Survival Study on Power Prediction Techniques with Wind Speed Data

Jency W.G, Dr J.E Judith
Research Scholar, Department of Computer Science and Engineering,
Noorul Islam Centre for Higher Education, Thuckalay, Kumaracoil, Tamil Nadu 629180, India
Associate Professor, Department of Computer Science and Engineering,
Noorul Islam Centre for Higher Education,Thuckalay, Kumaracoil, Tamil Nadu 629180, India.
jencywg@gmail.com
judithjegan@gmail.com
Assistant Professor, Department of Computer Science and Engineering,
John Cox Memorial C.S.I Institute of Technology, Kannammoola, Thiruvananthapuram 695011, Kerala, India.

***Abstract*— Wind power has gained large importance in worldwide. An accurate wind power forecast is vital one as wind has difficult and stochastic nature. The energy output of the wind farm is depending on the weather conditions. When output predicted accurately, energy suppliers determine the production of energy sources to avoid overproduction. The prediction of wind power is crucial for the maintenance of wind power units and energy reserve scheduling. Because of cleanness and wide availability, Wind power generation is improved in the large-scale industry. Due to changeable and intermittent features, it is not easy to guarantee stability when accessing the electricity grid. Many researchers carried out their research on wind power prediction. However, it does not provide better performance in terms of accuracy level and time consumption. To address these issues, conventional techniques are reviewed in this paper.**

***Keywords*— Wind power, prediction, cleanness, large-scale industry, intermittent, stability, electricity grid**

I. INTRODUCTION

Wind power is a well-liked renewable energy due to the high efficiency and low pollution. The fast growth of wind power has provided potential energy source in electrical power system. Many prediction methodologies forecasted the wind power depending on the historical meteorological data. Short and long-term wind forecasting are two kinds of wind power forecasting techniques. Short-term forecasting of wind power is described as the power prediction carried out before a day. Long-term wind power prediction is performed prior to year or months. Short-term based forecasting of wind power is more reliable for wind power forecasting.

This article is ordered as follows: Section II describes the review of various wind power prediction techniques, Section III explains study and analysis of existing wind power prediction techniques, Section IV describes the comparison between them. Section V discusses the issues of conventional techniques and conclusion of paper is discussed in Section VI.

II. LITERATURE REVIEW

In [1], a data mining approach was designed with K-means clustering and bagging neural network (NN) for wind power forecasting. However, forecasting accuracy was not improved. In addition, multidimensional clustering problem remained unaddressed. For wind power generation, a new prediction model was introduced in [2] with dilation and erosion (DE) clustering algorithm. However, the designed model was not suitable for the grayscale matrix conversion to cluster based on gray boundary. Depending on the decomposition and ensemble with fuzzy time series, a fuzzy data preprocessing scheme was presented in [3] to lessen the complexity. But, the feature selection rate was not enhanced using fuzzy data preprocessing scheme.

An advanced technique was designed in [4] for improving the wind power ramp prediction performance. The designed technique employed wind power curve to collect the wind power variation. Though prediction accuracy was enhanced, time consumption was not minimized. To predict the wind power, a Long Short-Term Memory-enhanced forget-gate network model (LSTM-EFG) was introduced in [5]. But, the designed method failed to forecast the wind power at certain time. In [6], Long-term wind power forecasting was performed with wind speed data. A machine learning (ML) algorithm was introduced to predict the wind power values. However, ML algorithm was not introduced before wind plant placement in geographical location.

III. WIND POWER PREDICTION TECHNIQUES

With fast growth of wind turbine capability in power systems, wind energy is most significant and efficient renewable energy. Wind power prediction is an essential one to guide grid dispatching and production planning of wind farm in efficient manner. Wind power forecasting deal with uncertainty principles. The intermittency and volatility of wind in diversity has large impact on forecasting accuracy. In power generation, the accurate prediction of wind power is a crucial one. Several factors which affect the predicted wind power namely wind speed variation and climatic conditions.

