Failure Rate Model of Electric Equipment Based on Meteorological Environment

Chen Song1,*, Yan Zhao2, Kaijin Xue3 and Yucai Li2
1School of Electric Power, Shenyang Institute of Engineering, Shenyang 110136, China
2School of Renewable Energy, Shenyang Institute of Engineering, Shenyang 110136, China
3State Grid Liaoning Information and Communication Company, Shenyang, Liaoning 110006, China
*Corresponding author email: 272636881@qq.com

Abstract. With the development and progress of equipment status monitoring and fault diagnosis technology, the status maintenance based on equipment status evaluation is gradually popularized and applied. By monitoring the quantity of equipment state and diagnosing the equipment state, a reasonable method of equipment state evaluation can be used to determine whether the equipment needs maintenance or not. This paper proposes a method of power equipment failure rate model based on meteorological environment. This method firstly uses the fuzzy c-means algorithm (FCM) to conduct fuzzy clustering of meteorological environment sequences in the historical sample data and forecast meteorological sequences in the maintenance period. Then, on the basis of clustering analysis, the grey relational degree analysis is carried out on the same kind of meteorological feature sequence, so as to obtain the power equipment failure probability with meteorological conditions in the maintenance period. The results show that the proposed method can more accurately calculate the failure probability of the equipment, and it is of practical significance to reduce the influence of maintenance on the reliability of power grid operation.

1. Introduction
With the progress of the society and the development of the economy, the scale, capacity and coverage of the power system are getting larger and larger, and the power failure accidents caused by the power grid failure will inevitably cause significant economic losses to the users and the society[1]. The power equipment is the key to ensure the safe and stable operation of the power grid. Power grid and equipment operation reliability is closely related to the equipment of meteorological environment, under different meteorological conditions of the probability of equipment failure, according to the weather forecast information, combined with the historical data to determine repair period operation equipment failure probability, to improve equipment utilization and availability, extend the service life of equipment, security system security and power supply reliability has important significance[2-3]. The failure rate of electrical equipment is an important reliability index, usually used to measure the performance of a device[4]. The common methods of fault diagnosis are based on analysis model and artificial intelligence. Among them, the method based on the analytical model is based on the accurate mathematical model of the system, and describes the dynamic and steady state behavior of the original system quantitatively or qualitatively by establishing a process model of the actual system[5]. The
uncertainty of meteorological environment results in a nonlinear relationship between equipment failure rate and meteorology. As a result, it is difficult to build accurate system model[6]. However, the method based on artificial intelligence does not need to establish accurate mathematical model for the system, and is good at establishing the relationship between feature space and fault class, so the diagnosis conclusion is easy to understand. This method mainly includes expert system, neural network method and fuzzy clustering fault diagnosis method[7]. The expert system simulates the reasoning of the expert in fault diagnosis to form a hybrid intelligent fault diagnosis system. However, its disadvantages lie in its strong dependence on expert knowledge and poor learning ability, and it can't deal with uncertain problems well. The use of artificial neural network does not require the operator to further understand the knowledge of expert domain, and it can be learned by using fault diagnosis examples. Its shortcoming is that the reasoning process cannot be fully explained and lacks transparency. Fuzzy clustering is a method to describe process change with fuzzy sets. It has a good advantage in fault diagnosis when system state and fault state are uncertain. The more mature fuzzy clustering method is the fuzzy C clustering algorithm, which successfully introduces the fuzzy concept into the equipment fault classification process, makes the target classification more accurate, and further improves the accuracy of the algorithm compared with other algorithms. In this paper, fuzzy C clustering algorithm is used to study the model of equipment failure rate[8].

Although FCM is able to in the heart of the uncertain relationship between meteorological environment and clustering data accurately, but the method in the determination of optimum clustering number is relatively difficult, clustering partition easily affected by the data distribution, do not produce uniform distribution of classes, in access to equipment failure rate can be more accurate estimate, at the same time can't more effectively solve the problem of "failure to gather" phenomenon, so I need combines FCM and grey correlation degree to use[9]. Therefore, this paper proposes a new method for fault diagnosis of power equipment:

- The FCM is used to determine the C optimal clustering centers of the entire fault space, so as to construct the state mode vector for diagnosing the fault of power equipment, which is used as the standard reference sequence of the grey relational degree model.
- The model was constructed with grey relational degree, and the meteorological data was used as the input and the relational degree coefficient as the output, which was compared with the traditional grey relational degree algorithm.
- Analyze the results and verify the feasibility of the method.

