

*Correspondence: Navid Atashfaraz, North Tehran Branch Azad University, Tehran, Iran, n.atashfaraz@iau-tnb.ac.ir

Deep Recurrent Neural Network Models for Forecasting Short-Term Wind Speed

Navid Atashfaraz¹, Mohammad Manthouri², Arash Hosseini³

¹North Tehran Branch Azad University, Tehran, Iran, n.atashfaraz@iau-tnb.ac.ir ²Shahed University, Tehran, Iran, mmanthouri@shahed.ac.ir ³North Tehran Branch Azad University, Tehran, Iran, a.hosseini@iau-tnb.ac.ir

Abstract

Wind speed/power has received increasing attention worldwide due to its renewable nature and environmental friendliness. Wind power capacity is rapidly increasing with the global installed, and the wind industry is growing into a large-scale business. We are looking for wind speed prediction to use wind power better. In this research, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (Simple RNN), and LSTM-GRU in the subset of artificial intelligence algorithms are used to predict wind speed. The data used in this study are related to the 10-minute wind speed data. In the first study on this dataset, we obtained significant results. To compare the deep recurrent models created, we implement four neural network models: Stacked Auto Encoder, Denoising Auto Encoder, Stacked Denoising Auto Encoder, and Feed-Forward presented in the research of others on this dataset. According to the RMSE statistical index, the LSTM network is worth 0.0222 for a short time and performs better than others in this dataset.

Keyword: Deep Learning, LSTM, GRU, RNN, Wind Speed Forecasting

1. Introduction

An urgent issue is to prevent environmental degradation by producing renewable energy. Wind energy can be expressed as fast, clean, flexible, and high-potential energy among various renewable energy sources. A power transmission operator can efficiently control the turbine when a wind farm can accurately predict wind turbine reactions (such as output and load-carrying power). However, predicting wind reactions is challenging due to the airflow's random and nonlinear characteristics and complex relationship with wind turbines. Designing and implementing an intelligent model for the accurate prediction of wind speed and power improves the performance of electrical systems. Although the widespread use of wind power is beneficial, wind farms still face challenges. For example, developing a comprehensive plan for balancing energy supply based on energy demand is a complex issue (Sun, S., Qiao, H., Wei, Y., & Wang, S., 2017). It is necessary to analyze the factors influencing the changes in power and wind speed and effectively evaluate wind energy prediction at different periods. It is necessary to analyze the factors influencing the changes in power and wind speed and effectively evaluate wind energy prediction at different periods. In general, wind speed prediction methods are divided into five categories: In general, wind speed prediction methods are divided into five categories: 1) Persistence model: In this approach, the future wind speed is considered equal to the wind speed at the time of forecasting. The performance of the sustainability method has declined rapidly when the horizon of the forecast time increases, so this model is reliable only for ultra-short-term goals. 2) Physical Methods: This approach requires information such as temperature, pressure, obstacles, and roughness to predict wind speed. These include numerical weather prediction (NWP). This approach was first introduced by (McCarthy, E. F., 1997). 3) Statistical Methods: Find math correlations between wind time series data. 4) Artificial intelligence (A.I.) methods: including artificial neural networks (ANN) that use two feed-forward neural network architectures to estimate wind speed (Philippopoulos, K., & Deligiorgi, D., 2012) and use wavelet neural network for wind speed forecasting (Liu, H., Tian, H. Q., et al., 2013), support vectors regression (SVR) (Yang, L., He, M., Zhang, J., & Vittal, V., 2015) and fuzzy methods (Eseye, A. T., Zhang, J., et al., 2017, March) led to new methods for predicting wind speeds. The advantage of these methods is that to predict future data without any mathematical model predefined. ANNs represent a complex nonlinear relationship for the task of approximating a function. These models learn the relationship between inputs and outputs by non-statistical methods. In the relevant literature, ANNs use different weather variables to predict different weather variables and provide satisfactory results in comparison with traditional algorithms (Philippopoulos, K., & Deligiorgi, D., 2012). 5)Hybrid models: These models are based on a combination of several methodologies to predict wind speed. The purpose of using such models is to take advantage of each approach's features to achieve the system's best performance (Wu, Y. K., & Hong, J. S., 2007).

