A Three-Layer Hybrid Model for Wind Power Prediction

Jian Gao
Computer Science & Engineering Department
New York University, New York, NY 10003
Email: jg4631@nyu.edu

Panitarn Chongfuangprinya, Yanzhu Ye, Bo Yang
Energy Solution Laboratory, R&D
Hitachi America Ltd, Santa Clara, CA 95054
Email: {joseph.chong, yanzhu.ye, bo.yang}@hal.hitachi.com

***Abstract*—Accurate wind power prediction (WPP) is important for stable operation of power systems. However, the intermittent nature and high variability of wind causes many challenges. This paper proposes a three-layer WPP model considering the data from historical power measurements and numerical weather prediction (NWP) systems. The first layer uses a linear model to learn the wind power generation equation. The second layer includes several non-linear models to learn the seasonality and the inertia of wind turbines. The third layer uses stacked regression to learn a hybrid combination of predictors in the previous layer. We compared the proposed approach against the state-of-the-art algorithm as well as two neural network models. Experiment results show that our approach has the best performance.**

***Index Terms*—wind power, time series, hybrid model, long-term forecasting, renewable energy.**

I. Introduction

Under the pressure of global warming and environment pollution, renewable energy systems have a rapid growth in the past decades. The increasing integration of wind and solar power generation leads to potential impacts on planning and operations of power systems. Utilities need to maintain the balance of demand and supply to ensure the stability of electric power operation. However, the intermittent nature and high variability of wind power generation causes many challenges to power system operators. A robust and accurate wind power prediction (WPP) system is very essential for economic and stable operation of the electricity markets.

According to several studies in this area [1], [2], WPP approaches can be classified based on the forecasting horizon into three categories [3]:

- Short-term forecasting (minutes to 8 hours-ahead)
- Mid-term forecasting (8 hours to day-ahead)
- Long-term forecasting (multiple-days-ahead)

A long-term wind power forecasting system gives utilities time to maintain system frequency and operating reserve, which benefits bidding in multiple-days-ahead electricity markets.

Wind power forecasting can be classified into physical approaches, statistical approaches, and hybrid approaches [4], [5], [6]. Physical approaches take into account the physical characteristic of the wind power generation process and establish a mathematical model from wind force to electric energy production, where accurate wind strength measurement is required. These approaches are effective in short-term forecasting and useful in dealing with operation problems. Statistical approaches learn the underlying non-linear relation between numerical weather prediction (NWP) wind forecasting and wind power generation through statistical methods. In general, NWP forecasts have major impact on the WPP performance.

Pierre Huyn et al. [7] developed a machine learning model based on support vector machines (SVMs) to forecast day-ahead wind power generation in 15-minute intervals. Shu Fan et al. [8] designed a model using Bayesian clustering and SVM to learn the NWP wind speed forecasting patterns. Bhaskar Kanna and Sri Niwas Singh [9] proposed an adaptive wavelet neural network (AWNN) to learn the mapping from NWP's wind speed and wind direction forecasts to wind power forecasts. Yao Liu et al. [10] provided a short-term wind forecasting model based on discrete wavelet transform and long short-term memory networks (DWT_LSTM). However, due to the large time scale in long-term forecasting and the noisy outputs of weather forecasting systems, short-term methods cannot be directly used. The industry has not established effective approaches for long-term forecasting [11].

This work addresses the problem of long-term wind power prediction. We proposed a three-layer WPP model considering the data from historical power measurements and NWP systems. The first layer learns the physical relation between wind force and wind power generation by modeling the wind generation process of wind turbines. It maps NWP's wind speed and wind direction forecasts to a basic wind power estimation based on a linear model. The second layer consists of several non-linear models that learn the seasonality and the inertia of wind turbines. The third layer is a stacked regression model that forms linear combinations of the predictors from the previous layer. The proposed approach is tested on a public dataset, in which the task is to predict 48-hour ahead hourly wind power generation at 7 wind farms. The prediction accuracy is evaluated by Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). We compared the proposed approach against the state-of-the-art algorithm as well as several neural network models. Experiment results show our approach has the best performance.