A. *Data Mining Approach Combining K-Means Clustering with Bagging Neural Network for Short-term Wind Power Forecasting*

The issues in the classification and prediction performances are addressed by utilizing the Data mining techniques. Data mining approach includes K-means clustering and bagging NN to forecast short-term wind power of individual wind turbine. To clean the unreasonable data, regularize training samples and select related variables as input for NN, Data preprocessing was performed in vector space. Through the K-means clustering, the data after preprocessing were gathered to select the training set. Wind power was forecasted through bagging NN to address the instability and over fitting issues of Back Propagation Neural Network (BPNN).

K-means clustering partition the original datasets (DS) into different categories with meteorological conditions and historical power to mitigate impact of diversity of training samples. To compute the similarity among each category and forecasting day, Person correlation coefficient was utilized. A bagging-based approach was combined into BPNN as forecasting engine. Through the Bootstrap sampling technique, the random sampling was performed on training sample to form 'N' subsets. BP algorithm was utilized in each subset and trained on individual network. Through calculating the average value of N networks, the final outcome was obtained.

B. *A novel clustering algorithm based on mathematical morphology for wind power generation prediction*

To ensure the transient stability of power grid, wind power prediction is essential due to the intermittency and randomness of wind power. The marketing power of wind producers are enhanced through the accurate forecasting of wind power. Throughout the world, Wind power has attained greater attention as renewable energy source. Depending on DE clustering algorithm, a new prediction model was introduced for wind power generation. Similar numerical weather prediction (NWP) information to predicted day was chosen by DE clustering algorithm with the basic operations in mathematical morphology. DE clustering algorithm clustered without any supervision.

Depending on DE, a new clustering algorithm determined the number of clusters. To achieve the new matrix, sample DS was processed in DE clustering algorithm that was dilated and eroded. Through categorizing the associated data points, number of clusters was determined with value '1' in matrix. The DS was partitioned into 'k' clusters by computing the distance among every data and cluster centers (i.e., data point mean values). The number of clusters was attained and designed algorithm partition the sample DS. For Medium-range Weather Forecasts (ECMWF), DE clustering-GRNN model efficiency was estimated.

C. *An innovative hybrid system for wind speed forecasting based on fuzzy preprocessing scheme and multi-objective optimization*

Depending on the decomposition and fuzzy time series, a fuzzy data preprocessing scheme was introduced to lessen the complexity and wind speed series chaos. The data preprocessing module, optimization module, forecasting module and evaluation module was utilized in fuzzy data preprocessing scheme. Advanced data preprocessing scheme was introduced in data preprocessing module with the Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) and FTS to observe the wind speed features and enhance the preprocessing efficiency. Data preprocessing was crucial for improving the performance. A data preprocessing scheme does not evaluate the internal characteristics of wind speed data and causes bias in forecasting outcomes. Multi-objective imperialist competitive algorithm (MOICA) was introduced in optimization module to optimize extreme learning machine (ELM) and equivalent hybrid predictor was utilized for wind speed forecasting. MOICA attained higher prediction accuracy and stability.

IV. PERFORMANCE ANALYSIS OF POWER PREDICTION TECHNIQUES WITH WIND SPEED DATA

For conducting the experiment, the wind power prediction techniques are compared with number of data points and number of features. The prediction of wind power generation is enhanced by utilizing the various parameters.

A. *Feature Selection Rate*

Feature selection rate is measured as the ratio of number of features which are correctly selected to total number of features. It is measured in percentage (%). It is formulated as,

$$\text{Feature selection rate} = \frac{\text{Number of features that are correctly selected}}{\text{Total number of features}} * 100 \qquad (1)$$

From (1), the feature selection rate is computed. The method is said to be more efficient when the feature selection rate is higher.

Table 1 Tabulation for Feature Selection Rate

Number of features (Number)	Feature Selection Rate (%)		
	Data mining approach	Prediction model	Fuzzy data preprocessing scheme
10	70	85	79
20	72	87	81
30	75	89	83
40	78	91	85
50	81	93	88
60	79	92	86
70	77	90	84
80	75	87	82
90	78	89	85
100	82	92	88

Table 1 explains the performance analysis of feature selection rate for different methods namely data mining approach, prediction model and fuzzy data preprocessing scheme. Feature selection rate is calculated with respect to the number of features in input DS. For experimental purpose, number of features is varied from 10 to 100.