Neural network architecture presented in recent work can be seen in two categories: shallow architectures and deep learning models. 1) Shallow models include feed-forward neural networks (FFNN) (Lee, D., & Baldick, R., 2013) and recurrent neural networks (RNN) (Cao, Q., Ewing, B. T., & Thompson, M. A., 2012), which have a hidden layer that limits the ability to generalize the neural network. The main reason for this inefficiency is that a classical neural network with more than one hidden layer cannot function efficiently using standard training methods due to the vanishing gradient. Unlike deep architectures, such models can not automatically learn the uncontrollable features of data. Deep learning algorithms such as deep belief network (DBN) neural networks and auto-encoder (A.E.) networks have recently been proposed to address this problem. 2) Deep learning architectures can teach multi-layers of high-capacity computing hidden units. In recent work, (Khodayar, M., & Teshnehlab, M., 2015, September) proposed a stacked auto-encoder (SAE) to predict wind speed. In addition, using the Rough set theory, they introduced neurons into the regression layer, the neural network. The use of the theory of the Rough set is applied to deal with uncertainty in wind data. Experimental results show that SAE has better generalizability than external networks, and there is no need for boring manual engineering approaches to select features. Reference (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017) provided a deep learning architecture using the SAE neural network Stacked Denoising Auto Encoder (SDAE) to predict wind speed. In order to improve prediction accuracy, Rough set theory was included in the proposed deep learning model to extend the development of SAE and SDAE. The DAE proposal has been corrupted to rebuild data. The DAE is capable of rebuilding data from input from corrupted data. This work introduces four architectures: SAE, SDAE, Rough stacked Auto Encoder (RSAE), and Rough Denoising stacked Auto Encoder (RSDAE). Experimental results show that the RSDAE has better predictive power than the low-power networks and the other three architectures introduced in this paper. Reference (Hu, Y. L., & Chen, L., 2018) presents a new nonlinear hybrid model for improving the performance of wind speed prediction—long Short-Term Memory Differential Evolution-Hysteretic Extreme Learning Machine (LSTM-DE-HELM).

In recent years hybrid models have been operating well on the issue of wind speed prediction. Reference (Liu, H., Mi, X., & Li, Y., 2018) provided a new hybrid model that combines wavelet transformation and two types of recurrent neural networks to predict wind speed. In the proposed model, wavelet transformation is applied to several sub-layers to analyze raw wind data. The LSTM, a deep learning algorithm, predicts low frequencies below the layers. Elman neural network, a classic recursive neural network, predicts the high frequencies below the underlying layers. Experimental results show that the LSTM is capable of nonlinear solid processing and is suitable for wind speed prediction. Elman also has good performance in memory and processing of nonlinear data. Reference (Khodayar, M., Wang, J., & Manthouri, M., 2018) presented a hybrid approach for predicting wind speed based on deep learning, rough set theory, and fuzzy set theory.

