A proposed Framework for a Forecasting System of Wind Energy Power Generation

Ahmed A. Abdullah, Ahmed E. Saleh, Mohamed S. Moustafa, Khaled M. Abo-al-Ez

Abstract—Wind energy generation is expected to increase in future electric grids. The generated wind power has an intermittent nature which may affect power system stability and increase the risk of blackouts. Therefore, a prediction system for wind power generation is essential for optimum operation of a power system with a significant share of wind energy conversion systems. In this paper, a hybrid neuro-fuzzy wind power prediction system is proposed. A wireless sensor network (WSN) is used to measure and transmit the required parameters for the prediction model at the operator center. Those parameters are the major factors affecting wind farm output power, namely air temperature, wind speed, air density and air pressure. Considering all these factors will increase the prediction accuracy of the proposed model. The proposed prediction model is designed using fuzzy rules with adaptive network. To decide the optimal number of fuzzy rules, the clustering of the data using modified Fuzzy C-Means (FCM) is used to implement hybrid optimization method.

Index Terms—prediction, Wireless sensor network, SCADA, Neuro fuzzy.

I. INTRODUCTION

Wind is one of the fastest growing energy sources since it is renewable, abundant and pollution-free. The power generated by a wind turbine (WT) generator varies randomly with time due to the variability of wind speed, temperature, and other factors. Uncertainty of wind power and the increase of wind power penetration level in future electric grids will affect power system stability and increase the risk of blackouts. Therefore, forecasting of wind power generation is beneficial for optimum operation of a power system with a significant penetration level of wind energy conversion systems [1-6].

Due to temperature and pressure difference, air density, topography and other factors, wind speed is one of the most difficult meteorological parameters to predict [7]. As a result, the power generated from WT will be difficult to predict. Therefore, the prediction model will inevitably be non-linear and must be more accurate. In recent years, with continuous increase of computer calculation speed, researchers proposed a number of power prediction models based on complex statistics and artificial intelligence techniques as found in [8]-[19], where the prediction models of wind power generation are categorized into direct and indirect models. Direct prediction models use historical information of wind power output as the prediction model's input and the output of the prediction model is the predicted value of wind power generation. Where, indirect prediction models predict wind power generation by predicting wind speed and then use the power curve to convert wind speed into power output.

Generally, statistical models have the advantage that they require no mathematical modeling and use available historical measurements for stochastic approximation between wind predictions and wind power output measurements. However, those models are not suitable for long-term prediction and it is very difficult for prediction model based on statistics to further improve prediction accuracy [7].

The indirect prediction models used the power curve to convert wind speed into power output. This may cause a delay and some error in calculations. The power curve depends on the wind speed not on the efficiency and other mechanical parameters of the wind turbine which could be changed with time. So the power curve should be correct every specified time interval.

Both direct and indirect methods are not complete prediction systems due to their limitations such as collecting, transmitting, and saving data and the display method of the prediction values to operator. The complete prediction system should consist of an efficient way to collect values of real time parameters from the wind turbine (specifically offshore wind turbines) using any communication media then save the collected data to a database. Thereafter, predicting the power generation and display this predicted value to the wind farm operator.

Real-time data collection technology is a challenging task, which requires not only to maximize efficiency, but also to take into account cost, reliability and other factors, to promote the security and stability of the grid. Wind power generation systems are installed in relatively abundant area of wind resources such as in the sea or the valley, with bad environment, complicated topography, transportation inconvenient factor. In order to get a real-time data from the WT, The wireless Sensor Networks (WSNs) has a feasibility to collect a real-time data from WT. A wireless sensor network is a technology for the collection and communication of the correlative Real-time data in wind

Manuscript received March, 2014.
Ahmed A. Abdullah, Communication and Computer Engineering, Delta University / Faculty of Engineering, (e-mail: aaleem82@gmail.com), Damietta, Egypt

Ahmed E. Saleh, Computer Engineering and systems, Mansoura University / Faculty of Engineering, Mansoura, Egypt, (e-mail: aisaleh@yahoo.com).

