The Validity of Deep Learning Computational Model for Wind Speed Simulation

Ali Danandeh Mehr[1]
[1] Civil Engineering Department
Antalya Bilim University
Antalya, Turkey
ali.danandeh@antalya.edu.tr

Reza Farhangi[2], Amir Rikhtehgar Ghiasi[2]
[2] Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
rezafarhanghi@gmail.com
agiasi@tabrizu.ac.ir

***Abstract*—Artificial intelligence (AI) approaches have been used successfully in many recent studies showing their high capability of recognizing underlying pattern in complex processes. In this study, we aimed at the use of various AI approaches to model wind approach which may result in improving the stability and reliability of wind power systems. To this end, we investigate K-Nearest Neighbor (KNN), Support Vector Regression (SVR), Bagging Regression (BR), and Long Short Time Memory (LSTM) techniques in order to model and forecast wind speed (WS) at Basel weather station over the lead time of 3 hours. The forecasting accuracy obtained by the techniques was compared in both training and testing data sets. The results showed the LSTM is slightly superior to its counterparts in our case study.**

***Keywords*—wind speed prediction, KNN, SVR, BR, LSTM**

I. Introduction

It is expected that the recent changes in the structure of global energy will have a significant impact on wind energy proportion in future energy [1–3]. Due to the stochastic attributes of wind speed (WS), it typically causes fluctuations in voltage and frequency in power systems [4,5]. Proper energy planning is challenging in grids since the level of uncertainty in wind power generation [6,7]. For proper grid planning, it is essential to pre-estimate the amount of wind energy generated as this will help the grid in making decisions regarding when to either increase or decrease power generation so that precious resources can be saved. WS should be predicted accurately due to its cubic relationship with the power it can generate and as early as possible as accurate WS prediction can reduce the chances of failures and ensure power system security and stability [8,9]. More power is generated at higher WS because it triggers faster rotation of the blades; hence, wind energy generation is dependent on the WS, mass, & kinetic energy [10,11].

The numerical weather models predict WS which these models based on dynamical equations; therefore, they are more appropriate to find out utmost important information about space and time. In recent years, there has been an increasing interest in predicting WS at specific locations. WS can be predicted by statistical models applying historical data; however, the advancements in data science have increased the popularity of statistical models. A multiple data model has been presented by Hoolohan et al. (2018) using Gaussian process regression and experimentally-sourced meteorological data [12]. Another study by Zhao et al. (2018) presented a probabilistic model for WS prediction; this model is based on the use of some parameter-free statistical frameworks and numerical data for weather prediction [13]. A study by Wang et al. (2018) presented a copula method-based model for WS prediction; this model was analyzed using combined data from the local wind farm and meteorological station. Despite numerous data models' capability to achieve good prediction performances, they may sometimes be associated with increased uncertainty and complexity. The increased availability of observational data has enabled the use of AI-based techniques to establish input-output relationships using non-statistical methods. Machine learning (ML) algorithms are rapidly growing in popularity [14–17].

The literature review has reported a noticeable progression in the WS prediction using various versions of ML models. For example, a hybrid prediction model has been developed by Wang et al. (2015). Also, Ak et al. (2015) investigated the potential of Multilayer perceptron (MLP) for interval WS prediction [18]. Furthermore, Wang et al. (2015) proposed a binary optimization algorithms Particle Swarm Optimization (PSO), with Gravitational Search Algorithm (GSA). In addition, Jiang et al. (2017) extracted similar information from adjacent wind turbine generators using grey correlation analysis before building an SVM model for WS prediction. The DL model has recently attracted much application interest recently; as a new branch of ML [19]. Literature suggests that DL models are more efficient in discovering the inherent and hidden features in data sets from the lowest to the highest level compared to the shallow models [20].