A new interval probability distribution learning (IPDL) model has been proposed for learning the nonlinear characteristics of the time series data. The IPDL model is presented as a Richard Boltzmann machine (RBM) graphical learning method and rough set theory to capture the unwanted control features of the input series. Also, the interval deep belief network (IDBN) has been proposed to predict wind speed. Empirical results indicate a significant improvement in the proposed IPDL model and its new learning algorithm. They compared with recent shallow and deep architecture, including DBN and hybrid methods. In addition, the effect of the proposed method in managing uncertainty in data improves the performance of this model. Reference (Qin, Y., Li, K., et al., 2019) provided prediction training based on a prediction of wind turbine signals. The proposed model includes a convolutional neural network (CNN), an LSTM network, and multi-task learning ideas in a frame of wind signals. In this method, CNN networks are used to exploit the spatial location of wind power. Also, LSTM is used to teach dynamic wind characteristics. Simulation results show that using a multivariate learning method to predict energy simultaneously with a neural network method reduces the complexity of the method. Therefore, the proposed model has shown good performance in predicting wind speed in the short term. Reference (Vinothkumar, T., & Deeba, K., 2020) has also worked on wind speed prediction whit recurrent neural networks and support vector machines. Reference (Kushwah, A. K., & Wadhvani, R., 2019) worked on wind speed prediction with linear and nonlinear statistical models, offering a powerful hybrid model. Reference (Kenarang, A., Farahani, M., & Manthouri, M., 2022) worked on BiGRU with the attention mechanism and CapsNet (BiGRUACaps) method. The GRU network outperforms LSTM because of fewer gates and, therefore, fewer parameters.

The first 660 kW pilot wind turbine in the LUTAK region, which shows good anomaly measurement, was commissioned in May 2006 with the assistance of Iran's New Energy Organization. Compared to similar turbines elsewhere in the country, the extraordinary situation with all its hassles generates over nine hundred thousand kWh (900 MW/h). Also, this is the first research into Deep Learning in the region and this wind turbine. In this work, we implemented LSTM, GRU, and SimpleRNN on this dataset to predict short-term wind speeds and obtained good results. We also implemented stacked auto-encoder neural networks, Denoising auto-encoder, stacked Denoising auto-encoder (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017), and feed-forward presented in the work of others to compare with recurrent neural networks. The reminder is as follows: Section II proposes wind speed prediction, which is explained in this section. Section III explains the study area, data sets, simulations, and results. Also, Section IV is the overall conclusion of the work.

2. Proposed Wind Speed Forecasting Models

This section explains LSTM, GRU, Simple-RNN, and LSTM-GRU and their architecture. We also look at formulas and their relationships and see their architecture in forms.

A. Long Short-Term Memory (LSTM)

LSTM, usually just called "LSTM," are a special kind of RNN, capable of learning long-term dependencies. They were introduced by (Hochreiter, S., & Schmidhuber, J., 1997) and were refined and popularized by many people in the following work and popularized by many people in the following work. LSTM network has a stable and powerful ability to solve long-term and short-term dependency problems. They work tremendously well on many problems and are now widely used. Due to the three gates, i.e., input gates, output gates, and forget gates, the LSTM network can add or remove information to the cell state. Updating the state of the cell and calculating the output of

the LSTM network can be computed as follows:

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \tag{1}$$

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \tag{2}$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \tag{3}$$

$$\tilde{C} = tanh(x_t U^g + h_{t-1} W^g) \tag{4}$$

$$C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C}_t) \tag{5}$$

Here i, f, and o are called the input, forget, and output gates. W is the recurrent connection between the previously hidden layer and the current hidden layer. U is the weight matrix connecting the inputs to the current hidden layer. C[~]is a "candidate" hidden state computed based on the current input and the previously hidden state, and C is the unit's internal memory. LSTM architecture is shown in Figure. 1.



Fig. 1: The architecture of the LSTM model

B. Gated Recurrent Unit (GRU)

GRU is the newer Recurrent Neural network and is similar to an LSTM. Based on the LSTM gating mechanism, GRU, shown in Figure. 2, was first proposed by (Chung, J., Gulcehre, C., Cho, K., & Bengio, Y., 2014). The GRU cell is computed as follows:

$$z_t = \sigma(x_t U^z + h_{t-1} W^z) \tag{6}$$

$$r_t = \sigma(x_t U^r + h_{t-1} W^r \tag{7}$$

$$\widetilde{h}_t = tanh(x_t U^h + (r_t * h_{t-1})W^h)$$
(8)

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}$$
(9)

Here r is a reset gate, and z is an update gate. Intuitively, the reset gate determines

how to combine the new input with the previous memory, and the update gate defines how much of the previous memory to keep around.1) Update Gate: The update gate acts similarly to forget and input gate of an LSTM. It decides what information to throw away and what new information to add.2) Reset Gate: The reset gate is another used to decide how much past information to forget. Moreover, that is a GRU. GRUs have fewer tensor operations; therefore, they train faster than LSTMs. There is not a clear winner which one is better. Researchers and engineers usually try both to determine which works better for their use case.



