Short-Term Wind Speed Prediction Based on the Weighted Regular Extreme Learning Machine

Da-cheng XING1, Ben-shuang QIN1,* and Cheng-gang Li2

1College of Electrical Engineering, Northeast Dianli University, Jilin 132012, Jilin, China
2Electric Power Research Institute, Jilin Electric Power Company Limited, Changchun 130021, China
*Corresponding author

Keywords: mRMR, Extreme learning machine, WRELM, Prediction model.

Abstract. Accurate wind speed is the basis of the wind power prediction. And it is of great significance to the parallel operation of the wind farm. So it is to the maintenance of the safety and stability of the power system. In the view of the strong volatility and randomness of the wind speed time series, it is difficult to be predicted. A new method of short-term wind speed prediction is established based on the weighted regular extreme learning machine (Weighted Regularized Extreme Learning Machine, WRELM). First, the wind speed time series and the wind direction time series which have high correlation with the wind speed are taken account. Besides the meteorological factors are also taken as the candidate sets such as temperature, pressure, humidity and so on. Then the maximal relevance minimal redundancy (mRMR) principle is used to select the maximum serial correlation properties. And those are taken as the prediction inputs. Afterwards, the train set and test set of the prediction network are fixed to establish WRELM. Then, the network parameters are trained with the train set data. And the WRELM prediction model is built. Finally, the WRELM network is adopted to predict the short-term wind speed and the future wind speed are obtained. The data from the wind farm is carried out to do the experiment. And it wants to prove the effectiveness of the new method.

Introduction

With the mass consumption of fossil energy and the growing worldwide energy crisis, countries all over the world are looking into clean energy[1]. Wind as a clean energy attracts more and more attention. However, the volatility, intermittent and randomness of the wind itself does harm to the security and the stability of the power system. And it hinders the large-scale application of wind power[1-3]. Reliable wind power and wind speed prediction is the effective measure to eliminate the adverse effect. In order to improve the short-term wind speed forecasting accuracy further, this paper proposes a new method of short-term wind speed forecasting model based on WRELM.

Feature Select

The wind is caused by pressure imbalance of the air flow in the air. So it is vulnerable to the effects of meteorological factors such as temperature, pressure, humidity[1, 3-5]. In a short period of time wind speed has a certain similarity at the same time around the same wind turbines. Due to the limitation of data collection, only the temperature, humidity and the air pressure 3 kinds of meteorological factors are taken into the wind velocity feature set. So it has been gained

\[ Q = \{S(t-1), S(t-2), \ldots, S(t-d), D(t-1), D(t-2), \ldots, D(t-d), T(t-1), T(t-2), \ldots, T(t-d), P(t-1), P(t-2), \ldots, P(t-d), H(t-1), H(t-2), \ldots, H(t-d)\} \]  

\[ S(t-i), D(t-i), T(t-i), P(t-i) \text{ and } H(t-i) \text{ represent the wind speed, wind direction, temperature, pressure and humidity the delay time value } t-i \text{ differently. } d \text{ is the total delay time and is taken 6 here in reference [5] and [6]. Therefore, the initial input data set has high dimension and is difficult to direct forecast computation. And it contains a large number of redundant information of correlation with the wind speed. The mRMR principle is taken as feature selection based on the mutual information. } \]
Define two random variables $x$ and $y$. So the probability density is $p(x)$ and $p(y)$ respectively. And the joint probability density is $p(x, y)$. So the mutual information between the $x$ and $y$

$$I(x; y) = \int \int p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \, dx \, dy$$

The maximal and minimum redundancy criterion respectively

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c)$$

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j)$$

Among them, $S$ and $|S|$ respectively represent features that are included in the set and its number. $c$ is the target category. Besides, $I(x_i; c)$ and $I(x_i; x_j)$ is the mutual information between $x_i$ and $c$ and $x_j$ respectively. $D$ is the average mutual information between each feature of $S$ and $c$. And it measures the degree of redundancy between feature set and between the feature set and category. $R$ is the mutual information of two different features in $S$. It represents the degree of redundancy between two features. Feature selection is to choose the category associated with the target as high as possible feature subsets, number of features and less as far as possible at the same time. It requires that the selected feature subsets and the target category of correlation, the biggest characteristics between the minimum redundancy. So mRMR is

$$\max \Phi(D, R), \Phi = D - R$$

As a result, the mRMR criterion is adopted to wind speed initial features set for feature selection, as shown in figure 1 choice results are obtained.