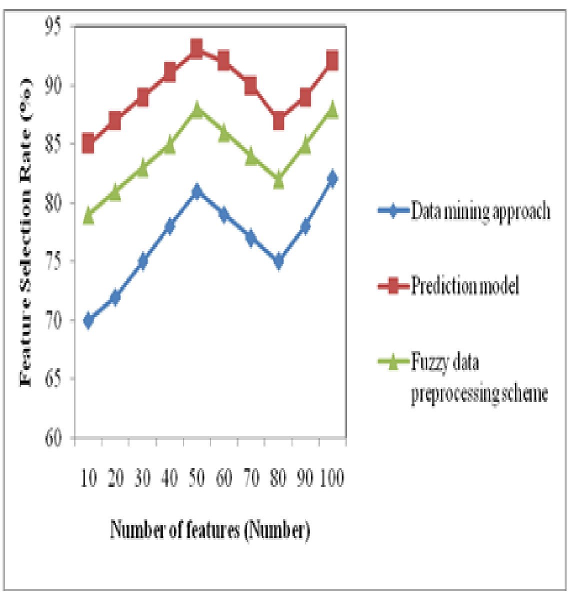

Figure 1 Measurement of Feature Selection Rate

Figure 1 portrays the graphical analysis of feature selection rate versus number of features in the input DS. From figure 1, it is evident that the prediction model provides higher feature selection rate than the data mining approach and fuzzy data preprocessing scheme. This is because of the utilization of DE clustering algorithm in wind power generation. DE clustering algorithm utilizes identical NWP information selection on predicted day by performing basic operations in mathematical morphology. For performing relevant feature selection, DE clustering algorithm clustered without any supervision. This in turn helps to improve the feature selection rate performance. Consequently, feature selection rate of prediction model gets improved by 17% and 6% when compared to data mining approach and fuzzy data preprocessing scheme respectively.

B. *Prediction Accuracy*

Prediction accuracy is measured as the ratio of number of data points are correctly predicted to total number of data points. It is calculated in percentage (%). It is given by,

$$PA = \frac{\textit{Number of data points that are correctly predicted}}{\textit{Total number of data points}} * 100 \qquad (2)$$

From (2), the prediction accuracy is calculated. Higher prediction accuracy, more efficient the method is said to be.

Table 2 Tabulation for Prediction Accuracy

Number of features (Number)	Prediction Accuracy (%)		
	Data mining approach	Prediction model	Fuzzy data preprocessing scheme
10	78	82	92
20	80	85	95
30	77	83	93
40	75	80	90
50	74	78	89
60	71	75	86
70	73	79	88
80	76	83	91
90	79	86	93
100	83	90	96

The performance analysis of prediction accuracy is portrayed in table 2 with different methods namely data mining approach, prediction model and fuzzy data preprocessing scheme. Prediction accuracy is calculated with respect to number of features in input DS. For experimental purpose, number of data points is varied from 10 to 100.

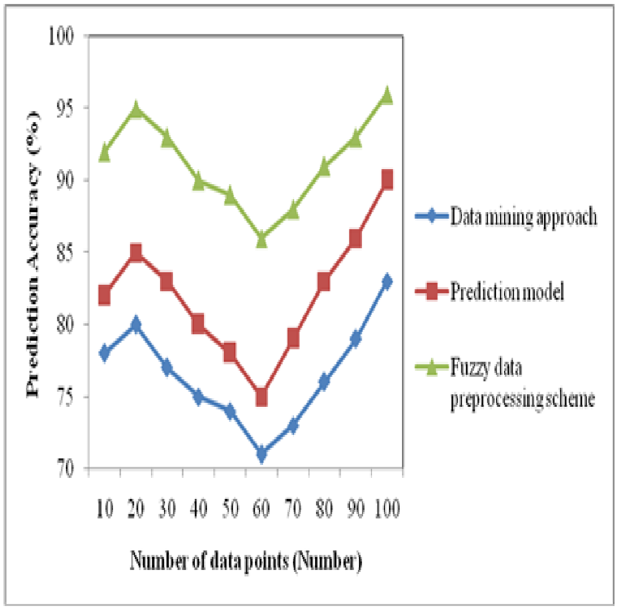

Figure 2 Measurement of Prediction Accuracy

Figure 2 explains the graphical representation of prediction accuracy versus number of data points in the input DS. From figure, it is observed that prediction accuracy using fuzzy data preprocessing scheme is higher than data mining approach and prediction model. This is due to the application of MOICA for optimizing ELM and equivalent hybrid predictor for wind speed forecasting. MOICA achieved higher prediction accuracy and stability. MOICA used decomposition and fuzzy time series to minimize complexity and wind speed series chaos. This in turn improves the prediction accuracy. Therefore, prediction accuracy of fuzzy data preprocessing scheme gets improved by 19% and 11% when compared to data mining approach and prediction model respectively.