Fig. 2: The architecture of the GRU model

C. Simple recurrent neural Network (Simple RNN)

This state is one of the recurrent neural networks where the output is to be fed back to the input. RNNs are neural nets that can deal with variable-length sequences (unlike feed-forward nets). They can do this by defining a recurrence relation over time steps which is typically the following formula:

$$h^{(t)} = f(h^{(t-1)}, x^{(t)}; \theta)$$
(10)

It says the current hidden state $h^{(t)}$ is a function f of the previous hidden state $h^{(t-1)}$ and the current input $x^{(t)}$. The theta is the parameter of the function f. The network typically learns to use as a loss summary of the task-relevant aspects of the past sequence of inputs up to t. Simple RNN architecture is shown in Figure. 3.



Fig. 3: Simple RNN architecture

D. Long Short-Term Memory and Gated Recurrent Unit (LATM-GRU)

The deep hybrid model is composed of an input layer, which in the first step is connected to an LSTM layer and, in the next step, to a GRU layer, and performs a prediction step in the final step shown in Figure. 4.



Fig. 4: The architecture of the LSTM-GRU model

3. Simulation and Results

This section presents the data sets, the recurrent neural networks' outputs, and the benchmark networks' outputs analyzed. Four nonlinear models are presented to obtain highly accurate and stable wind speed prediction results based on LSTM, GRU, Simple-RNN, and LSTM-GRU mechanisms.

A. Dataset, Study Area and Simulation's Parameter

This study is focused on the Lutak region from the city of Zabol, one of the cities of

Sistan and Baluchestan province in Iran. This region has located at latitude 30° 46' 23" and longitude 61° 25' 22" with warm and dry weather. This region has the most wind blowing in Iran. In Figure. 5, the location of this region within Iran is marked. This data set includes annual wind speeds. 10-minute intervals for 2006-2010. At this station, the power of 660 kW / h and 660 volts is injected into the 20-kV network by trans-terrestrial power. Fig. 6 shows some wind speed values in the data set. As shown in Figure. 6, wind time series are highly nonlinear and cannot be modeled with linear models; hence, we record the work of forecasting wind speed using deep neural networks that can model nonlinear data and models in the data set. In this process, the data are divided into two categories: training data and testing data; in this study, 80% of the total data belongs to the training, and the remaining 20% as test data to the model has been introduced. The training and testing set is shown in Figure. 7. There are 137000 wind speed values measured in 10-min intervals; therefore, sufficient data is available for training and testing the proposed approach.

We will look into the model and implementation details ahead in the forecast discussion without knowing any details about the weather conditions predicting the wind speed using the pattern it has followed in a certain period. For prediction, we use a Univariate model with one dataset column.tn our work, we use a wind speed maximum 40-meter parameter for short-term prediction.



Fig. 5: Map of Iran, including the case study area



Fig. 6: Original Wind Speed

Azerbaijan Journal of High Performance Computing, 5 (2), 2022



Fig. 7: Training and Testing for Forecasting

B. Evaluation Criteria

The root means square error (RMSE) and means absolute error (MAE) are employed as two evaluation metrics:

$$RMSE = \sqrt{\frac{1}{M} \sum_{n=1}^{M} e(n)^2}$$
(11)

Furthermore, the MAE is expressed as:

$$MAE = \frac{1}{M} \sum_{n=1}^{M} |e(n)|$$
 (12)

Here, e(n) = t(n) - y(n), and M is the number of samples in the testing set. t(n) and y(n) are the desired output and the actual output of the models for the nth sample, respectively.

