Research on the Method of Correlation Analysis Between Reliability Indices and Business Management Indices

CAO HUAZHEN1,2, WU YAXIONG1,2, JIN YEXUAN3, ZHANG JUNXIAO1,2, LIU YONG3
1.2. Guangzhou 501180, 3. Shanghai 200240
CHINA
jinyexuan777@163.com

Abstract: - The business management indices can present the most of the influence factors of reliability. In addition, business management indices are monitored and controlled by respective departments. It is possible to achieve the given annual reliability indices by controlling business management indices monthly. In this paper, the correlation between the reliability indices and the business management indices is studied. A comprehensive reliability evaluation index system based on business management indices is proposed, in which an improved grey relational model is used to mine the correlation between the reliability and factors. The simulation results prove the effectiveness of the method.

Key-Words: - reliability indices, business management indices, grey relational degree, dynamic resolution coefficient, fuzzy weighting

1 Introduction
The basic task of the distribution network is to provide users with safe, stable, and high-quality power. The distribution network has complex structures with a large number of devices and appliances, so that power supply reliability is the most important problem in distribution network. According to statistics, about 80% of power outage is caused by the faults of distribution network components [1]. Therefore, the research on the distribution network reliability is one of the core researches in the field of distribution.

Analytical method and simulation method are two major methods for reliability assessment. The typical simulation method is the Monte Carlo method, which is time consuming, low precision, and the scale of distribution network has little effect on the computational complexity. The analytical method, represented by the FMEA method, can realize the accurate calculation of the reliability indices of the complex distribution network. As both methods require the complete topology and equipment parameters of the distribution network, which cause too much calculation work when the distribution network has a large size, some rapid reliability assessment method are proposed.

With the continuous development of intelligent distribution, the data acquired, transmitted, and applied in various distribution management systems contain abundant information of operation, dispatch, maintenance, and marketing, which can provide the basis for distribution network analysis [2-4]. Studies on relationship between the network structure and reliability have been carried out. Principal component analysis and parallel association rule mining techniques are used to mine association relationships, and the future reliability indices are predicted based on artificial neural network in [5]. A method to use BP neural network to derive the analytic relation expression between reliability and driving factors is proposed in [6]. The grey correlation degree is used to obtain the correlation between reliability and network structure factors in [7]. A stepwise regression model is established to obtain the mathematical expression of the indicator association in [8]. This sort of method belongs to the data-driven method. It has high speed and extensive applicability because it does not rely on topology analysis.

The business management indices, such as line link rate, cable rate, and so on, can present the most of the influence factors of reliability involves network structure, equipment parameters, and management level. In addition, business management indices are monitored and controlled by respective departments. It is possible to achieve the given annual reliability indices by controlling business management indices monthly. In this paper, the relationship between the reliability indices and the business management indices is studied, which can find out the weakness in the business process. A comprehensive reliability
evaluation index system based on business management indices is proposed. An improved grey relational model is used to mine the correlation between the reliability of A+, A, B, C, D power supply area [9] and factors in the above reliability evaluation index system. The simulation results prove the effectiveness of the method.

2 Establishment of Comprehensive Reliability Evaluation Index System Based on Business Management Indices

The comprehensive reliability evaluation is intended to evaluate the level of technology and management of the power supply enterprises. The power supply company is a technology intensive enterprise, which involves planning, operation, maintenance, scheduling, marketing and other business. Different businesses have their own work processes. In addition to the level of power grid equipment, the professional skill of staff in each business process will also have an impact on the reliability of power supply. Through a comprehensive evaluation of the equipment and management level of the distribution network, the company looks for the influence factors affecting the reliability from all aspects of the business, the deficiencies in the reliability management, and improves the staff’s professional skill in order to improve the reliability and management of the distribution network [10].

Constructing a reliability evaluation index system is the basis of reliability evaluation of distribution network. By analyzing the business process the affects the reliability, the management of each business work is quantified in the form of index to establish the comprehensive evaluation index system.