This paper is organized as follows: Section II introduces the problem and a brief description of the dataset. Section III provides detailed explanation of the proposed algorithms. Section IV shows the performance evaluation and comparison. Section V concludes the paper.

II. Problem Description

A. The Problem

We study the problem of WPP. The task is to predict 48-hour ahead hourly wind power generation at 7 wind farms. Available information in the training data contains historical power measurements of these wind farms and meteorological forecasts of wind components from the NWP systems at each farm. The actual wind power generation was measured and recorded hourly and denoted by $wp[t]$. The NWP meteorological forecasts are formulated as vectors containing the predicted wind speed and wind direction (ws_k, wd_k), $k = 1, ..., K$, where K is the total number of forecasting records. The NWP forecasts were issued twice a day at time t_k with forecasting horizon of 48 hours ahead.

$$\begin{aligned} ws_k &= \{ws[t] \mid t = t_k + 1, ..., t_k + 48\} \\ wd_k &= \{wd[t] \mid t = t_k + 1, ..., t_k + 48\} \end{aligned} \tag{1}$$

Similarly, the wind power generation is predicted in hourly resolution. The WPP model will learn the mapping from the current NWP wind forecasts and the actual wind power measurement in the past. It can be described by

$$\hat{wp}[t] = f(ws[t], wd[t], wp[t-i], \Theta) \quad i = 1, 2, ... \tag{2}$$

where Θ denotes the model parameters learnt from the existing observations. When new NWP forecasts come, the model forecasts the wind power generation at the corresponding hour.

A major challenge of this problems comes from the unpredictable nature and variability of wind conditions, especially the difference between meteorological wind forecasts and the actual wind condition at specific wind farm locations and altitudes due to microclimate. Another issue is the time-series nature of the data that inherits the long-term dependencies and seasonal effects of wind. Traditional time-series models like Autoregressive Integrated Moving Average (ARIMA) cannot formulate such non-linear relationships and incorporate all these effects. We address those issues by training a hybrid model and extract features from the following three aspects:

1) Meteorological wind forecasts from NWP
2) Environmental influence of season and farm location
3) Historical data of past wind power measurement

For long-term wind power prediction (> 24 hours), features from aspect 1) and 2) show more contribution to the prediction performance. The prediction is made by solving a supervised learning problem, and no specific time-series model was used. Therefore, we need to design handcraft features for historical data to capture the time-series property in data streams.

B. The Data

The data is collected from Global Energy Forecasting Competition 2012 - Wind Forecasting [12]. It consists of NWP wind speed & direction forecasts and actual wind power measurement. All power values have hourly resolution and were normalized between 0 and 1. This enables a scale-free comparison of the forecasting results on various wind farms. The NWP forecast outputs are available twice daily at 00UTC and 12UTC and has forecast horizon of 48 hours ahead. Thus, for each datetime, there are 4 NWP forecasts with different forecast horizons. Figure 1 shows the NWP forecast pattern.

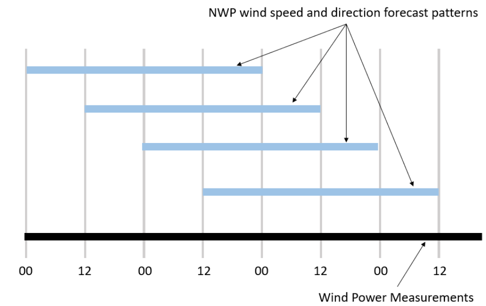

Fig. 1: NWP Forecast Patterns

There are two parts of available data (yyyy-mm-dd-hh):

- Series for the period 2009-07-01-00 to 2010-12-30-12
- Series for the period 2010-12-30-13 to 2012-06-28-12

In the first part, both actual power and wind forecasting are available at all datetimes. In the second part, a set of 48-hour periods with missing power observations are left for prediction. Each part can be split into 156 "$84-hour$ blocks".

III. Proposed Method

We propose a three-layer hybrid model to predict the wind power generation, in which each layer takes the output of previous layer as input data, and produce its own predictions. Figure 2 describes the entire system.