2. FCM Determines the Fault Standard Sequence of Power Equipment

2.1. FCM Basic Algorithm

According to relevant practical experience and theoretical literature, when electric equipment is working, it is affected by both electricity and heat, and the biggest factor is meteorological environment. Therefore, wind speed, temperature, humidity and rainfall were selected as the characteristic parameters, and the optimal cluster generated by FCM clustering was used as the standard sequence for fault diagnosis of electric equipment[10].

Let $U=(u_{ik})_{n \times c}$, is the fuzzy classification matrix, where $n$ represents the number of samples, $c$ represents the number of categories, and $u_{ik}$ represents the membership degree of the $i$th sample belonging to the $K$ classification. Let's say $X=\{x_1, x_2, \ldots, x_n\}$ is the set of classified samples, in which each sample $x_i$ has m characteristic indexes, namely $x_i=\{x_{i1}, x_{i2}, \ldots, x_{im}\}$. $P$ is used to represent the clustering center of $c(2 \leq c \leq n)$ subsets, and the optimization objective function of FCM is:

$$J_m(U,P) = \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij})^m d_{ij}^2$$

where $J_m$ represents the sum of squares of the distance between each data instance and the class center. $d_{ij}$ is a distance measurement function, and its selection must be determined based on the clustering distribution pattern that needs to be solved in practice, which is usually Euclidean distance. $m$
represents a fuzzy coefficient, which is also called the smoothing factor, and is used to restrict the fuzzy level of membership matrix $U$. The larger the $m$ value, the higher the fuzzy level.

2.2. FCM Iteration Steps

FCM is to obtain $\min \{J_m(U,P)\}$ of the objective function $J_m(U,P)$ under the condition of satisfying the constraint conditions, and obtain the optimal solution of the algorithm by obtaining the optimal clustering center matrix \( P = \{ p_i \}_{1 \leq i \leq c} \) and partition matrix \( U = (u_{ik})_{n \times c} \).

Because the columns in the matrix $U$ are independent, there are:

$$
\min \left\{ J_m(U, P) \right\} = \min \left\{ \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ik})^m (d_{ij})^2 \right\}
$$

(2)

Where the constraint conditions of the extreme value of the above equation are as follows:

$$
\sum_{i=1}^{c} \mu_{ik} = 1
$$

(3)

FCM algorithm achieves fuzzy classification of the data set through iterative optimization of the objective function $J_m(U,P)$, namely, iteration:

$$
\mu_{ik} = \frac{1}{\sum_{j=1}^{n} \left[ \frac{d_{ij}}{d_{jk}} \right]^{m-1}}
$$

(4)

$$
p_i = \frac{\sum_{j=1}^{n} (\mu_{ik})^m x_j}{\sum_{j=1}^{n} (\mu_{ik})^m}
$$

(5)

The specific steps are as follows:

- Initialization set the number of clustering categories as $c (2 \leq c \leq n)$, where $n$ represents the number of samples), and the membership degree matrix $U$ is initialized.

- Update the clustering center matrix $P$ according to equation (6):

$$
p_i^{(b)} = \frac{\sum_{j=1}^{n} (\mu_{ik}^{(b)})^m x_j}{\sum_{j=1}^{n} (\mu_{ik}^{(b)})^m}
$$

(6)

- According to formula (7), the membership matrix $U$ is updated:

$$
\mu_{ik}^{(b+1)} = \left\{ \frac{\sum_{j=1}^{n} \left[ \frac{d_{ij}^{(b+1)}}{d_{jk}^{(b+1)}} \right]^{m-1}}{\sum_{j=1}^{n} \left[ \frac{d_{ij}^{(b+1)}}{d_{jk}^{(b+1)}} \right]^{m-1}} \right\}
$$

(7)

- Determine whether $(U^{(b)} - U^{(b+1)}) < \varepsilon$ is true, if so, the algorithm stops, and the output matrix $P$ and $U$. Otherwise, let $b = b + 1$ go to step 2 to continue executing the iteration command.

It can be seen that FCM algorithm is the classification process of repeatedly modifying membership degree matrix and clustering central matrix, and can converge from any initial point along an iterative sub-sequence to a local minimum of its objective function $J_m(U,P)$, so as to obtain an optimization score.
3. Basic Grey Relational Degree Model of Meteorological Environmental Factors of Power Equipment Failure

3.1. The Overall Approach

After obtaining the new meteorological data, the new meteorological sequence was classified into the historical operation data through FCM, and the correlation degree between the meteorological sequence and each meteorological sequence in the class was calculated by combining the grey correlation analysis algorithm. According to the grey process of the system, the grey correlation analysis quantifies the dynamic change of the system, studies the process of data mapping to data, and can judge and process the random variables well. Through the analysis of the changes in the geometric shapes of the curves, the more similar the shapes of the curves are, the greater the correlation between different factors of the system will be. The more similar the shapes of the curves will be, the more consistent the change trend of the system will be. The specific process is shown in figure 1:

![Diagram](image)

Figure 1. Solving process of failure rate of electric equipment based on meteorological factors.