Based on the statistical data of management information system in recent years, this paper analyzes the business management and equipment factors that affect the reliability in four aspects like network frame, operation and maintenance, production scheduling, and construction operation. The various business processes and equipment factors that affect the reliability are analyzed and quantified in the form of index to constitute a comprehensive reliability evaluation index system. The comprehensive index system is divided into three levels, including evaluation indicators, business stratification indicators, and factor indicators. Among them, the first level is the evaluation index which represents power supply reliability. The second level are made up of four indicators, network level, operation and maintenance, production scheduling, and construction operation. The third level is enumerated according to the jurisdiction scope of the second layer. For example, the live working is related to the operation and maintenance, so the factor of live working rate is put into the operation and maintenance index. The comprehensive index system is shown in Fig.1.

3 Correlation Analysis of Reliability and Business Management Indices

Correlation analysis, regression analysis, variance analysis and principal component analysis can be used in mining association information, but the application of these methods requires a large number of samples, no correlation between the factors, and a certain distribution law. The grey relational analysis, measuring the proximity of each factor according to the proximity of the shape of the sequence curve of each factor, is a quantitative analysis method used to distinguish the complex relationship between variables and multiple factors [11]. The greater the degree of grey correlation is, the more the relative changes in the representative factors are in the development process. Because the grey correlation analysis is based on the analysis of the development situation, the method does not have much requirement on the number of samples, nor does the sample obey the typical distribution. The results are in good agreement with the qualitative analysis results. Therefore, the grey correlation analysis is used to calculate the correlation between factor index and reliability index.

2.1 Improved Grey Correlation Analysis
The traditional grey relational analysis method determines the proximity between elements $x_i$ of the comparison sequence $X_i = (x_{i1}, x_{i2}, \cdots, x_{in})$ and the elements $x_{0j}$ of the reference sequence $X_0 = (x_{01}, x_{02}, \cdots, x_{0n})$ by calculating the grey correlation coefficient $\zeta(x_{0j}, x_{ij})$ which is calculated as follows.

$$\zeta(x_{0j}, x_{ij}) = \frac{\min_{i} \Delta_{ij} + \rho \max_{j} \Delta_{ij}}{\Delta_{ij} + \rho \max_{j} \Delta_{ij}}$$  \hspace{1cm} (1)

Where, $i(i=1,2,\cdots,m)$ is the number of influencing factors sequence, $j(j=1,2,\cdots,n)$ is the dimension of each influencing factor sequence, $\rho \in [0,1]$ is the resolution coefficient, $\Delta_{ij} = |x_{i0} - x_{ij}|$.

The grey correlation coefficient $\zeta(x_{0j}, x_{ij})$ can be used to calculate the grey correlation between the comparison sequence $X_i$ and the reference sequence $X_0$.

$$\gamma_0 = \frac{1}{n} \sum_{j=1}^{n} \zeta(x_{0j}, x_{ij})$$  \hspace{1cm} (2)

It can be seen from formula (1) that the product of $\rho$ and $\max_{j} \Delta_{ij}$ has a greater influence on the value of the entire formula. The value of the resolution coefficient $\rho$ reflects the indirect influence of $\max_{j} \Delta_{ij}$ on the correlation. The default setting $\rho$ of the traditional grey correlation analysis is 0.5, which will result in the smaller correlation degree distribution, and the reduction of the degree of the correlation between the factors. Therefore, the dynamic resolution coefficient should be introduced into the traditional grey correlation analysis.

In reality, the importance of data varies from time to time. The closer the time, the greater the correlation with the current data. However, in the traditional grey relational analysis, it can be seen from formula (2) that the degree of grey correlation is obtained by calculating the mean value of the correlation coefficient at each different period. The treatment of averaging does not reflect the inconsistency in the degree of association. Therefore, the weights are introduced in the improved grey correlation analysis to reflect the importance of different historical data to current conditions [12].