A. First Layer Prediction

Accurate NWP wind forecasting is decisive to train a good WPP system [13]. The first layer model uses NWP data to learn the wind power generation equation of wind turbines.

Wind power is generated from the impact between wind and the blades of wind turbines. The rotating blades slow down the wind and convert it to mechanical energy that drives rotor to generate electricity. The speed-power curve can be split into three regions according to the convergence rate C_p:

1) Constant C_p region, power linearly increases with wind
2) Constant power region, power reaches a controlled limit
3) Region of power shutdown, wind exceeds the upper limit

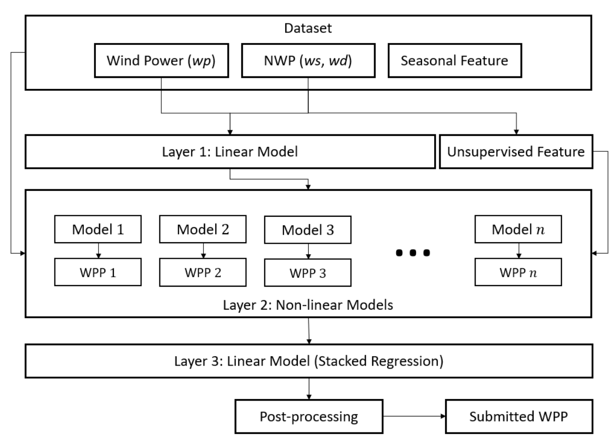

Fig. 2: The Three Layer Wind Power Prediction Framework

There is no introduction to Fig 3 (3 regions of turbine speed control). We may want to include one short sentence in the text body.
Note that your paragraph, table, figure positions are really good already and adding that might impact others.
Make Fig 3 smaller might allow you to add 1 sentence and minimize impact to other stuffs.

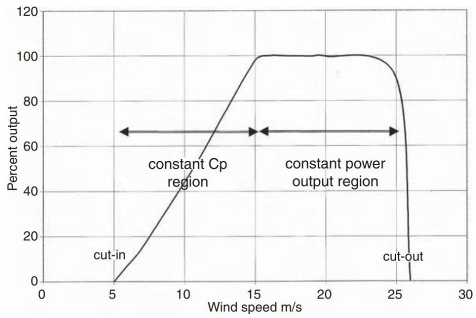

Fig. 3: Three Regions of Turbine speed Control [14]

The electric energy comes from the kinetic energy of wind, which is a function of the wind speed (ws) and air mass (m). The kinetic energy of wind is

$$KE = \frac{1}{2} * m * ws^2 \tag{3}$$

and momentum in the wind is $m * ws$, thus

$$power\ per\ unit\ area = KE * momentum \propto m^2 * ws^3 \tag{4}$$

It indicates the power extracted from wind is proportional to cube of wind speed. Taking into account the impact of wind direction (wd), we use the formula $wp \sim wd * (ws + ws^2 + ws^3)$ to learn this relation, where $*$ indicates the interaction operation in linear regression. The relation between NWP wind forecasts and wind power generation is learnt by a linear model, whose output will be used in the next layer models.

B. Second Layer Predictions

Besides the meteorological forecasts from NWP, there are many other kinds of information essential for wind power prediction. The second layer takes the first layer predictions as a new feature $wind$ and extracts features from other domains. The information captured by these features include seasonal pattern, historical observation, environmental influence, etc. We develop multiple non-linear statistical models to learn those kinds of information and combine them into a hybrid model to obtain better predictive performance.

1) Seasonal Pattern: Wind has a variation according to the season of the year. The underlying reason is that the season may affect the temperature, wind speed, and humidity, which have impact on the wind speed. Since the wind power generation equation is based on wind speed, wind direction, and air density, seasonal features like month, day, and hour can express these factors in a non-explicit way. Here we extract four kinds seasonal features: $year$, $month$, day, $hour$.