Mohamed S. Moustafa, Computer Engineering and systems, Mansoura University / Faculty of Engineering, Mansoura, Egypt, (e-mail: msmsksasy@gmail.com).

Khaled M. Abo-al-Ez, Electrical Engineering/ Faculty of Engineering/ Mansoura University, (e-mail: ezzkhalid@mans.edu.eg).
power prediction [20]. Researchers have done many effective jobs in WSN for monitoring the WT parameters [21, 23].

In this paper, a Wireless Sensor Network (WSN) for collecting the real-time data from wind turbine is proposed. Applying WSN in wind power prediction will be feasible [21]. The wireless sensors used in [21] are modified to measure output power, temperature, density, pressure, and wind speed sensors. The values of these parameters are inputs to the prediction model of wind power generation. The historical data are used as inputs to increase accuracy of the predicted wind power value. In order to achieve the highest creditable prediction accuracy, the modeling techniques is designed using the combination of a Fuzzy Classifier with a Temporal Neural Network techniques based on historical data and the real-time data which are vital for the wind power prediction.

This paper is organized as follows. A background for a WT modeling, WSN technology, and fuzzy logic system are firstly presented. Thirdly, the architecture of applying WSN in the proposed wind power prediction framework and data flow in the system is described at second. Finally, the flowchart of building and evaluate the proposed prediction model is designed.

II. BACKGROUND

A. Wind turbine modeling

In this section describes the details of wind turbine modeling. First, wind energy conversion system is discussed briefly. Then, the modeling and major factors affecting wind farm output power are presented in order to be used in the proposed wind power prediction model. Wind farm output power comes from wind power captured by WTs. The output power from ideal and practical wind turbines is discussed in the following section [22].

Essentially, wind energy is kinetic energy, and it can be calculated using equation (1).

\[
E = \frac{1}{2} m v^2 \]  
(1)

Where \( m \) is the air mass (kg) and \( V \) is the air speed in the upstream wind direction at the entrance of rotor blades (m / sec.), \( A \) is the cross-sectional area in \( m^2 \), \( \rho \) is the air density in \( kg/m^3 \), and \( x \) is the thickness of the parcel in m.

Wind power \( P_w \) (in Watts) is the time derivative of kinetic energy of wind using:

\[
P_w = \frac{dE}{dt} = \frac{1}{2} (\rho A) \frac{dx}{dt} v^2 = \frac{1}{2} (\rho A) v^2 \]  
(2)

Actually, the output mechanical power \( P_m \) extracted from WT blades can be obtained using equation (2) multiplied by coefficients as following [23].

\[
P_m = k C_p \frac{1}{2} (\rho A) v^3 \]  
(3)

Where:

\( P_m = \) Mechanical power output, kilowatts
\( C_p = \) Maximum power coefficient, ranging from 0.25 to 0.45, dimension less (theoretical maximum = 0.59)
\( A = \) Rotor swept area, \( ft^2 \) or \( \pi D^2/4 \) (D is the rotor diameter in ft, \( \pi \approx 3.1416 \))
\( V = \) Wind speed, mph
\( k = 0.000133 \) A constant to yield power in kilowatts. (Multiplying the above kilowatt by 1.340 converts it to horsepower [i.e., 1 kW = 1.340 horsepower]).
\( \rho = \) Air density, lb/ft$^3$.

The air density \( \rho \) is calculated by [36]:

\[
\rho = \frac{P}{R_{specific} T} \]  
(4)

Where \( P \) is absolute pressure, \( T \) is absolute temperature (K), and \( R \) is specific gas constant for dry air.

According to equation(3), the extracted mechanical power of a WT generator, is proportional to the air density \( \rho \), and the air speed in the upstream wind direction at the entrance of rotor blades \( V \), where \( \rho \) is determined mainly by air temperature as shown in equation (4) [36].

Therefore, the major factors affecting wind farm output power are air temperature, wind speed, air density and air pressure. These factors are assembled in the proposed model to predict the expected wind power.

