electrical power system using a hybrid forecasting system, *Renewable Energy* 122 (2018)533–550.
- [5] F. Wu, X. Liao, Short-term wind power forecast for wind farm base on artificial neural network, in 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), IEEE, 2017, pp. 260–264.
- [6] M. D. Mifsud, R. N. Farrugia, T. Sant, Investigating the influence of MCP uncertainties on the energy storage capacity requirements for offshore wind farms, in International Conference on Offshore Mechanics and Arctic Engineering, Vol.14059353, American Society of Mechanical Engineers, 2019, p. V001T01A030.
- [7] U.T. Salman, F. S. Al-Ismail, and M. Khalid, “Optimal sizing of battery energy storage for grid-connected and isolated wind-penetrated microgrid,” *IEEE Access*, vol. 8, pp. 91 129–91 138, 2020.
- [8] U. T. Salman, M. A. Abdulgalil, O. S. Wasiu, and M. Khalid, “Energy management strategy considering battery efficiency for grid-tied microgrids during summer in the Kingdom of Saudi Arabia,” 2019, doi: 10.1109/ICRERA47325.2019.8997000.
- [9] S. Rehman, U. T. Salman and L. M. Alhems; “Wind farm-battery energy storage assessment in grid-connected microgrids”, *Energy Engineering*, vol. 117, no. 6, pp. 343–365, 2020. [Online]. Available:<http://www.techscience.com/energy/v117n6/40336>
- [10] U. Salman, K. Khan, F. Alismail, and M. Khalid, “Techno-economic assessment and operational planning of wind-battery distributed renewable generation system,” *Sustainability*, vol. 13, no. 12, 2021.[Online]. Available:<https://www.mdpi.com/2071-1050/13/12/6776>.
- [11] G. Chen, J. Shan, D. Y. Li, C. Wang, C. Li, Z. Zhou, X. Wang, Z. Li, J. J.Hao, Research on wind power prediction method based on convolutional neural network and genetic algorithm, in 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), 2019, pp. 3573–3578. doi:10.1109/ISGT-Asia.2019.8880918.
- [12] L. Zhu, H. Shi, M. Ding, T. Gao, Z. Jiang, Wind power prediction based on the chaos theory and the gap neural network, in 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), 2019, pp. 4221–4224. doi:10.1109/ISGT-Asia.2019.8881549.
- [13] S. Rehman, S. A. Khan, L. M. Alhems, “The Effect of Acceleration Coefficients in Particle Swarm Optimization Algorithm with Application to Wind Farm Layout Design,” *FME Transactions*, vol. 48, no. 4, pp. 922–930, 2020, doi: 10.5937/fme2004922R.
- [14] B. Mohammed, O. I. Abdullah, A. I. Al-Tmimi, “Investigation and analysis of wind turbines optimal locations and performance in Iraq,” *FME Transactions*, vol. 48, no. 1, pp. 155–163, 2020, doi: 10.5937/fmet2001155B.
- [15] E. A. Z. Khazem, O. I. Abdullah, and L. A. Sabri, “Steady-state and vibration analysis of a WindPACT 1.5-MW turbine blade,” *FME Transactions*, vol. 47, no. 1, pp. 195–201, 2019, doi: 10.5937/fmet1901195K.
- [16] J. Svorcan, Z. Trivković, T. Ivanov, M. Baltić, and O. Peković, “Multi-objective constrained optimizations of VAWT composite blades based on FEM and PSO,” *FME Transactions*, vol. 47, no. 4, pp. 887–893, 2019, doi: 10.5937/fmet1904887S.
- [17] J. Jung, R. P. Broadwater, Current status and future advances for wind speed and power forecasting, *Renewable and Sustainable Energy Reviews* 31 (2014)762–777.
- [18] M. Marćciukaitis et al.: Non-linear regression model for wind turbine power curve, *Renewable Energy* 113 (2017) 732–741.165
- [19] F. Pelletier, C. Masson, A. Tahan, Wind turbine power curve modeling using artificial neural network, *Renewable Energy* 89 (2016) 207–214.1709
- [20] A. Marvuglia, A. Messineo, monitoring of wind farms’ power curves using machine learning techniques, *Applied Energy* 98 (2012) 574–583.