A hybrid system has been proposed by Wang et al. (2016) for probabilistic WS prediction based on DL, spine quantile regression, and wavelet transform. Hu et al. (2016) presented a transfer learning-dependent model that was built from a DL

978-0-7381-2405-6/21/$31.00 ©2021 IEEE

network pre-trained on data sourced from various wind farms. A DNN framework has been proposed by Khodayar et al. (2017) for ultrashort-term and short-term prediction of WS. The proposed model was equipped with a stacked autoencoder (SAE) and a stacked denoising autoencoder (SDAE).

In this study, we used four methods of DL and ML in order to WS prediction. K-nearest-neighbors (KNN), Support-Vector-Regression (SVR), Bagging-Regression (BR), and Long-Short-Time-Memory (LSTM). Data sets are normalized. We compare the result of the methods, and consequently, we found that mean of these methods can be used as another method for WS prediction.

The rest of the paper is organized as follows: Data is explained in section II. Methods are discussed in section III. Results IV. Finally, section V concludes the study.

II. Data

Observed data sets from Basel weather station (47.54694 Lat and 7.5689154 Long), Switzerland, were used in this study. The meteorological data includes hourly measurements of WS as well as temperature (T), relative humidity (H), and mean sea level pressure (P) during the year 2019. Figure 1 shows the observed WS variation versus the antecedent climate data. As it is seen, minimum value and maximum value of WS are 0 and 52. This data has been gathered along 4 different seasons.

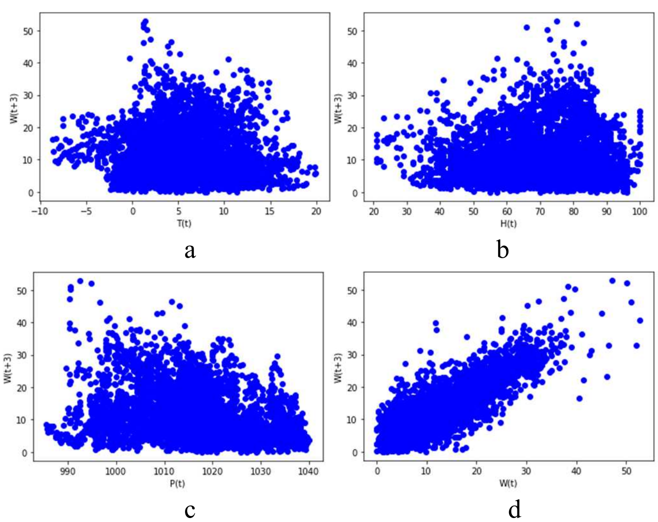


Figure 1. wind-speed in next 3 hours vs. a) temperature b) humidity c) pressure d) present wind-speed

III. Methods

In this study, the aim is to develop 3-hour ahead wind speed forecasting models using climate variables. First, 80 percent of data have been separated. This data is used as training data, and the rest of 20 percent data is used as test data to evaluate our model. The evaluation of these models is considered by the Mean Square Error (MSE), which the MSE is obtained through the equation (1):

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

In this equation, n is number of samples, and y_i is actual label of data, and $\hat{y}_i$ is predicted label of samples. The lower the mean squared error, the better the model used. The data is divided into three parts, 60% as train, 20% as validation, 20% as test.

A. *KNN model*

This model is one of the ML models, which has remarkable accuracy for this task. It's algorithm uses k number points of data and does the regression problem. In order to find the best hyperparameters of the model, which has been fitted on the training data and has been evaluated on the validation data. After the investigation of different values for k, it has been realized that the optimal choice for k is 6, and the hyperparameter k has been found. This algorithm uses six nearest points to the chosen point, analyzes the output of these 6 points, and makes an output for chosen point. This algorithm has been fitted on the combination of training data and validation data, and its MSE shows the accuracy of this model on the training data, and then this model has been evaluated on the test data. MSE value of this model on the test data shows the power of this model.

