Steam Turbine Performance Evaluation Based on Principal Component Analysis and Information Entropy

Chengbin Lu¹, Taiyan Zhang, Xiaofeng Liu, Yongling Yao, Huagui Chen
Jiangsu Frontier Electric Technology Co., Ltd.
²xxg@ncepu.edu.cn

Abstract. Information entropy and fuzzy analytical theory were combined and applied on the state assessment of power plant steam turbine unit. This method can be combined with the measured data of power plant, and made a comprehensive evaluation of the different working conditions. Fuzzy analysis of the original matrix can be a evaluation matrix, and the information entropy was introduced to solve the problem of calculating the weight of multi attribute factor set. Once the attribute weight is determined, the comprehensive evaluation value corresponding to the different working conditions can be defined and calculated. The results obtained by this method were compared with those obtained by the principal component analysis method. And the results obtained by information entropy method were objective and unique. The superiority of the method was verified. Under all conditions, the comprehensive evaluation value can be used as one index of the thermal archives as a benchmark for fault diagnosis, which provided a new idea and evaluation method for steam turbine fault diagnosis, and had a strong basic research and practical application value.

1. Introduction
Steam turbine was an important equipment of thermal power plant. Due to its complex and harsh working environment, it has become a "high-risk area" for power plant faults[1,2]. Experts and scholars have conducted researches on this subject with various research methods, such as principal component analysis, expert domain evaluation, fuzzy analysis. However, there were more or less problems in different methods. Principal component analysis was a multivariate statistical analysis method that converted multiple indicators into a few comprehensive indicators [3]. Although a small number of unrelated new variables reflected most of the information provided by the original variables, after all, there was a loss of information, which cannot restore the truth 100%.

Information entropy was the content of information theory. Similar to thermodynamic entropy, the concept of information entropy was widely used in various fields of society to describe the uncertainty of information sources. In this paper, information entropy was combined with fuzzy theory.

2. Steam Turbine Performance Evaluation
In this paper, 320MW unit of a power plant was selected as the research object. Theoretically, the operating conditions should be as comprehensive as possible to cover all operating conditions of the unit when establishing the thermal files of the unit, but in the actual application of the project, it was not realistic to cover the values of the entire operating conditions. In practical application, the interval should be 1 ~ 5MW. This paper was limited to the space and only took 200MW ~ 300MW every 10MW for analysis. In practical application, the thermal archives can be established by calculating and analyzing all operating conditions at a smaller interval. [3,4,5].
2.1. Selection of thermal parameters
In the SIS system of power plant, there were a lot of measuring points of thermal parameters. In this paper, 12 important parameters were selected, which were based on the reliability of measuring points. Therefore, 12 important thermal parameters were monitored. These parameters were selected as the original data in this paper. The specific thermal parameters were shown in Table 1.

Table 1. The original parameters

| Parameter                     | Description                   | Unit     |
|-------------------------------|-------------------------------|----------|
| $X_1$                         | Main steam pressure          | /MPa     |
| $X_2$                         | Main steam temperature       | /°C      |
| $X_3$                         | Reheat steam pressure        | /MPa     |
| $X_4$                         | Reheat steam temperature     | /°C      |
| $X_5$                         | Regulating stage pressure    | /MPa     |
| $X_6$                         | Regulating stage temperature | /°C      |
| $X_7$                         | First extraction pressure    | /MPa     |
| $X_8$                         | Eighth extraction pressure   | /MPa     |
| $X_9$                         | Eighth extraction temperature| /°C      |
| $X_{10}$                      | Main feed water temperature  | /°C      |
| $X_{11}$                      | Regulating stage pressure    | /MPa     |
| $X_{12}$                      | Deaerator pressure           | /MPa     |

2.2. Establishment and selection of membership function
There were five experimental methods to determine the membership function: horizontal method, inference method, vertical method, parameter estimation and fuzzy clustering method. The choice of these methods was related to the specific problems. In practical application, it was usually decided by the expression and recording method of the uncertainty of the problems. Limited to the space, this paper would not elaborate one by one. For the unit state evaluation, the conclusion to be obtained was the qualitative evaluation result of the unit under different operating conditions. Therefore, a semi-trapezoidal membership function can be chosen. In view of the characteristics of the factor set, two types of membership functions, large and small, were selected. The specific function selection was described later. [6,7,8]

Without considering the mutual coupling, the influence of each parameter on the overall condition of the unit was considered separately. The main steam pressure, the main steam temperature, the reheat steam pressure, the reheat steam temperature, the pressure and temperature of the regulating stage, and the deaerator pressure were all large membership functions. The larger the value was, the better the attribute was. The smaller the value of extraction pressure and temperature was, the better was, which belonged to a small membership function.