B. Wireless Sensor Network (WSN)

Recent years have witnessed an increased interest in the use of WSN innumerous applications which is designed and programmed depending significantly on the application, and it must consider factors such as the environment, the application’s design objectives, cost, power consumption, hardware, and system constraints to match the needs of the target applications before they are released to the field [24, 25]. These sensors are small, with limited processing and computing resources, and they are inexpensive compared to traditional sensors. These sensors can sense, measure, and gather information from the environment and, based on some local decision process, they encrypt the data in order to prevent eavesdropping and security then transmit the encrypted sensed data to server using equipped wireless interfaces which they can communicate with one another to form a network.

Generally, each sensor node consists of three sub-units as shown in figure (1): first, a sensing unit is used to acquire the target events or interesting data; second, a processing unit equipped with limited memory is used to manage the acquired data; third, a communication unit, usually a radio transceiver, is used to exchange information between nodes [26]. One of the most important WSN applications is monitoring the WT resources which are the fastest growing sources for power production in the world today and there is a constant need to predict the output power [27].

C. Fuzzy logic Systems

Fuzzy logic is the process of formulating the mapping from a given input to an output. The mapping then provides a basis from which decisions can be made, or patterns discerned. There are two types of fuzzy inference systems that can be implemented in Fuzzy Logic Toolbox: Mamdani-type and Sugenotype. These two types of inference systems vary somewhat in the way outputs are determined [28, 29, 30]. Because of its multidisciplinary
nature, fuzzy inference systems are associated with a number of names, such as fuzzy-rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and simply (and ambiguously) fuzzy systems. Figure (2) shows the process flow of fuzzy logic system.

The mamdani’s method was among the first control systems built using fuzzy set theory [29]. Mamdani-type inference, as defined for Fuzzy Logic, expects the output membership functions to be fuzzy sets. After the aggregation process, there is a fuzzy set for each output variable that needs defuzzification.

It is possible, and in many cases much more efficient, to use a single spike as the output membership functions rather than a distributed fuzzy set. This type of output is sometimes known as a singleton output membership function, and it can be thought of as a pre-defuzzified fuzzy set. It enhances the efficiency of the defuzzification process because it greatly simplifies the computation required by the more general Mamdani method, which finds the centroid of a two-dimensional function. Rather than integrating across the two-dimensional function to find the centroid, you use the weighted average of a few data points. Sugeno-type systems support this type of model. In general, Sugeno-type systems can be used to model any inference system in which the output membership functions are either linear or constant.

Figure (2) shows fuzzy logic process flow.

III. PROPOSED MODEL

Forecasting of WT power generation has two main aspects. The first aspect is the historical and real time data. The historical data are used to analyze and model the forecasting system where the current real time data are used with historical data to calculate the future value of the generated power depending on the model parameters and analysis. The historical data can be obtained from any data storage media in a monitoring center where the real time data need a measuring system that fits for the conditions of the WT. The second aspect is the modeling techniques that combine real time and historical data to predict the future value of WT power depending on the analysis with acceptable error.

A. Data description

WT data is usually collected by a Supervisory Control and Data Acquisition (SCADA) system in monitoring center. So, historical data can be obtained from a SCADA system installed at a wind farm. The real time data can be collected efficiently using the proposed wireless sensor network (WSN) which creates a new communication model for Real-time and reliable measurement. Using WSN in wind power generation prediction can achieve some highlighted benefits such as savings in cabling costs and rapid installation of the communication infrastructure [31], [32].

As mentioned earlier, the major factors affecting wind farm output power are air temperature, wind speed, air density and air pressure. Considering all these factors will increase the prediction accuracy of the proposed model. Therefore, the real time values for these factors could be measured in the proposed prediction model using WSN with acceptable period and accuracy. In order to improve the accuracy of prediction, the historical power output collected by the Energy Management System (EMS) can be used as input. Figure (1) show the WSN installed to WT to measure
real time values of the prediction parameters.