B. *SVR model*

Another method used for this task is SVR, which has acceptable accuracy for this Regression task. In order to achieve the best hyperparameters of this model, this model has been fitted on the training data and has been evaluated on the validation data. In this model, it is necessary to use a regulator as a hyperparameter, and after analysis of different values for this regulator, the best regulator is chosen as C=0.1. By examining this method, optimal values of this hyperparameter C has been found such that the model results in the best accuracy. At last, the model has been fitted on the combination of training and validation data, and the model has been evaluated on the test data. MSE value of model on the test data has been considered to see the power of this model.

C. *Bagging Regression model*

This is one of the Ensemble ML algorithms. This technique is useful for regression because of it's remarkable accuracy on the most data's. this model increases the stability of regression task and results stable values for WS prediction. BR model has been used with KNN subsets as base estimators, which has been implemented in previous steps, and 80 percent of data for each one of these base estimators, where it considerably increases the stability of models in the variance reduction and accuracy improvement. The training samples used for each subset of KNN are chosen by bootstrapping that choses 80 percent of data with replacement. This process omits the challenge of overfitting. Each KNN model uses 80 percent of training data randomly to have different results, and all of these results will be gathered together, and the best value will be chosen as predicted WS. This model has been trained on the training data and has been evaluated by test data. Finally, the MSE value has been calculated on the test data, which has been predicted by this model. The bagging Regression algorithm is shown in fig (2).

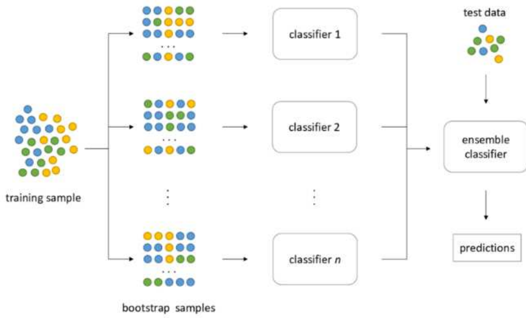


Figure 2. Bagging Regression algorithm

D. *LSTM model*

This model uses for time series data. First of all, the training data and test data are reshaped to (4,1), and 4 is the number of attributes. In the structure of this model, two LSTM layers are implemented, and after each LSTM layer, dropout layers are used to avoid overfitting network. The first LSTM layer has 50 units, and the second one has 25 units. Each one of the dropout layers will drop out 10 percent of data. Then, a flatten layer is used to flat the output of previous layer, and at last, two dense layers are constructed, that the first one has five units, and the second one has 1 unit.

After completing the model structure, it has been compiled. In the compile section, Adam optimizer is chosen as the optimizer function and MSE as the loss function. This network will be fitted on the combination of training and validation data by 250 epochs, and for this task batch size is 32. The best values of hyperparameters of this model is obtained by fitting model on the training data and evaluating model on the validation data. The MSE of this network will be considered on the combination of training and validation data, and also on the test data to evaluate the model. This model's summary is shown in table (1) and the structure of a LSTM network is shown in fig (3).

Table 1. LSTM model summary

Layer(type)	Output(shape)	Param
lstm (LSTM)	(None, 4, 50)	10400
dropout (Dropout)	(None, 4, 50)	0
lstm_1 (LSTM)	(None, 25)	7600
dropout_1 (Dropout)	(None, 25)	0
flatten_1 (Flatten)	(None, 25)	0
dense_9 (Dense)	(None, 5)	130
dense_10 (Dense)	(None, 1)	0

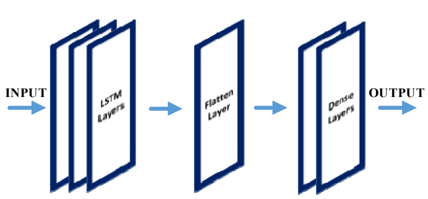


Figure 3. The LSTM neural network structure consists of LSTM layers, a flatten layer, and dense layers

This model has 18,136 parameters that all of them are trainable parameters. After training this model, each one of these parameters will have a specific weight.