3.2. Model Building

By analyzing the relationship between different influencing factors and the number of fault tripping, this paper determines the influence degree of different influencing factors on line fault tripping, that is, the most important influencing factors.

- Establish meteorological reference sequence and comparison sequence
  
  The reference sequence selected in this paper is the total number of line fault tripping of a power grid company in recent years. The comparison sequence is the number of fault tripping caused by wind speed, temperature, relative humidity, rainfall and other factors.

- The parameters are dimensionless
  
  Due to the different dimensions of different influencing factors in the system, the analysis results may not be correct, so the data of different parameters need to be processed dimensionless in the process of parameter correlation analysis.

\[
X_{jh} = \frac{X_{jh}}{X_j}
\]  

(8)

Where h corresponds to the time series and j to a feature in the comparison sequence

- Calculate the correlation coefficient
  
  The correlation coefficient is the correlation degree value of the comparison sequence and the reference sequence at each moment, and the calculation formula is as follows:

\[
\xi_{jh} = \frac{\min_{h \in \text{time}} \min_{k \in \text{com}} |X_{0h} - X_{jk}| + \rho \max_{h \in \text{time}} \max_{k \in \text{com}} |X_{0h} - X_{jk}|}{\max_{h \in \text{time}} \max_{k \in \text{com}} |X_{0h} - X_{jk}|}
\]

(9)

Where \( \rho \) is the resolution coefficient, the smaller \( \rho \) is, the larger the corresponding resolution is, generally \( \rho \in (0,1) \), usually \( \rho = 0.5 \).

- Calculated correlation
Because at different time series points, the correlation degree between the comparison sequence and the reference sequence is different, the correlation degree information is scattered. In order to accurately evaluate the correlation degree between the reference sequence and the comparison sequence, the average value of the correlation coefficient representing the same characteristic parameter at different time nodes is generally calculated as the correlation degree between the different comparison sequences and the reference sequence. The calculation formula of the correlation degree $r_j$ is as follows:

$$r_j = \frac{1}{n} \sum_{i=1}^{n} \xi_{ij}$$  \hspace{1cm} (10)

Where $\xi_{ij}$ represents the correlation coefficient of each moment of the same parameter.

4. The Example Analysis

In this paper, a probability model of power equipment failure based on meteorological forecast information is established, and the maximum wind speed, average temperature, average relative humidity and rainfall in unit time are selected as the characteristic parameters of meteorological environmental information. Based on the statistics of the historical operation data of a power supply line in a city, 10 groups of data sequences are selected for analysis, as shown in table 1.

| The time series | Maximum wind speed (m/s) | The average temperature (℃) | Mean relative humidity (%) | Rainfall (mm) |
|-----------------|--------------------------|-----------------------------|---------------------------|-------------|
| X₁              | 10.2                     | 25.1                        | 72                        | 71.1        |
| X₂              | 5.3                      | 24.4                        | 78                        | 32.4        |
| X₃              | 4.2                      | 32.4                        | 86                        | 19.7        |
| X₄              | 4.4                      | 30.6                        | 83                        | 117.6       |
| X₅              | 9.1                      | 24.8                        | 80                        | 75.2        |
| X₆              | 3.4                      | 30.3                        | 88                        | 25.4        |
| X₇              | 4.9                      | 32.6                        | 90                        | 29.7        |
| X₈              | 8.1                      | 24.7                        | 80                        | 82          |
| X₉              | 9.0                      | 26.3                        | 77                        | 86.3        |
| X₁₀             | 3.7                      | 31.5                        | 87                        | 23.1        |

Firstly, the meteorological characteristic data are standardized, and the data are standardized between [0,1] according to equation (8), as shown in table 2.

| The time series | Maximum wind speed (m/s) | The average temperature (℃) | Mean relative humidity (%) | Rainfall (mm) |
|-----------------|--------------------------|-----------------------------|---------------------------|-------------|
| X₁              | 1.0000                   | 0.0629                      | 0.0000                    | 0.5125      |
| X₂              | 0.2576                   | 0.0000                      | 0.3529                    | 0.1254      |
| X₃              | 0.0909                   | 0.9790                      | 0.7059                    | 0.0000      |
| X₄              | 0.1212                   | 0.8671                      | 0.6471                    | 1.0000      |
| X₅              | 0.8636                   | 0.0490                      | 0.4706                    | 0.5547      |
| X₆              | 0.0000                   | 0.8531                      | 0.9412                    | 0.0592      |
| X₇              | 0.2121                   | 1.0000                      | 1.0000                    | 0.0993      |
| X₈              | 0.7121                   | 0.0420                      | 0.4118                    | 0.6239      |
| X₉              | 0.8485                   | 0.0699                      | 0.2941                    | 0.6690      |
| X₁₀             | 0.0455                   | 0.9301                      | 0.8824                    | 0.0341      |