According to the results of figure 1, from big to small on basis of mRMR one added feature vector, the test result is shown in figure 2. When the number of features set $\Phi$ collected at 7, MAPE obtains the minimum value, namely the best prediction effect.
WRELM

WRELM is proposed on the basis of ELM [5]. Define \( g \) is the ELM activation function. The network can be described as

\[
f_L(x) = \sum_{i=1}^{L} \beta_i g(\omega_i x + b_i) = z
\]

Among them, \( \omega_i \) and \( \beta_i \) represent the connection weights of the input and output vector between the hidden layer neurons vector respectively. \( b_i \) is the offset value of the \( i \) hidden layer neurons.

ELM has more advantages, but there also exists the following problems [4-6]. Because of the lack of structural risk assessment standard, the ELM model is not the optimal model.

In order to overcome the above shortcomings and to enhance the generalization ability of ELM network, some scholars on the basis of ELM introducing weighting factor and coefficient of regular, build WRELM [4, 6]. The objective function for WRELM is

\[
\min E = \min_{\beta} \left( \frac{\lambda}{2} \| w \|_2^2 + \frac{1}{2} \| \beta \|_2^2 \right)
\]

Among them, \( \lambda \) is the coefficient of the regular type. \( \varepsilon_j = \sum_{i=1}^{L} \beta_i g(\omega_i \cdot x_j + b_i) - z_j \) is the training error, \( j = 1, 2, \cdots, N \); \( w = \text{diag}(w_1, w_2, \cdots, w_N) \) is the weight of the diagonal matrix. \( \| w \|_2^2 \) and \( \| \beta \|_2^2 \) represent the experience and structural risk respectively. The Lagrange equations solvable output weight matrix are obtained as follow

\[
\hat{\beta} = (H^T w^2 H + \frac{I}{\gamma})^{-1} H^T w^2 T
\]

Among them \( I \) is the matrix for the unit. Then it can be fitted regression model based on RELM wind speed forecasting using formula (6).

\[
y = \sum_{i=1}^{L} \hat{\beta}_i g(\omega_i x + b_i)
\]

Known from the analysis of the above, when predicting network's input, output is determined, short-term forecasting model can build wind speed.

**Forecasting Network Building**

Considering the MRMR results in the new method, it selects 7 features as the wind forecasting
inputs and the wind speed forecasting time point as the output with the test experiment results. And it takes single step to build the short-term wind speed forecasting model. The flow chart of the new method is as Fig.3. New method in reference [4] and [5], and cross validation method is used to determine WRELM network hidden layer nodes, the activation function, regular coefficient and weight function and other parameters.

In order to cover the wind information fully and enhance the applicability of the model, the data is used of the temperature, pressure, humidity, wind speed, wind direction and other meteorological information data as the training data from 39.91 ° north latitude and longitude 105.29 of wind power technology centers in the United States in 2004. The NWTC record data intervals is 1 min and the experimental data for the corresponding data the average of 1 hour[4]. Choose every an hour’s wind speed, wind direction, 1 interval time before interval values of temperature, humidity and wind speed, wind direction, air pressure before 6 time value as the network input characteristics of prediction and the output moment to be forecast wind speed to form a training set. Forecast for moment characteristic into seven has trained WRELM network, the output is the moment to forecast wind speed. Using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) as indicators of forecast, the new method has higher accuracy and its validity is proved[1,3].

**Conclusion**

In order to improve the short-term wind speed forecasting accuracy further, this paper proposes a short-term wind speed forecasting model using WRELM. And to overcome the shortcomings in the high wind speed characteristic dimension, not easy to predict, mRMR feature selection is taken in the new method and to select the optimal prediction of the input set. Compared with the traditional model of ELM, WRELM considers the structure of risk and risk model. So it obtains the better network model. At the same time the training weights are joined to avoid the outliers to predict the results of interference. All in all it improves the network generalization and fitting ability.

**Reference**

[1] Sfetsos A. A novel approach for the forecasting of mean hourly wind speed time series[J]. Renewable Energy, 2002, 27(2):163-174.

[2] Fan S, Liao J R, Yokoyama R, et al. Forecasting the wind generation using a two-stage network based on meteorological information[J]. IEEE Transactions on Energy Conversion, 2009, 24(2):474-482.

[3] Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(8):1226-38.
[4] Huang G B. What are extreme learning machines? Filling the gap between Frank Rosenblatt’s Dream and John von Neumann’s Puzzle[J]. Cognitive Computation, 2015, 7(3):263-278.

[5] Deng W, Zheng Q, Chen L. Regularized Extreme Learning Machine[C]// Computational Intelligence and Data Mining, 2009. CIDM ’09. IEEE Symposium on. IEEE, 2009:389-395.

[6] Huynh H T, Won Y. Weighted least squares scheme for reducing effects of outliers in regression based on extreme learning machine[J]. International Journal of Digital Content Technology & Its Applications, 2008, 2(3):40-46.