C. Numerical Result and Comparison for Prediction

In this study, the wind speed prediction model has been compared for the short term in order to evaluate general abilities. This work compares the performance of the LSTM, GRU, SimpleRNN, and LSTM-GRU algorithms with recent deep and shallow models as a benchmark for predicting Ultra-short-term and short-term wind speeds. Tables I and II show that the RMSE and MAE criteria are used to predict wind speed. Performance of our approach compared with deep and shallow methods. RMSE generally increases with predicted horizons. The SDAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017) and DAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017) because noise cancellation is used in this method, and the noise factor applied to the RMSE of data sets is improved and compared to SDAE and DAE, the DSAE method is better Because it is stacked and deeper. The LSTM model works better than our other recursive models, which include GRU, SimpleRNN, and LSTM-GRU, because this model solves the vanishing gradient problem in the first place, as well as memory, compared to other recursive networks.

Comparison of the all-recurrent neural network, with the FFNN as the benchmark, indicates a significant improvement in the RMSE value of this deep return structure

compared to the FFNN. The main reason behind this algorithm's best-known is its memory architecture, which can best represent the time series data, model, and representation. Also, the combined model of the LSTM and GRU neural network, because it uses its memory state in the first layer of LSTM and output as input to the GRU, generates the output of the GRU end, the value of the RMSE criterion for LSTM network is different. As specified in the table, the RMSE of the LSTM network is 0.0222, and the combined network RMSE is 0.0224. It is known that this is very small due to the GRU network. Also, it was tested on this issue: that is, if GRU is in the first layer and produces output LSTM, because of GRU, the RMSE value is different. Comparing LSTM with SDAE as the best auto encoder refinement indicates a significant improvement in the RMSE value of this deep return structure compared to the DSAE. Finally, Figures 8 and 9, 10, 11, 12, and 1 show the actual wind speed and predicted values by all RNNs, FFNN models, and SDAE.

Mathad	Time Steps		
Method	10 MIN	20 MIN	1h
FFNN	0.0310	0.0328	0.0532
SAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0271	0.0305	0.0509
DAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0253	0.0284	0.0487
SDAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0234	0.0262	0.0466
Simple RNN	0.0224	0.0258	0.0453
LSTM-GRU	0.0224	0.0252	0.0445
GRU	0.0223	0.0253	0.0446
LSTM	0.0222	0.0251	0.0445

Table 1: RMSE of forecasting methods for different time horizons

Table 2: MAE of forecasting methods for different time horizons

Method	Time Steps			
	10 MIN	20 MIN	1 h	
FFNN	0.0243	0.0260	0.0411	
SAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0209	0.0238	0.0388	
DAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0189	0.0218	0.0367	
SDAE (Khodayar, M., Kaynak, O., & Khodayar, M. E., 2017)	0.0170	0.0196	0.0346	
Simple RNN	0.0159	0.0192	0.0333	

Azerbaijan Journal of High	Performance	Computing, 5	(2),	2022
----------------------------	-------------	--------------	------	------

Mathad	Time Steps			
Method	10 MIN	20 MIN	1h	
LSTM-GRU	0.0163	0.0181	0.0323	
GRU	0.0159	0.0182	0.0324	
LSTM	0.0156	0.0179	0.0320	