E. *Mean model*

Analysis of previous models showed that better results are achievable by the combination of previous models results. This algorithm will use the average of predicted WS from each previous model to predict WS. This model is called the voting model, which uses the same data for train, but it uses different learners to learn, and combines the results of each learner to achieve better and stable result. The learners that have been used for this algorithm are KNN, SVR, BR, LSTM. Results of each one of these models on the test data are gathered and the mean of these values is output of this model.

IV. Results

In this study, WS at Basel is forecasted using both univariate and multivariate time series modeling over the lead time of 3 hours. As previously mentioned, it is necessary to derive WS's joint probability with the independent parameters at current and past times together with that of WS with its antecedent values before calculation. Some values of predicted WS by models and actual value of WS is shown in table (2):

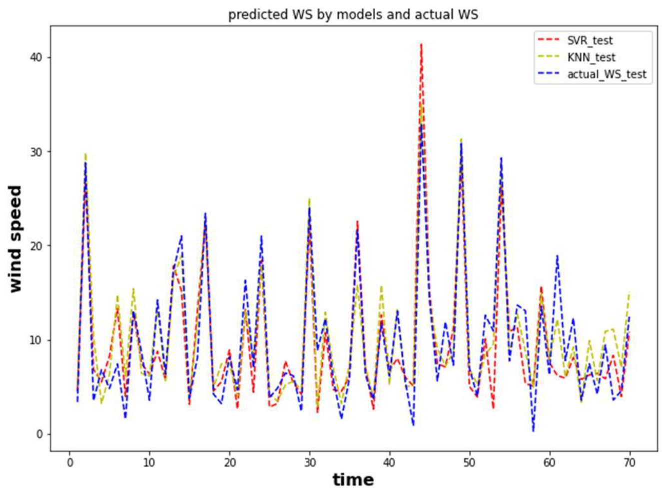
Table 2. actual value and predicted values of WS

WS actual	WS KNN	WS SVR	WS BR	WS LSTM	WS Mean
3.369723	5.105902	4.505571	5.081656	6.950080	5.40
28.74775	29.81256	27.15941	29.67308	28.27261	28.72
3.503945	10.35422	6.837650	10.67352	10.62069	9.61
6.791894	3.213513	5.493486	3.348953	5.129228	4.29
4.802594	6.279059	8.259749	6.811372	6.590673	6.98
…	…	…	…	…	…
3.473431	4.152016	4.681220	4.811478	4.259640	4.47
14.099080	8.907576	10.851933	10.652269	9.499921	9.96
2.677287	4.431412	2.617350	4.273949	3.951381	3.81
9.448753	8.678595	12.637446	9.160078	9.818278	10.06
1.598400	3.079086	2.526638	3.284064	5.207019	3.51

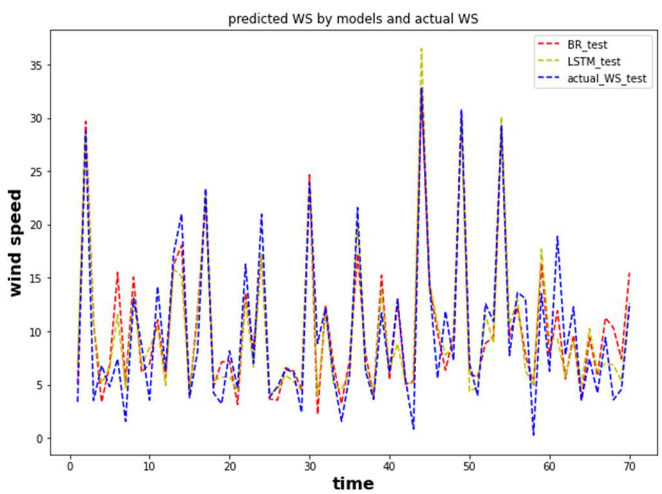
As it is seen in table (2), the mean model seems to be better than others for test data, but for more information, the accuracy of these models will be investigated by MSE.

