A Study of Wind Turbine Comprehensive Operational Assessment Model Based on EM-PCA Algorithm

Minqiang Zhou, Bin Xu, Yangyan Zhan*, Danyuan Ren, Dexing Liu

State Key Laboratory of Wind Power System, Zhejiang WonderCon, Ltd, Hangzhou, China

*Corresponding author e-mail:zhanyy@chinawindey.com

Abstract. To assess wind turbine performance accurately and provide theoretical basis for wind farm management, a hybrid assessment model based on Entropy Method and Principle Component Analysis (EM-PCA) was established, which took most factors of operational performance into consideration and reach to a comprehensive result. To verify the model, six wind turbines were chosen as the research objects, the ranking obtained by the method proposed in the paper were 4#>6#>1#>5#>2#>3#, which are completely in conformity with the theoretical ranking, which indicates that the reliability and effectiveness of the EM-PCA method are high. The method could give guidance for processing unit state comparison among different units and launching wind farm operational assessment.

1. Introduction

Due to the extreme operational environment and complicated conditions, there are deviations between the actual operation and the preliminary design. A comprehensive operational assessment including actual operating conditions of wind turbines from the perspective of performance, power performance, reliability and so on, and explore the gap between operational stage and preliminary design. It provides the theoretical basis for targeted optimization or upgrading of the wind turbine.

At present, many researchers have studied the comprehensive assessment methods in operational statuses or reliability of wind turbines. The articles [1-2] introduced a systematic comparative ranking method based on entropy value method; in literature [3], a comprehensive assessment method combining matter-element analysis and deterioration degree variable weight analysis was introduced; in document [4], the reliability of the wind turbine is evaluated based on the grey entropy analytic hierarchy process (AHP) and the method of distance of good-gad solutions, and obtained the results of wind turbine condition assessment, the reliability index simulation was applied to a wind farm of 9 wind turbines, which provided a new approach for making scientific maintenance plan of wind turbines. Document [5] based on the traditional entropy method of positive and negative indicators, the intermediate index is introduced to calculate the information entropy and weight of each subsystem parameter, and the state of the unit is evaluated according to the actual wind field data; Document [6] used fuzzy comprehensive assessment method to identify abnormal working conditions of wind turbines, and determined the membership degree and fuzzy boundaries according to the degree of deterioration, realized a comprehensive assessment. Even though all the studies above provided some
ideas to make operational assessment, an accurate and practical comprehensive assessment methods are less mentioned and still hard to obtain.

To accurately assess wind turbine performance, a comprehensive operational assessment model of wind turbine based on EM-PCA is presented. The statistics of operational data from Supervisory Control and Data Acquisition System (SCADA) are used to calculate the values of index. Combined with EM-PCA algorithm, the quantitative assessment of the wind turbines’ comprehensive performance is realized.

2. Assessment Index System
Currently, many operational assessment indexes are widely used in wind industry, some of which are not completely independent. Therefore, a process of selecting operational assessment indicators is prerequisite for simplifying assessment process and reach to more accurate results. The calculation of indexes in this paper is based on operational data from SCADA.

According to attributes of common-used indexes, an index system of comprehensive operational assessment model was constructed, as shown in Fig. 1

![Figure 1. Framework of comprehensive operational assessment indexes](image)

The index system has three layers, which are comprehensive index, sub-indexes and basic indexes. The meaning of each basic index is described as follows.

K, the contractual index of power curve, is the ratio of theoretical power generation corresponding to the measured power curve and the design power curve. According to the general requirement of contract, the output performance of wind turbine is qualified if the value of K is greater than 95%, otherwise it is unqualified. PPSD is short for Power Performance Standard Deviation, which reflects the deviation between the measured power curve and the guaranteed power curve in the wind speed bin between cut-in and cut-out. The industrial standard value is less than or equal to 5%.