2.1.1 Determination of segmentation dynamic resolution

The formula for calculating the correlation degree coefficient is obtained as follows,

$$\zeta(x_{0j}, x_{ij}) = \frac{\min_{i} \Delta_{ij} + \rho \max_{j} \Delta_{ij}}{\max_{i} \Delta_{ij} + \rho \max_{j} \Delta_{ij}}$$  \hspace{1cm} (3)

When the sequence of influencing factors is not abnormal and stable, $\max_{j} \Delta_{ij} >> \Delta_{ij}$ and $\max_{j} \Delta_{ij} >> \min_{i} \Delta_{ij}$ will appear. In this case, if the resolution coefficient $\rho$ is larger, $\zeta(x_{0j}, x_{ij})$ will be dominated by $\rho$, and its value will be close to 1, which makes it difficult to reflect the effect of $\Delta_{ij}$. Therefore, the range of the grey correlation degree obtained by the traditional grey correlation analysis is narrow, and the wrong conclusion that the reference sequence and each comparison sequence are highly correlated will be obtained. In the case, the resolution factor $\rho$ should be smaller, thereby reducing the dominant effect of $\max_{j} \Delta_{ij}$ on $\zeta(x_{0j}, x_{ij})$.

When the sequence of influencing factors is not abnormal and stable, $\max_{j} \Delta_{ij}$ and $\Delta_{ij}$ are close to each other. If the resolution coefficient $\rho$ is smaller, $\zeta(x_{0j}, x_{ij})$ will be smaller. Therefore, the grey correlation degree of each influencing factor will be very similar, and the distribution interval is small, which make it difficult to distinguish the correlation degree between the reference sequence and each comparative sequence. In this case, the value of $\rho$ should be increased as much as possible to increase the dominant effect of $\max_{j} \Delta_{ij}$ on $\zeta(x_{0j}, x_{ij})$, so the integrity of the correlation degree can be fully reflected.

Therefore, it is necessary to establish a dynamic resolution coefficient according to the smoothness of the sequence to obtain the correlation degree between the sequences more accurately. Define $\lambda_i$ as the sequence smooth coefficient, which is calculated by the flowing formula,
\[ \lambda_i = \frac{\sum_{j=1}^{n} \Delta_{ij}}{\max_j \max_i \Delta_{ij}} \]  

As \( \rho_i \in (0,1] \), using \( \rho_i = 0.5 \) as the boundary value, the value of \( \rho_i \) can be determined as follows [13],

1) When \( \lambda_i < 1/3 \) (the data sequence is not stable), \( \rho_i \) should take a smaller value less than 0.5 to reduce the dominance of \( \max_i \Delta_{ij} \) on the degree of grey correlation. So \( \rho_i \) should be in \( [\lambda_i, 1.5 \lambda_i] \), generally taking \( \rho_i = 1.5 \lambda_i \).

2) When \( \lambda_i \geq 1/3 \) (the data sequence is stable), in order to increase the distinction between sequences, \( \rho_i \) should be in \( [1.5 \lambda_i, 2 \lambda_i] \), generally taking \( \rho_i = 2 \lambda_i - 1/6 \).

By using segmented dynamic resolution coefficient instead of static resolution coefficient, a new correlation coefficient formula based on dynamic resolution coefficient is obtained as follows,

\[ \zeta(x_{0j}, x_{ij}) = \frac{\min_j \min_i \Delta_{ij} + \rho_i \max_j \max_i \Delta_{ij}}{\Delta_{ij} + \rho_i \max_j \max_i \Delta_{ij}} \]  

2.1.2 Time weighting based on fuzzy matrix

The calculation formula for the horizontal weighted correlation degree between the \( i^{th} \) influencing factor sequence and the reliability level sequence \( y_{oi} \) is

\[ y_{oi} = \sum_{j=1}^{n} \omega_j \zeta(x_{0j}, x_{ij}) \]

Where, \( \omega_j \) is horizontal weighting factor.

The horizontal weighting coefficient is calculated as follows [14].

1) According to the principle that the closer the time of the data is, the greater the degree of correlation is and the more important the data is, the fuzzy complementary priority relationship matrix \( F = (f_{h_{j1}})_{n \times n} \) in historical periods \( j_1 \) and \( j_2 \) can be formed. \( f_{h_{j1}} \) indicates that the data of the historical point \( j_1 \) is more important than the data of \( j_2 \). When \( j_1 > j_2 \), it means that the data of the historical point \( j_1 \) is more important than that of \( j_2 \), so that \( f_{h_{j1}} = 1 \). On the contrary, when \( j_1 < j_2 \), it means that the data of the historical point \( j_1 \) is less important than that of \( j_2 \), so that \( f_{h_{j1}} = 0 \). When \( j_1 = j_2 \), the data of the historical point \( j_1 \) and \( j_2 \) have the same degree of importance to the current situation, so that \( f_{h_{j1}} = 0.5 \).