2.3. Information entropy for comprehensive evaluation value
Due to the large number of thermal parameters selected, it was difficult and not objective for the weight of traditional expert’s evaluation. The information entropy method introduced above can solve the problem of the weight well. [9,10] The specific results were shown in Table 2, the results of comprehensive evaluation value of each operating condition were shown in Table 3, and the trend of evaluation value was shown in Figure 1.

Table 2. Information entropy and weight of each parameter

| Parameter | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | $E_6$ | $E_7$ | $E_8$ | $E_9$ | $E_{10}$ | $E_{11}$ | $E_{12}$ |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|---------|---------|
| Entropy   | 0.545 | 0.452 | 0.553 | 0.349 | 0.546 | 0.530 | 0.565 | 0.592 | 0.643 | 0.493   | 0.560   | 0.522   |
| Weight    | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ | $w_6$ | $w_7$ | $w_8$ | $w_9$ | $w_{10}$| $w_{11}$| $w_{12}$|
### 2.4. Principal component analysis for comprehensive evaluation value

The original matrices of 5 operating conditions were calculated and analyzed according to the method introduced in 1.3. The first two eigenvalues of the matrix were $\lambda_1 = 10.659$, $\lambda_2 = 1.135$ respectively. The cumulative contribution rate of the first two eigenvalues of the correlation matrix was as high as 97.8%. The corresponding eigenvector was taken and then the corresponding principal components were calculated under each operating condition. [11,12,13]

The specific calculation results of each operating condition were shown in Table 3 and Figure 1.

| Operating conditions | Comprehensive value of information entropy | Comprehensive value of principal component analysis |
|----------------------|-------------------------------------------|--------------------------------------------------|
| 200MW                | 0.296                                     | -4.25                                           |
| 210MW                | 0.466                                     | -1.15                                           |
| 220MW                | 0.535                                     | 0.251                                           |
| 230MW                | 0.626                                     | 1.903                                           |
| 240MW                | 0.674                                     | 3.25                                            |

#### 2.5. Comparative analysis of the two methods

It can be seen from Figure 1 that the comprehensive evaluation value increased with the increase of load, which was consistent with the actual operation experience. The slope of the increasing curve of the two methods was almost the same, while the threshold value of the principal component analysis (the range of the comprehensive value) was broader. Therefore, the principal component analysis had a higher identification degree compared with the fuzzy information entropy under different operating conditions[14,15]. On the surface, principal component analysis was more appropriate in the practical application process. However, with the increase of operating points, the situation would change. In this paper, the author took 8 operating points between 200MW ~ 270mw and 11 operating points between 200MW ~ 300MW as examples, and analyzed them respectively according to the above two methods. The specific process will not be repeated. Now the final results were shown as follows:

Table 4. Eight conditions comprehensive value of the information entropy and principal component analysis
Operating conditions | Comprehensive value of information entropy | Comprehensive value of principal component analysis
--- | --- | ---
200MW | 0.380 | -3.960
210MW | 0.392 | -2.764
220MW | 0.446 | -1.655
230MW | 0.473 | -1.039
240MW | 0.571 | 0.080
250MW | 0.586 | 2.216
260MW | 0.612 | 4.017
270MW | 0.663 | 3.104

Table 5. Operating conditions comprehensive value of the information entropy and principal component analysis

From the comparison of Table 5 and Figure 2, it can be seen that the comprehensive evaluation value under 8 operating points obtained by the fuzzy information entropy was still increasing, which was consistent with the actual situation. However, the comprehensive evaluation value at 8 operating points obtained by the principal component analysis has decreased at 270MW, which was no longer a monotonic increasing function. Similarly, the comprehensive evaluation value at 11 operating points has decreased at 280MW (see Figure 3), which was no longer a monotonic increasing function. In both cases, the comprehensive value obtained by the information entropy was increasing all the time.