Considering these values, it is clear that the predicted WS by mean model has better result, because the other models predict WS for different samples of test data, and some of these predicted values are bigger than it's actual value and some other predicted values are smaller than it's actual value. So mean model calculates the mean of these results. It is obviously clear that the predicted WS values by mean model is more stable than the others.

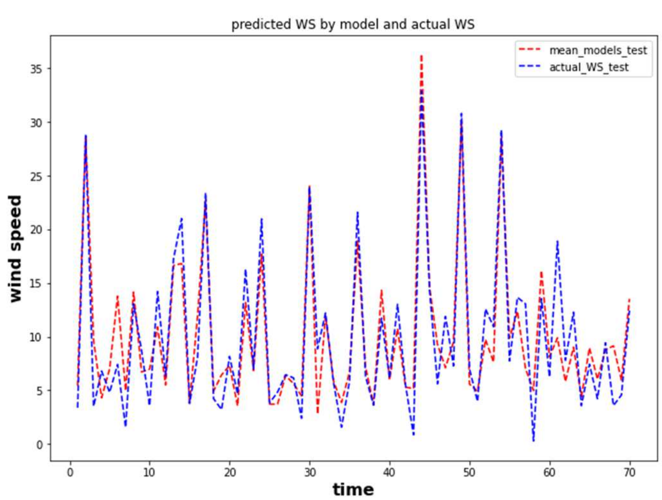
The predicted WS of models on some samples from test data and actual values of WS are shown in Fig (4).



a



b



c

Figure 4. Predicted WS models and observed WS, a) KNN and SVR b) BR and LSTM c) Mean

In table (3), RMSE values for each model on the training data are shown as:

Table 3. Models performance on the training data

Model	KNN	SVR	BR	LSTM	Mean
RMSE(kh/h)	3.338	4.256	3.363	3.576	3.470

In table (4), RMSE values for each model on the test data are shown as:

Table 4. Models performance on the test data

Model	KNN	SVR	BR	LSTM	Mean
RMSE(kh/h)	3.736	3.824	3.677	3.661	3.559

As shown in table (3) and table (4), evaluated criterion for these models is RMSE. In Table (3), the accuracy of each model is examined on the training data. Among these methods, KNN has performed better on the training data, and it has less RMSE value than the other models. SVR model has the worst RMSE value. On the other hand, the mean model has shown an appropriate MSE value.

In Table (4), the accuracy of each model is evaluated on the test data. As it is clear, the LSTM model has a good performance on these data among four main methods. Regarding the LSTM model as a DL model, the constructed network by LSTM has different results in each implementation of this model on data. Consequently, computed RMSE from the implementation of this model on data is changeable between 13 and 14. Finally, by investigating the mean model, it is found that this model has the best performance on test data, and its RMSE is less than the others, but this model has more complexity than the others. Hence, when it is needed to have fast calculations, it is better to use the LSTM model. But the mean model is used when the best performance is needed in order to have the predicted values for WS as possible as it is near to the it's actual value.

According to these RMSE values that are found for each one of these models, the actual WS of each moment could be predicted by having 3 hours' time lag of all needed attributes. This predicted value for WS, eventually will have an average error of [3.559, 3.824] km/h. Considering the min and max values of WS that are 0 and 52, the predicted value for WS using these models is close enough to it's actual value. So these prediction algorithms can be used for power generation systems by wind turbines. Having a vision for wind speed in the coming hours is a critical issue for power generation systems using wind turbines.

V. CONCLUSION

We modeled WS by KNN, SVR, BR, LSTM, and obtained ensemble results. The results showed that the KNN-based model has the best performance for the training data, however the ensemble mean shows better forecasts for the testing data set. The ensemble mean model is complex, so when we want to predict WS as fast as it is possible, we could use LSTM that has better performance than the others. The mean model on testing data has RMSE value equivalent to approximately 3.56, and LSTM on the testing data has RMSE value equivalent to approximately 3.66 km/h, but in each run of the LSTM model, this value is changeable as the model is undeterministic.

