

Comprehensive Evaluation to Distribution Network Planning Schemes Using Principal Component Analysis Method

Wang Ruilian¹, Gao Shengjian²

¹School of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou, Henan, China

²School of Civil Engineering and Communication, North China University of Water Resources and Electric Power, Zhengzhou, China

Abstract

This paper proposes a new comprehensive evaluation method by nonlinear principal component analysis in allusion to the problem in distribution network planning, such as a large number of factors, intense fuzziness and the nonlinear relationship between the factors. Firstly, according to pre-existing achievement and considering long-range development of distribution network, the comprehensive evaluation index system taking into account the technical, economic, environmental and adaptability factors and so on is constructed. Secondly, by disposing all of the factors in the index system using fuzzy consistent matrix model to get the relative membership degree matrix of every scheme, the initial data matrix in fuzzy nonlinear principal component analysis is obtained. Thirdly, the weight of every principal component is acquired by means of the entropy conception. The superiority of the method in the paper introduced is the method can dispose the fuzzy problem in complicated planning, simplify the computation process in comprehensive evaluation, reduce the subjectivity and arbitrariness, and can make the conclusion more scientific and more reasonable. A combination of a flow diagram based and a concrete example, the algorithm is proved to be correctly and practically.

Keywords: *distribution network planning, comprehensive evaluation, principal component analysis, fuzzy consistent matrix, entropy weight*

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

Linking power network and users, distribution network reflect directly the user's requirements in the security, reliability, utility and other aspects. The scientific and rational distribution networks planning play an important role in the power network construction, improvement and operation. So the comprehensive evaluation in distribution network planning taking into account lots of factors such as technology, economic, environment and the adaptability to future development and others, especially in the uncertainty environment and adaptability factors, is very difficult.

At present, the comprehensive evaluation methods for distribution network planning include AHP, entropy decision, group decision-making and data envelopment analysis and so on, as well as the integrations of two or more methods [1-5]. For example, a combination of expert subjective experience and objective entropy to get comprehensive weight used for scheme optimization [1-2], preference order and traditional AHP to establish comprehensive evaluation model [3], group decision-making theory and AHP to evaluate comprehensively distribute network [4], data envelopment analysis method and AHP to evaluate planning scheme of distribution network [5], all of that methods inseparable from the experts experience are difficult to deal with the miscellaneous index and for the too high experts experience and technical requirements have large of computing in distribution network planning.

Principal components analysis method, by weighting and synthesizing several less uncorrelated variables that is linear combination of more original indexes to obtain the final conclusion, can simplify the analysis system structure, reduce the calculated amount and keep the information of original variables as much as possible[6]. But the method not considering the nonlinear relationship between the original indexes is not accord with the practical engineering.

On the basis of traditional principal component analysis method, nonlinear principal component with good development and application, can reflect the nonlinear relationship between the original variables, extract the high-orders relations from original variables, also can reduce the conclusion dependencies on original variable. However, the principal component analysis is hard to qualitative indexes with vague language in the complex distribution network evaluation system, and the weights of every principal component using the variance contribution rate can only reflect the information ratio but the amount carried in the original variables.

To qualitative indexes with vague language, on the basis of fuzzy consistent matrix, fuzzy decision method by rigorous mathematical calculation can get the relative membership degree can be presented [9-11] and entropy can quantify information contained in indexes or variables [12-14]. By summarizing the research results, this paper constructs distribution network planning indexes system, introduces the relative membership degree and entropy utility values to nonlinear principal component analysis, and quickly and accurately evaluate comprehensively distribution network planning.

1.2. The Basic Ideas of Fuzzy Nonlinear Principal Component Analysis

Nonlinear principal component analysis (NLPCA) based on principal component analysis method is a nonlinear analysis method, including four processes: firstly, the initial data matrix from the evaluations of every index to schemes is gotten using fuzzy consistent matrix model, secondly, nonlinear principal components by handling the initial data matrix using Aitchison.J logarithmic transformation to get the covariance matrix and obtain the principal components from the eigenvalues and eigenvectors of the covariance matrix, thirdly, the weights of every principal component using concept of entropy, At last, by a combination of the weight and principal component, decisions can be made.

Nonlinear principal component analysis, taking into account the nonlinear relationship between every index in evaluation system, reducing dimensionality, can settle the problem about comprehensive evaluation results in direct proportion to the correlation degree between the original variables, but to part of the quantitative description indexes in system, the Acheson logarithmic transformation is powerless leading to the initial data matrix hard to be established. The fuzzy consistent matrix based analysis can quantify qualitative description indexes using membership, but it exists defects in fuzzy consistent matrix-based such as difficult to incarnate the difference between two schemes by the elements in the established fuzzy precedence relation matrix, poor similarity degree between obtained fuzzy consistent matrix and fuzzy precedence relation matrix. In view of advantages and disadvantages in the principal component and fuzzy consistent matrix analysis, nonlinear principal component analysis-based total process is completed after all of the indexes is disposed using the fuzzy consistent matrix model to get the new initial data matrix, and the information entropy is introduced in the analysis to get the weight of every principal component.