2) Time-Series Property: Most machine learning models assume observations to have independent and identically distributed (i.i.d.) distribution. However, the temporal dependence of time-series data violates this assumption. For each timestamp, we integrate the $wind$ feature with its previous and next m-hours observations as a vector $(p_1, \cdots, p_m, n_1, \cdots, n_m)^T$. It learns the inertial behavior of wind turbines by capturing the temporal dependence in time-series observations.

3) Historical Observation: The main purpose of time-series modeling is to forecast the future by studying the past observations. Statistical models use previous wind power observations to generate prediction over the next few hours. The prediction performance falls significantly as the forecasting horizon increases. Statistical models like auto regressive integrated moving average (ARIMA) and neural network models like long short-term memory (LSTM) are only good at small horizons. The partial auto-correlation of wind power observation is small at larger horizons. For each 48-hours missing period, we extract the actual wind power observation before and after it: $h1 = wp[t_k]$, $h2 = wp[t_k - 1]$, $wph1 = wp[t_k]$, $wph49 = wp[t_k + 49]$. These values are the nearest available observations to the missing block and are shared by all 48 predictions in set k.

4) Recursive Forecast: Recursive forecast is another way to use historical data. Ordinary machine learning approaches train independent models for each horizon and perform forecasting in parallel. Recursive forecasting models are trained sequentially so that the predictors at adjacent horizons can help each other. Suppose we want to forecast y using its past observation and feature x, the recursive model would be:

$$y_{t+1} = \alpha_0 + \alpha_1 y_t + \alpha_2 x_{t+1} + \epsilon_{t+1} \tag{5}$$

One step ahead forecast is

$$\hat{y}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 y_t + \hat{\alpha}_2 x_{t+1} \tag{6}$$

Two steps ahead forecast is

$$\hat{y}_{t+2} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{y}_{t+1} + \hat{\alpha}_2 x_{t+2} \tag{7}$$

The forecasting models are trained recursively at each horizon by including the previous output as the historical feature.

5) Environmental Influence: The location of wind farms have influence on wind power generation. Farms that locate close to each other may have similar patterns. We extract unsupervised features to discover the environment influences. Training data shows strong correlation between farm 4, 6, and 7, and weak correlations between other farms. First, we extract unsupervised features $pos12$, $start$, $cluster\_all$, $cluster\_farm$ as in [15] to learn the weak correlations. Then, we add a post-processing procedure to smooth the output of the predictions. Denote (y_4, y_6, y_7) as the raw prediction vector at farm 4, 6, 7, then the vector after smoothing is the weighted average $(y_4, y_6, y_7) * (a_1, a_2, a_3)^T$. The smoothing coefficient matrix (a_1, a_2, a_3) is learnt by linear regression on the validation set. Another smoothing operation is conducted on the time sequence. We use moving average of a 3-hour window size to smooth the predictions. Experiments show the smoothing trick can always reduce the prediction error.

6) Summary: We have designed nine models in the second layer with different features and different type of training data. Some models are trained separately at 7 farms and/or 48 horizons, others are trained on all farms and/or all forecasting horizons. All models are trained by gradient boosting regression [16]. The summery of model variants and the feature used in layer two are listed in Table I, Table II and Table III.

Please change all "fxx" to Italics to make it consistent with Table III

TABLE I. Prediction Models in the Second Layer

feature / data	7 farms, all hours	7 farms, 48 hours	all farms, 48 hours	all farms, all hours
forecast only	f0	N/A	N/A	f10
forecast + history	f1	f12	f11, f15	N/A
environment only	N/A	f6, f7	N/A	N/A
forecast + history + recursive	N/A	N/A	f13	N/A