The parameters: air temperature $T(t)$, wind speed $V(t)$, air density $\rho(t)$, pressure $P(t)$, and output power $S(t)$ are the real time measured values coming from WSN. Figure (3) shows the framework for the prediction system data flow. The data is measured using the WSN and sent to the SCADA system over wireless communication media. The SCADA system receives the data using specific receiver module and save these data to database (DB) for historical data analysis and archive then display these data on human machine interface (HMI) interface for the operator. The prediction model brings the historical data from DB and real-time data from the receiver module to predict the output power for WT.

The prediction area shows in figure (3) begins with computing the membership function for each input coming from data set. The member ship function is computing using the parameters centriod and width $(\alpha_{ij}, \sigma_{ij})$. The fuzzification step convert the inputs value to crisp inputs then calculate the output using a rule base. The output of the prediction area is sent to HMI for the operator and making decision.

B. Proposed prediction model

The data set which is coming from WSN is representing the input data space and it is used to train and validate in the proposed model. The training data set is clustered using a modified Fuzzy C-Means (FCM), which is a data clustering algorithm to decide the optimal number of fuzzy rules by grouping the dataset into $n$ clusters with every data point in the dataset belonging to every cluster to a certain degree and assign a rule for each cluster [33]. Based on the data in each cluster, number of rules and membership function is adjusted. Hybrid optimization method, which is a combination of least-squares and back propagation gradient descent method, is used to train the membership function parameters to emulate the training data [34]. The steps of the prediction model methodology are discussed in the next section.

C. Clustering input space

In this section, each input in the training dataset is clustered into $k$ subspaces. Based on factors selected for clustering as follows:
- Range of influence: specifies a cluster center’s range of influence in each of the data dimensions.
- Squash factor: This factor determine the neighborhood of a cluster center.
- Accept ratio: This factor sets the potential above which another data point is accepted as a cluster center.
- Reject ratio: This factor sets the potential below which a data point is rejected as a cluster center.

The clustering operation returns the cluster centers in the matrix $C$; each row of $C$ contains the position of a cluster center. The returned $S$ vector contains the sigma values that specify the range of influence of a cluster center in each of the data dimensions. All cluster centers share the same set of sigma values.

D. Model extraction and validation

After clustering the training data to $C$ clusters, each input

---

**Figure (3): framework for the prediction system data flow**

![Framework Diagram](attachment:image.png)
is illustrated to C clusters and C membership function with Gaussian membership function. So, for \( n \) input and \( C \) clusters with specific center, the equation of membership function is shown in equation (5).

\[
    u_{ij}(x_j) = \exp\left(-\frac{(x_j - \alpha_{ij})^2}{\sigma_{ij}^2}\right)
\]

(5)

Where \( \alpha_{ij} \) and \( \sigma_{ij} \) are the center and width of the \( j \)th membership function in the \( i \)th input, \( i=1, 2, \ldots, n \), \( j=1, 2, \ldots, C \).

A fuzzy model is designed as a set of rules depending on the number of the clusters number. The rules are set in the IF-THEN form to describe input-output relations using Takagi-Sugeno with linear function to output. A multi-input and single-output fuzzy model is represented as a collection of fuzzy rules in a Sugeno fuzzy model which has the following form.

\[
    R_k: \text{IF } X_1 \text{ is } A_{1j} \text{ and } X_2 \text{ is } A_{2j} \ldots \text{X}_n \text{ is } A_{nj} \text{ then output } = f_k(X_i)
\]

where \( X=(X_1,X_2,\ldots,X_n) \) are the inputs to fuzzy model, \( A_{1j},A_{2j},\ldots,A_{nj} \) are linguistic variables, \( R_k \) represents \( k \)th rule, \( k,j=1,2,\ldots,C \) which represent the cluster number, \( f_k(X_i) \) is Linear relation between inputs and output which is illustrated in equation (6).

\[
    f_k(x) = b_{0j} + \sum_{i=1}^{n} b_{ij} x_i
\]