Fig. 8: LSTM Output: 10 min forecasting



Fig. 9: GRU Output: 10 min forecasting



Fig. 10: SimpleRNN Output: 10 min forecasting



Fig. 11: LSTM-GRU Output: 10 min forecasting



Fig. 12: SDAE Output: 10 min forecastingforecasting

Fig. 13: FFNN Output: 10 min forecasting









IV. Conclusions

Accurate wind speed and wind energy prediction are essential in wind farms' operation and risk management. Therefore, in order to evaluate the importance of the subject, in order to evaluate the performance of deep learning models, we examined four deep models in predicting wind speed for Sistan Wind Power Plant. Wind speed, air temperature, humidity, and sunlight were also recorded every 10 minutes during the statistical period (2006-2010). According to the results obtained in this study, the LSTM neural network at Sistan Pilot Wind Power Plant is a relatively efficient model for predicting wind speed using processed data. The LSTM model performs better than our other recursive models, including GRU, SimpleRNN, and LSTM-GRU, as it eliminates the slope and memory loss problem compared to other recursive networks. The most well-known reason for this algorithm is its memory architecture, which can best represent temporal data, models, and representations. We also run three autoregressive neural networks in addition to the feed-forward network presented in the work of others to compare with recurrent neural networks. However, none of them perform well on recurrent neural networks on wind time series data for future work, considering the capacities of the study area in this paper and the fact that this is the first turbine research in this field. LSTM neural network with regression layers can exist and use the attention mechanism to improve wind speed prediction accuracy.

References

Cao, Q., Ewing, B. T., & Thompson, M. A. (2012). Forecasting wind speed with recurrent neural networks. *European Journal of Operational Research, 221*(1), 148-154.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Eseye, A. T., Zhang, J., et al. (2017, March). A double-stage hierarchical ANFIS model for short-term wind power prediction. In 2017 IEEE 2nd International Conference

on Big Data Analysis (ICBDA) (pp. 546-551). IEEE.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Hu, Y. L., & Chen, L. (2018). A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. *Energy conversion and management*, *173*, 123-142.

Kenarang, A., Farahani, M., & Manthouri, M. (2022). BiGRU attention capsule neural network for persian text classification. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.

Khodayar, M., & Teshnehlab, M. (2015, September). Robust deep neural network for wind speed prediction. In *2015 4th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)* (pp. 1-5). IEEE.

Khodayar, M., Kaynak, O., & Khodayar, M. E. (2017). Rough deep neural architecture for short-term wind speed forecasting. *IEEE Transactions on Industrial Informatics*, *13*(6), 2770-2779.

Khodayar, M., Wang, J., & Manthouri, M. (2018). Interval deep generative neural network for wind speed forecasting. *IEEE Transactions on Smart Grid*, *10*(4), 3974-3989.

Kushwah, A. K., & Wadhvani, R. (2019). Performance monitoring of wind turbines using advanced statistical methods. *Sādhanā*, 44(7), 1-11.

Lee, D., & Baldick, R. (2013). Short-term wind power ensemble prediction based on Gaussian processes and neural networks. *IEEE Transactions on Smart Grid, 5*(1), 501-510.

Liu, H., Mi, X., & Li, Y. (2018). An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm. *Renewable energy*, *123*, 694-705.

Liu, H., Tian, H. Q., et al. (2013). Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. *Applied Energy*, *107*, 191-208.

McCarthy, E. F. (1997). *Wind speed forecasting in the central California wind resource area* (No. CONF-970608-PROC.). American Wind Energy Association, Washington, DC (United States).

Philippopoulos, K., & Deligiorgi, D. (2012). Application of artificial neural networks for the spatial estimation of wind speed in a coastal region with complex topography. *Renewable Energy*, *38*(1), 75-82.

Qin, Y., Li, K., et al. (2019). Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. *Applied energy, 236,* 262-272.

Sun, S., Qiao, H., Wei, Y., & Wang, S. (2017). A new dynamic integrated approach for wind speed forecasting. *Applied energy*, *197*, 151-162.

Vinothkumar, T., & Deeba, K. (2020). Hybrid wind speed prediction model based

on recurrent long short-term memory neural network and support vector machine models. *Soft Computing, 24*(7), 5345-5355.

Wu, Y. K., & Hong, J. S. (2007). A literature review of wind forecasting technology in the world. *2007 IEEE Lausanne Power Tech*, 504-509.

Yang, L., He, M., Zhang, J., & Vittal, V. (2015). Support-vector-machine-enhanced markov model for short-term wind power forecast. *IEEE Transactions on Sustainable Energy*, 6(3), 791-799.

Submitted: 01.09.2022 Accepted: 26.10.2022