TABLE II. Features Created in the Proposed Algorithm

Feature	Description	Type	Range
$wind$	layer one prediction	float	[0,1]
$p_1, \cdots, p_m, n_1, \cdots, n_m$	lag of $wind$	float	[0,1]
$month$	month	int	[1,12]
$year$	year	int	[2009,2012]
$hour$	hour in a day	int	[0,23]
day	date difference to July 15	int	[0,195]
$dist$	forecasting horizon	int	[1,48]
set	batch number	float	[1,313]
$h1, h2, wph1, wph49$	historical feature	float	[0,1]
$pos12$	($dist$-1)%12	categorical	$\{0, 1, \cdots, 11\}$
$start$	start hour of forecast	categorical	{0,12}
$cluster_farm$	cluster in one farm	int	[1,6]
$cluster_all$	cluster in all farms	int	[1,24]
wd_c8	categorical wind direction	categorical	$\{0, 1, \cdots, 8\}$
wd_c12	categorical wind direction	categorical	$\{0, 1, \cdots, 11\}$
$ws2, ws3$	wind speed square & cube	float	[0,17]
$r1$	recursive feature	float	[0,1]

C. Third Layer Prediction

Since no individual forecasting approach can capture all the information, we use hybrid method to combine the knowledge learnt by single models. This layer takes the prediction outputs from the previous layer to learn a hybrid ensemble model. We introduce three varieties of hybrid model to learn the ensemble: choose best, simple average, and stacked regression.

1) Choose Best: Model with different types of feature shows different performance at each forecasting horizon. For example in Figure 4, model $f0$ has lower prediction error in the middle while model $f1$ performs much better in the first and last few hours. The reason is, $f1$ includes historical features before and after the missing hours. This type of feature is benefit to the adjacent hours, but may deteriorate the prediction at larger distance. The ensemble model is obtained by combining experts at different forecasting horizons. Based on the prediction performance on validation set, a hybrid model $f01$ is formed by choosing $f1$ at horizons close to the available data and $f0$ at other forecasting horizons.

2) Simple Average: The mean-squared-error (MSE) of an predictor $\hat{\theta}$ with respect to the real value θ is defined as

$$MSE(\hat{\theta}) = E_{\hat{\theta}}[(\hat{\theta} - \theta)^2] = Var_{\hat{\theta}}(\hat{\theta}) + Bias_{\hat{\theta}}(\hat{\theta}, \theta)^2 \quad (8)$$

The first hybrid approach is aimed at reducing the bias by choosing the best model at different horizons. The variance can be simply reduced by averaging the predictions of multiple

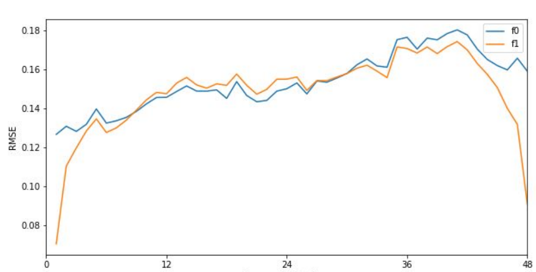

Fig. 4: Error of predictor $f0$, $f1$ over horizons

TABLE III. The Layer 2 Models and Associated Features

Model	Features
$f0$	ws, wind, season, lag of wind, dist
$f1$	ws, wind, season, lag of wind, dist, h1, h2
$f6$	wind, season, pos12, start, cluster_farm, cluster_all
$f7$	wind, season, pos12, start, cluster_all
$f10$	ws, wind, season, lag of wind, dist, farm, wd_c12
$f11$	ws, wind, season, lag of wind, dist, farm, wd_c12, wph1, set
$f12$	ws, wind, season, lag of wind, dist, farm, h1, h2, set
$f13$	ws, wind, season, lag of wind, dist, farm, h1, h2, set, r1
$f15$	ws, wind, season, lag of wind, dist, farm, wd_c12, wph1, wph49, set

models. Table IV shows the averaged prediction has less error than the best single model in layer 2.

3) Stacked Regression: Suppose we have a set of predictors $f_1(x), \cdots, f_K(x)$, instead of selecting a single one from the set, a more accurate predictor can be obtained by combining them. We restrict attention to linear combinations

$$f(x) = \sum_{k=1}^{K} \alpha_k f_k(x) \quad (9)$$

In the simple average approach, $\alpha_k = \frac{1}{K}$ is a constant. Given samples $\{(x_n, y_n), n = 1, \cdots, N\}$, we learn the coefficient α_k to minimize $\sum_{n=1}^{N}(y_n - \sum_{k=1}^{K} \alpha_k f_k(x))^2$. Since the single model predictions are highly correlated, we use ridge regression with regularization term $\sum_k \alpha_k^2 = s$.