Because, having a vision for wind speed in the coming hours is a critical issue for power generation systems using wind turbines, WS should be predicted more accurately. So, this field should be checked more extensively and some feature works are suggested in follow.

For the feature works, it is suggested to consider the lstm model, and make some changes in the structure of this model to have better predictions. Another issue that is suggested to investigate is the influence of attributes. In order to do this investigation, the effect of each attribute at different time lags on the WS of the moment to be predicted should be investigated.

REFERENCES

[1] Jung C, Taubert D, Schindler D, “The temporal variability of global wind energy–Long-term trends and inter-annual variability,” *Energy Conversion and Management 2019*;188:462–72.

[2] Carvalho D, Rocha A, Gómez-Gesteira M, Santos CS, “Potential impacts of climate change on European wind energy resource under the CMIP5 future climate projections,” *Renewable Energy* 2017;101:29–40.

[3] de Jong P, Barreto TB, Tanajura CAS, Kouloukoui D, Oliveira-Esquerre KP, Kiperstok A, et al, “Estimating the impact of climate change on wind and solar energy in Brazil using a South American regional climate model,” *Renewable Energy* 2019;141:390–401.

[4] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ, “Current methods and advances in forecasting of wind power generation,” *Renewable Energy* 2012;37:1–8.

[5] Kim D, Hur J, “Short-term probabilistic forecasting of wind energy resources using the enhanced ensemble method,” *Energy* 2018;157:211–26.

[6] Ervural BC, Evren R, Delen D, “A multi-objective decision-making approach for sustainable energy investment planning,” *Renewable Energy* 2018;126:387–402.

[7] Yadav A, Pal N, Patra J, Yadav M, “Strategic planning and challenges to the deployment of renewable energy technologies in the world scenario: its impact on global sustainable development,” *Environment, Development and Sustainability* 2018:1–19.

[8] Qian Z, Pei Y, Zareipour H, Chen N, “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications,” *Applied Energy* 2019;235:939–53.

[9] Lydia M, Kumar SS, Selvakumar AI, Kumar GEP, “Wind farm power prediction based on wind speed and power curve models,” *Intelligent and Efficient Electrical Systems, Springer*; 2018, p. 15–24.

[10] Patel MR, “Wind and solar power systems: design, analysis, and operation,” *CRC press*; 2005.

[11] Bernstein J, “Wind Resource Analysis Program 2002,” *Minnesota Department of Commerce*: Saint Paul, MN, USA 2002.

[12] Hoolohan V, Tomlin AS, Cockerill T, “Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data,” *Renewable Energy* 2018. doi:10.1016/j.renene.2018.04.019.

[13] Zhao X, Liu J, Yu D, Chang J, “One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data,” *Energy Conversion and Management* 2018. doi:10.1016/j.enconman.2018.03.030

[14] Noorollahi Y, Jokar MA, Kalhor A, “Using artificial neural networks for temporal and spatial wind speed forecasting in Iran,” *Energy Conversion and Management* 2016;115:17–25.

[15] Santamaría-Bonfil G, Reyes-Ballesteros A, Gershenson C, “Wind speed forecasting for wind farms: A method based on support vector regressio,” *Renewable Energy* 2016;85:790–809.

[16] Liu H, Mi X, Li Y, “An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm,” *Renewable Energy* 2018;123:694–705.

[17] Hu Q, Zhang R, Zhou Y, “Transfer learning for short-term wind speed prediction with deep neural networks,” *Renewable Energy* 2016;85:83–95.

[18] Khosravi A, Koury RNN, Machado L, Pabon JJG, "Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system," *Sustainable Energy Technologies and Assessments* 2018;25:146–60.

[19] Chen Y, Lin Z, Zhao X, Wang G, Gu Y, "Deep learning-based classification of hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2014;7:2094–107.

[20] Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, et al, "A review of deep learning with special emphasis on architectures, applications and recent trends," *Knowledge-Based Systems* 2020:105596.