![Figure 2. Trend of comprehensive value at 8 operating points](image-url)
The reason for the above situation was the defect of the principal component method itself, because the method can be used only when the cumulative contribution of eigenvalues was greater than 90%. The first two eigenvalues of the principal component analysis at the 8 operating points were $\lambda_1 = 10.96$, $\lambda_2 = 1.106$, and the cumulative contribution rate (credibility) was 97.5%. The first two eigenvalues of the principal component analysis at the 11 operating points were $\lambda_1 = 10.716$, $\lambda_2 = 0.828$, and the cumulative contribution rate (credibility) was 96.2%. Although the reliability was very high, compared with the fuzzy information entropy method, there was still information loss, which made the calculation result "distorted". Then there would be a problem that the same comprehensive evaluation value corresponds to two or more operating conditions, which was not conducive to practical application.

3. Conclusions
In this paper, fuzzy theory and information entropy were introduced into the unit evaluation of power plant. With the different access of the membership function, the final comprehensive evaluation index value was also different, but the difference of the index value did not affect the order of the unit's operating conditions. The determination of the weight of each parameter was the key problem of fuzzy analysis. Some classical methods had more or less "human" subjective factors, which made the results less objective. The principal component analysis showed its superiority in higher identification when analyzing fewer operating conditions. However, with the increase of the number of operating conditions, the method would appear one-to-many phenomenon. When the selection of operating condition interval and the number of operating conditions were different, the comprehensive evaluation value would be different, so the established reference value of the thermal files under the whole operating condition would also be different.

References
[1] ZHOU ZY. (2015) Multi-fault analysis method for the steam turbine flow based on principal component analysis and neural network [D]. Beijing: North China Electric Power University.
[2] Xie JX. (2010) Study on Fault Analysis and Application of Turbine Flow in Power Plant [D]. North China Electric Power University.
[3] FU ZhG, WANG LP, GE ZH. (2008) Evaluation of operation status of power station unit by principal component analysis [J]. Journal of Power Engineering, 04: 548-551.
[4] XU ZS. (2004) Uncertain multi attribute decision making method and application [M]. Tsinghua University Press, PP.38-90.
[5] XU JP. (2006) Multi attribute decision making method and application [M]. Tsinghua University Press, PP. 20-50.
[6] MAN RY, FU ZG. (2009) Evaluation of thermal power plant based on fuzzy comprehensive evaluation [J]. Proc Chin Soc Elect Eng, 05: 5-10.

[7] JIN T, FU ZG, MAN RY. (2010) Fuzzy evaluation method for comprehensive evaluation of thermal power plant [J]. Thermal Power Engineering, 05: 473-477 + 571.

[8] Dusmanta Kumar Mohanta, Pradip Kumar Sadhu, Chakrabarti R. (2004) Fuzzy reliability evaluation of captive power plant maintenance scheduling incorporating uncertain forced outagerate and load representation[J]. Electric Power Systems Research, 72(1):73-84(in Chinese).

[9] QI MF, FU ZF, JING Y. (2013) A comprehensive evaluation method of thermal power unit based on information entropy and principal component analysis [J]. Proc Chin Soc Elect Eng, 02: 58-64.

[10] FU XQ, LIU GL, JIANG J. (2005) Based on the information entropy of the unit operating degradation degree of comprehensive index[J]. Automation of Electric Power Systems, 07: 75-78.

[11] Jiang HY, Wang WX. (2003) Application of principal component analysis in synthetic appraisal for multiobjects decision making[J]. Journal of Wuhan University of Technology. 28(3): 467-470.

[12] LI XX, FAN ZY, WANG HJ. (2016) Application of principal component analysis in optimal operation of thermal power units [J]. Power Station Auxiliary, 04:33-36.

[13] MA R, KANG R, LUO B. (2013) Method of energy consumption identification for thermal power units based on improved principal component analysis [J]. Power Grid Technology, 05: 1196-1201.

[14] FAN XL, FENG HH, YUAN M. (2013) Feature selection algorithm based on mutual information based on mutual information[J]. Control and Decision. 06:915-919.

[15] Han M. (2016) Multivariate statistical analysis: from data to conclusions [M]. Shanghai University of Finance and Economics Press, PP.68-90.