IV. Experiment Result

The dataset comes from a Kaggle competition[1], in which hundreds of teams have submitted their prediction results. The performance of the proposed approach is compared with the winner's approach [15] as well as two neural network models Node Decoupled Extended Kalman Filter trained Recurrent Neural Network (NDEKF_RNN) [17] and AWNN [9]. The performance of the prediction models is evaluated by RMSE and MAE. Note that all power values were normalized to 0-1 range, which enables scale-free comparison on multiple farms.

In this work, we proposed a three-layer hybrid model for wind power prediction. Experiment results on the same public dataset show that our model has the best performance compared with several existing approaches (Table IV). It also shows the prediction error for single models in each layer and for the hybrid model. RMSE of the proposed hybrid model is 0.14508, which outperforms the state-of-the-art approach and the neural network models. The performance of persistence model and random guess are also listed for reference.

In this task, the WPP model predicts 48-hour ahead hourly wind power generation at 7 wind farms. Figure 5 plots the prediction error over 48 forecasting horizons, in which $f7$ and $f12$ are the worst and the best model in the second layer. Including historical observations significantly improves the performance at horizons close to the available data, that is, the first and last few hours.

Model performance varies over forecasting horizons. A comparison over horizon period $[13, 36]$ with AWNN [9] and NDEKF_RNN [17] is displayed in Table V. Note that the

[1]https://www.kaggle.com/c/GEF2012-wind-forecasting

Please change all "fxx" to Italics to make it consistent with Table III

TABLE IV. Prediction Error in RMSE for Comparison.

model	note	RMSE
Linear Regression	layer 1	0.17390
f0	layer 2	0.15435
f1	layer 2	0.14997
f6	layer 2	0.15704
f7	layer 2	0.15777
f10	layer 2	0.15297
f11	layer 2	0.15390
f12	**layer 2**	**0.14914**
f13	layer 2	0.14949
f15	layer 2	0.15045
f01	layer 3	0.14873
simple average	layer 3	0.14695
Proposed	**layer 3**	**0.14508**
Leustagos [15]	winner's approach	0.14567
DuckTile [12]	local linear regression	0.14719
Duehee Lee [12]	neural network & Gaussian process	0.15501
AWNN [9]	wavelet neural network	0.15014
NDEKF_RNN [17]	recurrent neural network	0.15347
PERSIST	persistence model	0.35366
Random	random guess	0.46070

horizon period $[13, 36]$ is where the neural network models have the largest improvement over the persistence baseline. Figure 6 shows actual and predicted wind power on farm 2.

Please change to bold text to make it consistent with Table IV

TABLE V. Prediction error over horizon period $[13, 36]$

model	note	RMSE	MAE
Proposed	layer 3	**0.1496**	**0.1055**
AWNN [9]	wavelet neural network	0.1531	0.1172
NDEKF_RNN [17]	recurrent neural network	0.1690	0.1280
PERSIST	persistence model	0.3710	0.2725

V. Conclusion

This paper proposed a three-layer hybrid model for wind power prediction. The useful information comes from two aspects: the NWP wind forecast and the historical wind power measurement. We developed multiple models targeting at different aspects of knowledge. The hybrid model integrates the physical and statistical models specialized for short and long forecasting horizons. Experiment results on a public competition dataset show that the proposed prediction model has the best performance compared with the state-of-the-art approach as well as several neural network models. Future work is to incorporate discrete wavelet transform with LSTM network for short-term forecasting. Another direction is to develop power prediction models for solar energy and produce probabilistic predictions.