Time Based Availability (TBA) refers to the ratio of the unit runtime to the assessment period, the general requirement is greater than 95%. Service Factor (SF) refers to the ratio of the number of hours in which unit is in operation to the total hours of assessment period. RMH is routine maintenance hours during statistical period that mainly includes 500-hour maintenance, half yearly maintenance, one-year maintenance, daily inspection and alignment. RMT is routine maintenance times during the statistical period. Failure Hours (FH) refers to breakdown time during the statistical period. FT refers to failure times during the statistical period. MTBF is short for mean time between failures, which reflects the reliability of unit. The less MTBF is, the better reliability is.
Mean Logistic Delay Time (MLDT) refers to the average value of the total time spent by the on-site failure handlers during the assessment period, and represents the speed and ability of the personnel to respond to the failure. MLDT reflects the rapid response of staff at the scene. Mean Time to Repair (MTTR) refers to the average value of the total time spent by the on-site personnel in handling the failures during the assessment period. Generally, the smaller MTTR is, the higher efficiency of the on-site staff is. Effective Wind Rate (EWR) refers to the ratio of the duration of the wind speed between a cut-in and a cut-out at (or near) the hub height to the number of calendar hours in the statistical period. Off-site Downtime (OD) refers to all the downtime due to external problems (not relate to wind turbine) such as grid failure, extreme weather and personal order from owner.

For all the indexes above, the smaller values of PPSD, RMH, RMT, FH, FT, OD, MLDT, MTTR, the unit performance would be better, and therefore these indexes are defined as the cost index; The remaining indexes are defined as the profitable index that the greater the value, the better the performance of the unit.

3. Comprehensive Operational Assessment Model

3.1. EM and PCA

Entropy Method (EM) is a process of attributing objective weight according to the amount of information contained in the observations of different indexes, and is a measure of system uncertainty. The smaller entropy of an index, the greater amount of information is provided, as well as the greater weight in the comprehensive assessment, and vice versa.

Principal component analysis (PCA) is a statistical procedure of dimension reduction with little loss of information (the general requirements to retain more than 90% information of original variables), as far as possible the feature extraction of the observation data, and achieve accurate assessment.

3.2. Assessment model based on EM-PCA algorithm

Based on Fig. 1, the comprehensive assessment model of wind turbines based on EM-PCA algorithm can be divided into two steps, as shown in Fig. 2

![Figure 2. Procedure of EM-PCA algorithm](image)

3.2.1. Sub-index model based on EM algorithm. If the number of study subjects is $n$ and any sub-index is made up with $m$ basic indexes, thus the decision matrix of sub-index is $C_{ik} = [c_{ij}]_{ik}$, which in columns are basic indexes of different units and in rows are different indexes of one unit, $k$ is the number of sub-indexes.

To ensure the consistency of basic index in comprehensive assessment, and eliminate the dimensional effects between indicators, the decision matrix is standardized as showed below.

$$
c'_{ij} = \begin{cases} 
\frac{\max(c_{ij}) - c_{ij}}{\max(c_{ij}) - \min(c_{ij})}, & c_{ij} \text{ is cost-type indicator} \\
\frac{c_{ij} - \min(c_{ij})}{\max(c_{ij}) - \min(c_{ij})}, & c_{ij} \text{ is efficiency indicator} 
\end{cases}
$$

(1)
Here, \( i = 1, 2, ..., n \), and \( c_{ij} \geq 0 \).

The standardized decision matrix is denoted as \( C'_{kb} = (c'_{ij})_{nm} \), the entropy \( E_j \) of index \( j \) is:

\[
E_j = -e \sum_{i=1}^{n} c_{ij} \ln(c_{ij}), \quad j = 1, ..., m
\]  

(2)

In equation, \( k = 1/\ln(n) \) is a constant related to the number of study subjects, with the purpose \( E_j \in [0, 1] \), and \( 0 \leq c_{ij} \leq 1 \). When \( c_{ij} = 0 \), set \( c_{ij} \ln(c_{ij}) = 0 \). The EM algorithm indicates the disorder degree of information, and the larger \( E_j \) is, the fewer information is contained in the target index \( j \), as well as the weaker ability on assessing the performance of unit. Therefore, the valid information \( d_j \) and weight \( \omega_j \) of index \( j \) are showed below:

\[
d_j = 1 - E_j
\]

(3)

\[
\omega_j = \frac{d_j}{\sum_{j=1}^{m} d_j} = \frac{1 - E_j}{m - \sum_{j=1}^{m} E_j}, \quad \sum_{j=1}^{m} \omega_j = 1, (j = 1, ..., m)
\]

(4)

With the respect to sub-index, the weight reflects the importance of each basic index to the sub-index. The greater the weight \( \omega_j \) of the index \( j \) is, the greater the impact on the sub-index \( B_k \), and the greater contribution to the assessment result. Therefore, the assessment value of sub-index is:

\[
B_{ij} = \sum_{j=1}^{p} \omega_j c_{ij}, \quad i = 1, ..., n
\]

(5)

In equation: \( 0 \leq B_{ij} \leq 1 \), the larger \( B_{ij} \) is, the unit’s performance is better. Repeating the above steps and constructing the decision matrix \( B_{n,p} \), \( p \) is the number of sub-indexes.