(6)

The crisp output is calculated using the equation (7). The output level \( f_p \) of each rule is weighted by the firing strength \( w_p \) of the rule.

\[
    \text{output} = \frac{\sum_{p=1}^{k} w_p f_p}{\sum_{p=1}^{k} w_p}
\]

(7)

The \( w_p \) is weight of and memberships value for linguistic variable for the inputs in IF statement of the rules \( P \). the weight \( w_p \) relation is illustrated in equation (8).

\[
    w_p = \text{and method}(\mu_p)
\]

(8)

The and-method may be product or minimum of membership value. In this paper, minimum and-method is used. The block diagram of prediction model used is shown in figure (5). The data is coming from SCADA database and WSN to fuzzify and calculate the output using the rule base. The fuzzy logic system with wtaverde-fuzzifier, product-inference rule and Gaussian-minimum fuzzifier is proposed to predict the output power of the WT.

An Adaptive Network based Fuzzy Inference System can incorporate fuzzy if-then rules and also, provide fine-tuning of the membership function according to a desired input output data pair. Figure (4) shows the adaptive network with 5 layers for the fuzzy model of the prediction model. As mentioned before, a first order linear sugeno fuzzy model is used as a means of modeling fuzzy rules into desired outputs.

The adaptive network consists of 5 layers. Layer 1 corresponds to a linguistic label and the membership function of this linguistic label. Layer 2 estimates the firing strength of a rule, which is found from the and-method (multiplication or minimum) of the incoming signals. Layer 3 estimates the ratio (\( w_i \)) of the \( i \)th rule's firing strength to the total firing strength. Layer 4 and Layer 5 are used for defuzzification and output calculation.

\[
    \sum_{i=1}^{n} w_i = \sum_{p=1}^{k} w_p
\]

Figure (4): adaptive network with 5 layers for the fuzzy model of the prediction model

- **Layer 1**: Input membership functions
- **Layer 2**: Rules
- **Layer 3**: Weight ratio of rule
- **Layer 4**: Defuzzification in Sugeno model
- **Layer 5**: Output
sum of the firing strength of all rules. C. Layer 4 is the product of the previously found relative firing strength of the ith rule and the rule. The final layer computes the overall output as the summation of all incoming signals from layer 4 [35].

For train the prediction model, data could be selected to cover all seasons of the year i.e. four subsets data could be selected from data collected to train the prediction model in different periods of year as show in table (1).

Table (1): subsets periods to be used in train the prediction model

<table>
<thead>
<tr>
<th>Subset No.</th>
<th>Period From</th>
<th>Period To</th>
<th>season</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15. April</td>
<td>15. May</td>
<td>Spring</td>
</tr>
<tr>
<td>2</td>
<td>15. July</td>
<td>15. August</td>
<td>Summer</td>
</tr>
<tr>
<td>3</td>
<td>15. October</td>
<td>15. November</td>
<td>Winter</td>
</tr>
<tr>
<td>4</td>
<td>15. January</td>
<td>15. February</td>
<td>1000</td>
</tr>
</tbody>
</table>

The flow chart for building trains and evaluates the prediction model is shown in figure (5). The flow chart is divided into four steps. First step, prepare data for clustering. Second step, set the clustering parameters create the fuzzy rule and start the training on 1000 samples of data called epochs. Third step is for training the model with acceptable RMSE. Forth step is for Evaluate the model and compute error.

IV. CONCLUSION

In this paper, a complete system is proposed for prediction of wind power generation. The system uses a neuro-fuzzy model based fuzzy inference system to predict the generated power. The prediction system uses the parameters measured by WSN. The proposed prediction model considers the major factors affecting the wind output power, namely air temperature, wind speed, air density and air pressure. Considering all these factors increases the prediction accuracy of the proposed model. Training and evaluating flow chart is proposed for four season data to cover year time.

REFERENCES

[1] A. Sfetsos, “A comparison of various forecasting techniques applied to...