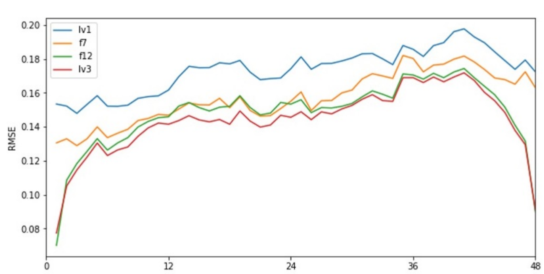

Fig. 5: Prediction error over forecasting horizons

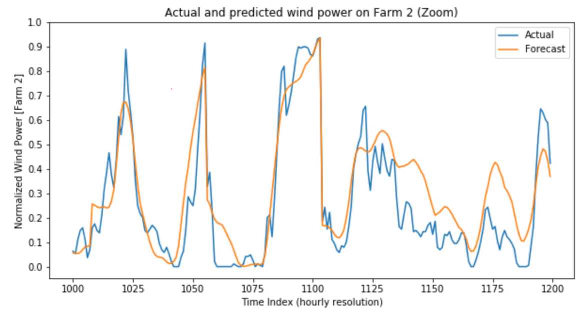

Fig. 6: Actual and Predicted Wind Power

References

[1] L. Lazić, G. Pejanović, M. Živković, and L. Ilić, "Improved wind forecasts for wind power generation using the eta model and mos (model output statistics) method," *Energy*, vol. 73, pp. 567 – 574, 2014.

[2] Q. Chen and K. A. Folly, "Comparison of three methods for short-term wind power forecasting," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–8.

[3] X. Wang, P. Guo, and X. Huang, "A review of wind power forecasting models," *Energy Procedia*, vol. 12, 2011, the Proceedings of International Conference on Smart Grid and Clean Energy Technologies.

[4] L. LI, Y. qian LIU, Y. ping YANG, S. HAN, and Y. mei WANG, "A physical approach of the short-term wind power prediction based on cfd pre-calculated flow fields," *Journal of Hydrodynamics, Ser. B*, vol. 25, no. 1, pp. 56 – 61, 2013.

[5] J. Gao and H. Tembine, "Distributed mean-field-type filters for big data assimilation," *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 1446–1453, 2016.

[6] D. Alencar, C. de Mattos Affonso, R. Oliveira, J. Moya Rodríguez, J. Leite, and J. C. Reston Filho, "Different models for forecasting wind power generation: Case study," *Energies*, vol. 10, p. 1976, 11 2017.

[7] P. Huyn, P. Chongfuangprinya, S. Raju, and C. Venkatraman, "A case-study of wind turbine power forecasting using machine learning techniques," *Asia Power Week Conference*, 2017.

[8] S. Fan, J. R. Liao, R. Yokoyama, L. Chen, and W. Lee, "Forecasting the wind generation using a two-stage network based on meteorological information," *IEEE Transactions on Energy Conversion*, vol. 24, no. 2, pp. 474–482, June 2009.

[9] B. Kanna and S. N. Singh, "Long term wind power forecast using adaptive wavelet neural network," in *2016 IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering (UPCON)*, Dec 2016, pp. 671–676.

[10] Y. Liu, L. Guan, C. Hou, H. Han, Z. Liu, Y. Sun, and M. Zheng, "Wind power short-term prediction based on lstm and discrete wavelet transform," *Applied Sciences*, vol. 9, no. 6, 2019.

[11] S. Han, Y. hui Qiao, J. Yan, Y. qian Liu, L. Li, and Z. Wang, "Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network," *Applied Energy*, vol. 239, pp. 181 – 191, 2019.

[12] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, 2014.

[13] D. Wang, H. Luo, O. Grunder, and Y. Lin, "Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction," *Renewable Energy*, vol. 113, pp. 1345 – 1358, 2017.

[14] M. Patel, *Wind and Solar Power Systems: Design, Analysis, and Operation, Second Edition*. CRC Press, 2005.

[15] L. Silva, "A feature engineering approach to wind power forecasting: Gefcom 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 395 – 401, 2014.

[16] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *CoRR*, vol. abs/1603.02754, 2016.

[17] B. Kanna and S. Niwas Singh, "Improved rnn and awnn based wind power forecast using meteorological inputs," in *1st Annual International Conference on Power, Energy and Electrical Engineering (PEEE 2013)*, 08 2013.