C. Prediction Time

Prediction time is measured as the amount of time consumed to predict the wind. It is the difference of starting time and ending time of prediction process. It is calculated in milliseconds. It is formulated as,

$$\textit{Prediction time} = \textit{Ending time} - \textit{starting time of prediction} \quad (3)$$

From (3), prediction time is computed. Minimum prediction time, more efficient the method said is said to be.

Table 3 Tabulation for Prediction Time

Number of features (Number)	Prediction Time (ms)		
	Data mining approach	Prediction model	Fuzzy data preprocessing scheme
10	18	27	34
20	20	29	36
30	23	31	38
40	26	34	40
50	28	37	43
60	30	39	45
70	33	41	47
80	36	44	50
90	38	47	53
100	41	50	55

Table 3 describes the performance analysis of prediction time for different methods namely data mining approach, prediction model and fuzzy data preprocessing scheme. Prediction time is calculated with respect to number of data points in input DS. For experimental purpose, number of data points is varied from 10 to 100. The graphical representation of prediction time results is shown in figure 3.

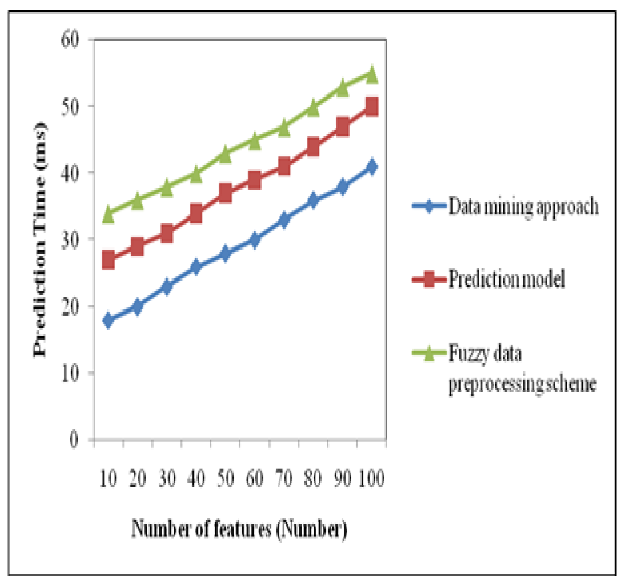

Figure 3 Measurement of prediction time

Figure 3 illustrates the graphical analysis of prediction time versus number of data points in input DS. From figure 3, it is evident that the Data mining approach provides lesser prediction time than prediction model and fuzzy data preprocessing scheme. This is due to the application of K-means clustering to categorize original DSs into different categories based on meteorological conditions. Person correlation coefficient calculates similarity level between every category and forecasting day. A bagging-based approach combined with BPNN to address over fitting issues for minimizing the time consumption. Consequently, prediction time of data mining approach gets reduced by 24% and 35% when compared to prediction model and fuzzy data preprocessing scheme respectively.

V. DISCUSSION AND LIMITATION ON WIND POWER PREDICTION TECHNIQUES

A bagging-based approach was used to address the instability and over fitting problems. BP algorithm was utilized in every subset in individual network. The designed approach reduced the computational complexity. But, the forecasting accuracy was not improved at required level as effective meteorological forecasting was not introduced. An optimal method for back propagated NN was not introduced. The multidimensional clustering issue was not addressed. DE clustering algorithm clustered repeatedly without any supervision. DE clustering-GRNN model was an adaptive one. The clustering algorithm carried out clustering routinely for wind farm applications. The designed model was not appropriate one for grayscale matrix conversion and for clustering data consistent with gray boundary. The prediction error was not minimized when weather pattern was diverse for attaining additional information relationship between wind power and meteorological data. MOICA improves the accuracy and stability to extract uncertain features of wind speed. For wind speed forecasting, Point and interval forecasting mined uncertainty of wind speed via the various options. But, the feature selection rate was not improved by fuzzy data preprocessing scheme.