After the meteorological characteristic parameters are standardized, the variation trend of the correlation coefficient of tripping fault caused by the serious deterioration of each parameter relative
to the total fault of the line is calculated and compared with the traditional method, as shown in the figure.

Figure 2. Diagram of correlation coefficient of wind speed.

Figure 3. Diagram of temperature correlation coefficient.

Figure 4. Diagram of humidity correlation coefficient.

Figure 5. Diagram of correlation coefficient of rainfall.

It can be seen from the figure above that the correlation coefficient of transmission line trip failure caused by wind speed, humidity and rainfall is relatively high at each time node, while that of other related factors is generally relatively low at different time nodes. The average method is used to calculate the correlation degree of each influencing factor, as shown in the figure:

Figure 6. Correlation degree of different factors.

Thus, the correlation degree between transmission line tripping faults caused by wind speed in this region and total tripping faults is $r=0.8613$, that is, the influence of wind speed on transmission line tripping faults in this region is the most significant and closely related. At the same time, it can be seen
that the correlation coefficient between humidity and rainfall and the total number of transmission line fault trip is also higher in different time series. In this paper, the prediction error of the improved grey relational analysis method is within the allowable range, and the accuracy is increased by nearly 5% compared with that of the unimproved grey relational analysis method, which verifies the effectiveness of the improved method proposed in this paper. In the actual site, operation and maintenance personnel can predict the line failure rate according to the weather forecast data in the future for the next step of prevention and maintenance work.

5. Conclusion
In this paper, based on the known historical operation data of transmission lines, an improved grey correlation analysis algorithm is proposed to establish the equipment failure rate model. Finally, an example is given to show that the prediction accuracy of the improved method is improved compared with that of the grey relational analysis method. At the same time, the selected meteorological influencing factors are analyzed to determine the degree of correlation between the tripping fault caused by different influencing factors and the total tripping fault of the transmission line. The results show that the influence of wind speed, humidity and rainfall on the fault trip of transmission lines in this region is high, and it can be considered as the key factors in the construction and monitoring of transmission lines. The prediction model can provide reference for the operation decision of power system, and the prediction method will obtain more accurate prediction results with the increase of the number of transmission line historical operation samples.

Acknowledgements
This paper was supported by the Key R&D Program of Liaoning Province (2018220017, 2019JH8/1010062, 2019JH8/1010066).

References
[1] Y Ren and H Zhu 2019 A General Inversion Method Based on Magnetic Flux Leakage Inspection Data Driven Control and Learning Systems Conference.
[2] F M Qu 2019 Wind Turbine Condition Monitoring Based on Assembled Multidimensional Membership Functions Using Fuzzy Inference System IEEE Transactions on Industrial Informatics.
[3] A A Bulatov and I K Andronchev 2017 Combined approach to an assessment of maintenance of electrical equipment on traction rolling stock Russian Electrical Engineering.
[4] R Vinayagamoorthy and M A Xavior 2014 Parametric Optimization on Multi-Objective Precision Turning Using Grey Relational Analysis Procedia Engineering.
[5] H Marko 2019 Research of the optimal variable defects of the preventive maintenance of medicinal equipment Tehnički Glasnik.
[6] F W Zhang and S H Xu 2017 Remarks to “Fuzzy multicriteria decision making method based on the improved accuracy function for interval-valued intuitionistic fuzzy sets Soft Computing.
[7] B Ismat and R Tabasam 2017 A Clustering Algorithm Based on Intuitionistic Fuzzy Relations for Tree Structure Evaluation International Journal of Applied and Computational Mathematics.
[8] J Sathiamoorthy and B Ramakrishnan 2017 A Reliable Data Transmission in EAACK MANETs Using Hybrid Three-Tier Competent Fuzzy Cluster Algorithm Wireless Personal Communications.
[9] H Ayyaz and H Muhammad 2017 A new cluster based adaptive fuzzy switching median filter for impulse noise removal Multimedia Tools and Applications.
[10] Y Li 2017 A reconstructed variable regression method for thermal error modeling of machine tools The International Journal of Advanced Manufacturing Technology.