3.2.2. The comprehensive assessment model with PCA algorithm. The PCA algorithm is applied to the decision matrix of the comprehensive index \( B_{n,p} \), and the correlation coefficient matrix \( R \), \( p \) characteristic roots and corresponding eigenvectors are calculated:

\[
\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_p \geq 0; \quad e_j = (l_{ij}, l_{j2}, ..., l_{jp}), \quad j = 1, ..., p
\]

(6)

According to the obtained eigenvalues, calculate the variance contribution \( \alpha_l = \lambda_l / \sum \lambda_i \) and sort it by descending order. When former \( l (l \leq p) \) index’s cumulative variance contribution rate arrives the request that \( \alpha_{cum} = \sum_{l=1}^{l} \alpha_l > 90\% \), it is considered that the \( l \) principal component can comprehensively represent the \( P \) indexes. Therefore, according to the former \( l \) main components and calculating the comprehensive index as below, which can represent each wind turbine’s performance:
Finally, sort $F_i$, which represent the rank of the wind turbines, quantitatively.

4. Comprehensive assessment model verification

To verify the effectiveness and feasibility of model, took six 1.5MW wind turbines (1#~6#) as the research objects and January to October of 2016 as the assessment period. Based on the operational data, basic indexes of each assessed unit are calculated, as shown in Table 1.

| Units | PPSD (%) | K (%) | MLDT | MTTR | OD | EWR (%) | RMT | RMH | FH | FT | TBA (%) | SF (%) | MTBF |
|-------|----------|-------|------|------|----|---------|-----|-----|----|----|--------|--------|------|
| 1#    | 2.11     | 110.9 | 4.83 | 1.7  | 297.6 | 75.9  | 12   | 14.15 | 126.4 | 20 | 98.2   | 94.0   | 312.8 |
| 2#    | 2.70     | 107.8 | 5.90 | 1.4  | 298.3 | 76.7  | 11   | 12.83 | 127.9 | 18 | 98.2   | 94.4   | 354.5 |
| 3#    | 2.85     | 109.0 | 9.11 | 3.3  | 346.9 | 77.6  | 12   | 13.45 | 198.1 | 16 | 97.2   | 93.2   | 379.0 |
| 4#    | 2.38     | 111.6 | 0.37 | 2.2  | 301.0 | 74.2  | 14   | 16.22 | 28.0  | 11 | 99.6   | 94.6   | 577.0 |
| 5#    | 2.46     | 107.2 | 2.50 | 3.7  | 296.1 | 73.9  | 13   | 15.70 | 65.6  | 12 | 99.0   | 94.7   | 494.9 |
| 6#    | 2.36     | 114.7 | 3.66 | 2.2  | 306.3 | 79.3  | 8    | 12.30 | 86.1  | 15 | 98.7   | 94.1   | 405.0 |

Based on Table 1, hybrid EM-PCA algorithm is applied in the operational assessment, and sub-indexes of each wind turbine are calculated, as shown in Table 2.

| Units | Performance | Reliability | External factors | O&M |
|-------|-------------|-------------|------------------|-----|
| 1#    | 0.7399      | 0.3468      | 0.7755           | 0.7666 |
| 2#    | 0.1367      | 0.5172      | 0.8142           | 0.6038 |
| 3#    | 0.1203      | 0.2529      | 0.2189           | 0.0000 |
| 4#    | 0.6089      | 0.7226      | 0.6251           | 0.8022 |
| 5#    | 0.2577      | 0.6543      | 0.6778           | 0.6015 |
| 6#    | 0.8355      | 0.6812      | 0.8648           | 0.7049 |

From table 2, several conclusions were obtained:
1) In the view of performance, the order is 6# > 1# > 4# > 5# > 2# > 3#;
2) In the view of reliability, the order is 4# > 6# > 5# > 2# > 1# > 3#;
3) In the view of external factors, the order is 6# > 2# > 1# > 5# > 4# > 3#;
4) In the view of operation and maintenance, the order is 4# > 1# > 6# > 2# > 5# > 3#;