- [21] Y. Chen et al., "Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history," *Energy Convers. Manag.*, vol. 227, p. 113559, Jan. 2021, doi: 10.1016/j.enconman.2020.113559.
- [22] H. Liu, X. Mi, and Y. Li, "Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM," *Energy Convers. Manag.*, vol. 159, pp. 54–64, Mar. 2018, doi: 10.1016/j.enconman.2018.01.010.
- [23] Y. L. Hu and L. Chen, "A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm," *Energy Convers. Manag.*, vol. 173, pp. 123–142, Oct. 2018, doi: 10.1016/j.enconman.2018.07.070.
- [24] K. U. Jaseena and B. C. Kovoov, "A hybrid wind speed forecasting model using stacked autoencoder and LSTM," *J. Renew. Sustain. Energy*, vol. 12, no. 2, p. 23302, Apr. 2020, doi: 10.1063/1.5139689.
- [25] X. Liao, Z. Liu, and W. Deng, "Short-term wind speed multistep combined forecasting model based on two-stage decomposition and LSTM," *Wind Energy*, 2021, doi: 10.1002/we.2613.
- [26] S. Gangwar, V. Bali, and A. Kumar, "Comparative Analysis of Wind Speed Forecasting Using LSTM and SVM," *ICST Trans. Scalable Inf. Syst.*, vol. 0, no. 0, p. 159407, Jul. 2018, doi: 10.4108/eai.13-7-2018.159407.
- [27] M. Lazaridis, "First principles of meteorology," in *First Principles of Meteorology and Air Pollution*. Springer, 2011, pp. 67–118.
- [28] H. R. Maier, G. C. Dandy, *Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications*, *Environmental modelling & software* 15 (1) (2000) 101–124.
- [29] F. S. Duckworth, J. S. Sandberg, "The effect of cities upon horizontal and vertical temperature gradients," *Bulletin of the American Meteorological Society* 35 (5) (1954) 198–207.
- [30] A. R. Desai, J. A. Austin, V. Bennington, G. A. McKinley, "Stronger winds over 220a large lake in response to weakening air-to-lake temperature gradient," *Nature Geoscience* 2 (12) (2009) 855–858.
- [31] A. S. Alahmed, S. U. Taiwo, M. A. Abido, and M. M. Almuahini, "Intelligent flexible priority list for reconfiguration of microgrid demands using deep neural network," 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), IEEE, 2019, pp. 3490–3495.
- [32] A. H. Elsheikh, S. W. Sharshir, M. Abd Elaziz, A. Kabeel, W. Guilan, and Z. Haiou, "Modeling of solar energy systems using artificial neural network: A comprehensive review," *Solar Energy*, vol. 180, pp. 622–639, 2019.
- [33] M. A. Getahun, S. M. Shitote, and Z. C. A. Gariy, "Artificial neural network based modelling approach for strength prediction of concrete incorporating agri-cultural and construction wastes", *Construction and Building Materials*, vol. 190, pp. 517–525, 2018.
- [34] R. Ghasemiyeh, R. Moghdani, and S. S. Sana, "A hybrid artificial neural network with metaheuristic algorithms for predicting stock price", *Cybernetics and Systems*, vol. 48, no. 4, pp. 365–392, 2017.
- [35] H. Moeeni, H. Bonakdari, "Impact of normalization and input on armax-ann model performance in suspended sediment load prediction," *Water resources management* 32 (3) (2018) 845–863.
- [36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [37] Understanding LSTM Networks. [online Kernel Description. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Accessed: January 2021.
- [38] Y. Liu, C. Yang, K. Huang, and W. Gui, "Non-ferrous metals price forecasting based on vibrational mode decomposition and lstm network," *Knowledge-Based Systems*, vol. 188, p. 105006, 2020.

---

**КРАТКОРОЧНО ПРЕДВИЋАЊЕ БРЗИНЕ  
ВЕТРА БАЗИРАНО НА МРЕЖАМА ЗА  
ДУГОТРАЈНЕ-КРАТКОРОЧНЕ МЕМОРИЈЕ  
(LSTM)**

**У.Т. Салман, Ш. Рехман, Б. Алаводе,  
Л.М. Алхемс.**

Електропривредне компаније, програмери и инвеститори се залажу за већи продор производње енергије ветра и соларне енергије у постојећи енергетски микс. Рад је посебно окренут коришћењу енергије ветра у Саудијској Арабији. Профитабилан развој коришћења енергије ветра подразумева прецизно познавање брзине ветра како у временском тако и у просторном домену. Брзина ветра је параметар са највише прекида и флукуација у поређењу са свим метеоролошким променљивим. Неизвесна природа брзине ветра отежава временско предвиђање снаге ветра. Брзина ветра зависи од метеоролошких фактора као што су притисак, температура и релативна влажност. Предвиђање брзине ветра је од значаја за управљање мрежом, ценом енергије, квалитетом снабдевања енергијом. У раду се даје предлог за краткорочно, вишедимензионално предвиђање брзине ветра коришћењем LSTM. Аутори су развили пет модела обуком мрежа на основу измерених вредности брзине ветра на час у периоду 1980-2019. укључујући егзогене инпуте (температуру и притисак). Утврђено је да је LSTM моћан алат за краткорочно предвиђање брзине ветра. Међутим, LSTM може бити и недовољно прецизан метод када се у обуку мрежа укључе егзогени фактори и предвиђања дужине трајања унапред.