2) Sum the lines of the matrix \( F = (f_{h_{j1}})_{n \times n} \) to get

\[ s_h = \sum_{j=1}^{n} f_{h_{j1}}, j = 1,2,\cdots,n \]

The fuzzy consensus matrix \( S = (s_{h_{j1}})_{n \times n} \) is calculated, where

\[ s_{h_{j1}} = \frac{s_h - s_{h_{j1}} + 0.5, j = 1,2,\cdots,n}{2(n-1)} \]

3) Calculating horizontal weight \( \omega_j \)

\[ \omega_j = \frac{1}{n} - \frac{1}{2a} + \frac{1}{na} \sum_{s=1}^{n} s_{h_{j1}}, j = 1,2,\cdots,n \]

Where, \( a \geq \frac{n-1}{2} \).

2.2 Correlation degree algorithm of reliability and business management indices

In this paper, an improved grey correlation analysis method is used to calculate the correlation between reliability and business management indices, which steps are as follows.

1) Data trend processing.

The grey correlation degree measures the consistency of the development trend of factors. When both factors are negatively correlated, the trend of the curve is the opposite, which will lead to a small value of grey correlation. Therefore, in this paper, the comparison sequence of the data trend and the reference sequence is compared. If their trend are opposite, the data of comparison sequence should be reversed to adjust the trend to the same as the reference sequence.

2) Data standardization processing

Standardize influencing factors sequence \( H_i = (h_{j1}, h_{j2},\cdots,h_{jn}) \) and reliability
indicators sequence \( H_0 = (h_{i01}, h_{i02}, \ldots, h_{ij}, \ldots, h_{ioa}) \) that have undergone trend processing [15].

\[
X_j = \frac{h_{ij} - \min h_{ij}}{\max h_{ij} - \min h_{ij}} \quad X_{0j} = \frac{h_{0j} - \min h_{0j}}{\max h_{0j} - \min h_{0j}}
\]

3) Determination of the dynamic resolution coefficient

\[ \Delta_j = |X_{0j} - X_j| \] can be calculated by \( x_j \) and \( x_{0j} \). The sequence stability coefficient \( \lambda_i \) can be obtained according to equation (4), and determine the value of the dynamic resolution coefficient \( \rho_j \) by the value of \( \lambda_i \).

4) Grey correlation coefficient calculation

Using the obtained \( \Delta_j \) and \( \rho_j \), calculate the grey correlation coefficient \( \zeta \left( x_{0j}, x_j \right) \) according to formula (5).

5) Horizontal weighting factor calculation

According to the time-weighting method based on fuzzy matrix, the fuzzy consensus matrix \( S = (s_{ij})_{axa} \) can be calculated according to equation (7) and (8). Then the horizontal weighting factor \( \omega_j \) can be calculated according to formula (9).

6) Correlation coefficient calculation

Using the grey correlation coefficient \( \zeta \left( x_{0j}, x_j \right) \) obtained in step 4 and the horizontal weight factor \( \omega_j \) obtained in step 5, the horizontal weighted correlation degree \( \gamma_{ji} \) between the sequence of the first influencing factor and reliability can be calculated according to equation (6).

3 Case Study

This paper takes an example from a coastal area as case analysis. The region provided a total of 8 factor indicators for the 2013-2016 years of class A, B, C, and D power supply areas, including cable rate \( X_1 \), distribution automation rate \( X_2 \), line link rate \( X_3 \), average number of segments \( X_4 \), average line load rate \( X_5 \), average distribution transformer load rate \( X_6 \), average repair time \( X_7 \), live working ratio \( X_8 \). The data is shown in Table 1.