In sum, unit 3# performs the worst; 4# performs the best in reliability and maintenance management, which have a significant influence in comprehensive assessment; 6# performs the best in external factors and performance and 2nd in reliability. By comparative analysis, the four sub-indexes cannot achieve a comprehensive assessment. Therefore, PCA was applied to realize a comprehensive result, and the results are showed in table 3.

| Units | 1# | 2# | 3# | 4# | 5# | 6# |
|-------|----|----|----|----|----|----|
| Com-Index | 0.3579 | 0.1001 | 0.0748 | 0.5095 | 0.2591 | 0.4877 |
The comprehensive assessment result is 4# > 6# > 1# > 5# > 2# > 3#. To verify the accuracy of the results, wind energy equation was adopted, as shown in equation (8).

\[ P = \frac{1}{2} \rho V^3 A \eta * C_p \]  

(8)

where \( P \) is power, kW; \( \rho \) is air density, kg/m\(^3\); \( V \) is wind velocity, m/s; \( A \) is the sweeping area, m\(^2\); \( \eta \) is mechanical efficiency, which can be ignored in calculation; \( C_p \) is the efficiency of power. Therefore, the energy production of wind turbine is \( P \) times \( t \), as shown in equation (9):

\[ E = P \times t \]  

(9)

Where \( E \) is the energy production, kWh; \( t \) is the generating hours of wind turbine. According to the equation (8) ~ (9), the wind turbine operational performance bias mainly decided by \( C_p \) and \( t \) when \( \rho \), \( A \) are constants in a wind farm. Thus, the ratio of the energy production and the three times of the wind speed, which equal to \( C_p \) multiply \( t \), is taken as the reference of operational performance. As a result, the theoretical ranking of each unit is calculated, as shown in Table 4.

Table 4. The theoretical ranking of each unit

| Units | 1#   | 2#   | 3#   | 4#   | 5#   | 6#   |
|-------|------|------|------|------|------|------|
| \( P/V^2 \) | 19.58 | 18.34 | 17.17 | 20.40 | 18.65 | 20.18 |
| Order | 3rd  | 5th  | 6th  | 1st  | 4th  | 2nd  |

Table 4. shows that the theoretical ranking order is 4# > 6# > 1# > 5# > 2# > 3#, which exactly matches with the result of comprehensive assessment.

5. Conclusion

The paper presented a hybrid EM-PCA model by combining Entropy Method with Principle Component Analysis algorithms, which applied to assess the comprehensive performance of wind turbine units. the model was verified by wind farm operational result. For one thing, the model provides a new method that compare and analysis the comprehensive performance of units in one or more wind farms. Besides, sub-indexes can pinpoint the reasons why different units performs variously, and provide theoretical basis for the subsequent improvement.

Acknowledgments

This work was financially supported by Chinese National Programs for Science and Technology Development Fund (No. 2015BAA06B01)

References

[1] QI Minfang, FU Zhongguang, JING Yuan, et al. A comprehensive evaluation method of power plant units based on information entropy and principal component analysis [J]. Proceedings of the CSEE, 2013,3 (32): 58-64. (in Chinese)

[2] Tang Wenbin, Han Zhijun. A comprehensive evaluation method of financial position based on information entropy [J]. Journal of Nanjing University of Science and Technology, 2001, 25 (6): 650-653. (in Chinese)

[3] LI Hui, HU Yaogang, YANG Chao, et al. A matter-element assessment method of a grid-
connected wind turbine driven generator system under on-line operating conditions [J]. Automation of Electric Power Systems, 2011, 35 (6): 81-85. (in Chinese)

[4] Di Shuai, Song Peng, Yao Wanye, et al. Research on reliability evaluating of wind turbine based on grey entropy AHP and TOPSIS method [J]. North China Electric Power, 2016 (8): 24-28. (in Chinese)

[5] Li Dazhong, Xu Bingkun, Chang Cheng. Comprehensive evaluation of operation state of large wind turbine based on entropy method [J]. Journal of Engineering for Thermal Energy and Power, 2016, 31 (4): 1-5. (in Chinese)

[6] MAN Ruoyan, FU Zhongguang. Condition assessment of fossil-fired plant based on fuzzy comprehensive evaluation [J]. Proceedings of the CSEE, 2009 (5):5-10. (in Chinese)