A. *Related Works*

Based on gated recurrent unit NN, a numerical weather prediction wind speed error correction model was presented in [7] for short-term wind power prediction. However, forecasting time was not reduced. In [8], a new forecasting model was introduced with convolution NN and LightGBM. A new feature sets were employed through examining the features of raw time series data from wind field. However, robustness was not simple to guarantee for abnormal and false data. In addition, model generated by ensemble learning (EL) consumed large amount of resources. A data preprocess strategy was introduced in [9] depending on feature extraction for minimizing fluctuations of wind power generation. A fuzzy set theory selection technique discovers the solution from Pareto front set deriving from optimization phase. But, the multi-objective optimization issues were not resolved in the energy sector.

B. *Future Direction*

Future direction of the work is to perform efficient prediction of wind power by using ML and EL techniques with higher accuracy and lesser time consumption.

VI. CONCLUSION

A comparative study of different wind power prediction techniques is carried out. From the survival study, the feature selection rate was not improved by fuzzy data preprocessing scheme. In addition, prediction error was not minimized for attaining additional information relationship between wind power and meteorological data. The forecasting accuracy was not improved at required level. The wide experiment on conventional techniques evaluates the results of different wind forecasting techniques and discusses its issues. From the result analysis, the research work can be done with the assist of ML and EL techniques for wind power prediction with higher accuracy and lesser time consumption.

REFERENCES

[1] Wenbin Wu and Mugen Peng, “A Data Mining Approach Combining K-Means Clustering With Bagging Neural Network for Short-Term Wind Power Forecasting”, IEEE Internet of Things Journal, Volume 4, Issue 4, August 2017, Pages 976-986

[2] Ying Hao, Lei Dong, Xiaozhong Liao, Jun Liang, Lijie Wang and Bo Wang, “A novel clustering algorithm based on mathematical morphology for wind power generation prediction”, Renewable Energy, Volume 136, 2019, Pages 572-585[3] Chen Li, Zhijie Zhu, Hufang Yang, Ranran Li, “An innovative hybrid system for wind speed forecasting based on fuzzy preprocessing scheme and multi-objective optimization”, Energy, Elsevier, Volume 174, Pages 1219-1237

[3] Tinghui Ouyang, Xiaoming Zha, Liang Qin, Yusen He, Zhenhao Tang, “Prediction of Wind Power Ramp Events Based on Residual Correction”, Renewable Energy, Elsevier, Volume 136, June 2019, Pages 781-792

[4] Min Ding, Hao Zhou, Hua Xie, Min Wu, Yosuke Nakanishi and Ryuichi Yokoyama, “LSTM-EFG for wind power forecasting based on sequential correlation features”, Future Generation Computer Systems, Elsevier, Volume 93, April 2019, Pages 33-42

[5] Halil Demolli, Ahmet Sakir Dokuz, Alper Ecemis, Murat Gokcek, “Wind power forecasting based on daily wind speed data using machine learning algorithms”, Energy Conversion and Management, Elsevier, Volume 198, 2019, Pages 1-12

[6] Min Ding, Hao Zhou, Hua Xie, Min Wu, and Yosuke Nakanishi and Ryuichi Yokoyama, “A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting”, Neurocomputing, Elsevier, July 2019, Pages 1-8

[7] Yun Ju, Guangyu Sun, Quanhe Chen, Min Zhang, Huixian Zhu and Mujeeb Ur Rehman, “A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-Short-Term Wind Power Forecasting”, IEEE Access, Volume 7, Pages 28309 – 28318

[8] Ping Jiang, Ranran Li and Hongmin Li, “Multi-objective algorithm for the design of prediction intervals for wind power forecasting model”, Applied Mathematical Modelling, Elsevier, Volume 67, March 2019, Pages 101-122

[9] Ping Jiang, Ranran Li and Hongmin Li, “Multi-objective algorithm for the design of prediction intervals for wind power forecasting model”, Applied Mathematical Modelling, Elsevier, Volume 67, March 2019, Pages 101-122