Table 1 Factor data of each power supply area in 2013-2016

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>2014</td>
<td>2015</td>
<td>2016</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>31.04</td>
<td>38.07</td>
<td>45.47</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0</td>
<td>0</td>
<td>1.51</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>56.51</td>
<td>58.61</td>
<td>62.49</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>2.93</td>
<td>3.34</td>
<td>3.99</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>21.77</td>
<td>24.57</td>
<td>27.96</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>34.53</td>
<td>34.26</td>
<td>35.47</td>
</tr>
<tr>
<td>( X_7 )</td>
<td>318</td>
<td>302</td>
<td>294</td>
</tr>
<tr>
<td>( X_8 )</td>
<td>0.86</td>
<td>5.26</td>
<td>10.15</td>
</tr>
</tbody>
</table>

Firstly, according to the determination method of the segmented dynamic resolution coefficient, the resolution coefficient is determined by the degree of smoothness of the influencing factor sequence. The resolution coefficients are shown in Table 2.

Table 2 Resolution coefficient of each influencing factor sequence

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>0.3356</td>
<td>0.0551</td>
<td>0.2785</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0.1319</td>
<td>0.3259</td>
<td>0.7756</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>0.4329</td>
<td>0.2085</td>
<td>0.3583</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>0.4329</td>
<td>0.2458</td>
<td>0.2773</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>0.4456</td>
<td>0.2105</td>
<td>0.4306</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>0.7196</td>
<td>0.758</td>
<td>0.8333</td>
</tr>
<tr>
<td>( X_7 )</td>
<td>0.2314</td>
<td>0.0267</td>
<td>0.0851</td>
</tr>
<tr>
<td>( X_8 )</td>
<td>0.0764</td>
<td>0.0942</td>
<td>0.2689</td>
</tr>
</tbody>
</table>

Then, according to the fuzzy matrix weighting method, the weights of the historical period are obtained. Table 3 shows the weight coefficients.

Table 3 The weights of historical period

<table>
<thead>
<tr>
<th>Year</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>0.0833</td>
<td>0.1944</td>
<td>0.3056</td>
<td>0.4167</td>
</tr>
</tbody>
</table>

According to the improved grey relational degree model, the grey relational degree between the influencing factor index and the reliability index is calculated. The results of the correlation degree are shown in Table 4.

Table 4 Gray correlation degree between each influencing factor and SAIDI

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>0.7184</td>
<td>0.7119</td>
<td>0.7233</td>
</tr>
</tbody>
</table>
same time, the ability of the maintenance team construction of distribution automation, and at the same time, the ability of the maintenance team should be improved to shorten the repair time. In B-class power supply area, the average number of segments and the average repair time have a great influence on SAIDI. Therefore, the enterprises in charge of such area should give priority to upgrading the grid structure and pay more attention to the fault repair procedure optimization. In the D-class power supply area, the line link rate and the average number of segments have a great influence on SAIDI. Therefore, enterprises in charge of such area should pay attention to increasing the average number of segments and line link rate.

From Table 4, it can also be seen that the correlation between power supply reliability and line link rate in A-class region is lower than that in C-class and D-class regions. In addition, the correlation between the power supply reliability and distribution automation is higher than line link rate in A-class region. This is because the line link rate of the A-class region during 2013-2016 is higher than that of the C-class and D-class regions, and a small increase in the line link rate will have little impact on the improvement of power supply reliability. However, in 2013-2016, the distribution automation rate in the A-class region is generally not high, and there is a large space for improvement, which has a great impact on the reliability of power supply. The above analysis verifies the rationality of the method from the side.

### 4 Conclusion

This paper proposes a reliability mining algorithm based on business management indicators. The main conclusions are as follows.

1) Through the analysis of business processes, a three-tier comprehensive reliability evaluation index system is established. The system includes four aspects, grid level, operation and maintenance, production scheduling, and construction operation, which fully reflects the technology and equipment level of power grid and the management level of power supply business.

2) An improved grey relational degree model is adopted in this paper. Compared with the traditional grey relational model, the disadvantages of the set value of the resolution coefficient and the importance difference of the historical data are taken into consideration, and the dynamic segmentation resolution coefficient and fuzzy weighting are introduced to improve it. The result of the improved model is more realistic and provides more accurate guidance for the business management and reliability improvement.

3) Compared with the traditional reliability calculation method, this method does not depend on the topology of the power grid and does not require a large number of calculations. It can directly analyze the main factors that affect the reliability of the power supply.

### References:


[5] Yu Qing Feng, Jian Hua Yang, Lei Huang, Bin Ji, Jian Su, Intellectualization and Reliability