2. Calculation Process

2.1. Calculation Process Chart

The nonlinear principal component analysis calculation process is shown in Figure 1:

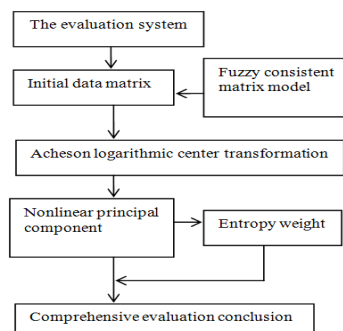


Figure 1. Fuzzy Nonlinear Principal Component Analysis Based Calculation Flow Chart

2.2. Initial Data Matrix

The initial data matrix is constituted by the evaluation of every index to schemes. To complicated evaluation system, the initial data matrix from the relative membership degree matrix by fuzzy consistent matrix model to dispose every evaluation is presented. Assuming that n schemes expressed in $x_i (i = 1, 2, \dots, n)$ and m evaluation indexes expressed in $u_k (k = 1, 2, \dots, m)$ constitute a certain system, the initial data matrix is gotten as follows:

(1) Fuzzy preference relation matrix

The order of a certain index u_k to schemes is expressed with the natural number, the greater numbers the worse schemes, the equal numbers the same importance of schemes. Thus the fuzzy precedence relationship matrix is shown:

$$C = (c_{ij}^k)_{n \times n} \quad (k = 1, 2, \dots, m; i, j = 1, 2, \dots, n)$$

In the formula: c_{ij}^k are the precedence degree coefficients about scheme x_i and scheme x_j to the index u_k . Assuming that the order of scheme x_i to some an index $u_k (k = 1, 2, \dots, m)$ is $a_i (i = 1, 2, \dots, n)$, so m fuzzy preference relation matrixes of m indexes are shown:

$$C^k = (c_{ij}^k)_{n \times n} = 0.5 + \frac{0.5 \times (a_j^k - a_i^k)}{\max_{l=1}^n a_l^k - \min_{l=1}^n a_l^k} \quad (1)$$

(2) Fuzzy consistent matrix

Fuzzy consistent matrix $R_k = (r_{ij}^k)_{n \times n} (k = 1, 2, \dots, m)$ converted from fuzzy preference relation matrix in the condition of $(\max_{p=1}^n r_p^k - \min_{p=1}^n r_p^k) \leq (n - 1)$ is computed by the formula:

$$r_{ij}^k = \frac{r_i^k - r_j^k}{2 \left(\bigvee_{p=1}^n r_p^k - \bigwedge_{p=1}^n r_p^k \right)} + 0.5 \quad (2)$$

In the formula: $r_i^k = \sum_{l=1}^n c_{il}^k (j = 1, 2, \dots, n)$ and r_{ij}^k reflect relative superiority degree about scheme $x_i (i = 1, 2, \dots, n)$ and $x_j (j = 1, 2, \dots, n)$ to the evaluation index u_k .

(3) Optimization value of single index

The optimization value s_i^k about every scheme x_i to index u_k is computed by the formula:

$$s_i^k = \frac{1}{n(n-1)} \left[2 \sum_{j=1}^n r_{ij}^k - 1 \right] \quad (3)$$

(4) Optimization membership matrix of every index

Optimization membership matrix of every index in evaluation system from the optimization value of single index is shown:

$$S = [s_{ij}]_{n \times m} \quad (4)$$

2.3. Nonlinear principal components

(1) Initial data matrix

Optimization membership matrix of every index is the initial data matrix from the formula (4). Thus the nonlinear principal components analysis by Acheson logarithmic center transformation method is presented.

(2) Acheson logarithmic transformation

All of the data in the initial data matrix is converted by Acheson logarithmic transformation shown:

$$y_{ij} = \ln s_{ij} - \frac{1}{p} \sum_{l=1}^m \ln s_{il} \quad (5)$$

Thus the new data matrix is $Y = (y_{ij})_{nm}$.

(3) Covariance matrix

Covariance matrix from the new data matrix by logarithmic center is gotten, expressed in:

$$Z_{ij} = (z_{ij})_{m \times m} \quad (6)$$

In the formula: $z_{ij} = \frac{1}{n-1} \sum_{p=1}^n (y_{pi} - \bar{y}_i)(y_{pj} - \bar{y}_j)$ where $\bar{y}_i = \frac{1}{n} \sum_{p=1}^n y_{pi}$ and $\bar{y}_j = \frac{1}{n} \sum_{p=1}^n y_{pj}$.

(4) Nonlinear principal components

The Nonlinear principal components from the formula $Z_{ij} = (z_{ij})_{m \times m}$ are obtained shown:

$$F = (f_{ij})_{n \times q} = SA^T = s_{nm} a_{mq} \quad (7)$$

In the formula: Matrix $A = (a_{mq})$ is corresponding eigenvectors of every positive eigenvalue of matrix $Z = (z_{ij})_{m \times m}$. Usually, the positive eigenvalues is less than all of the eigenvalues, so the corresponding eigenvectors is less too. That is to say, the corresponding principal components are less than the original variables. The dimensionality reduction is successful.

2.4. Entropy Weight

Usually, the variance contribution rate of principal components is the weight distribution. But the variance contribution rate reflects the relative value of information carried in every original variables but the actual quantity. Information entropy can quantify information, so it is used to determine the weight of t principal components.

(1) Nonnegative processing of principal components

In order to acquire entropy, the possible negative quantity existing in principal component matrix needs nonnegative processing. After nonnegative processing, the matrix is F :

$$F = F + g \quad (8)$$

In the formula: g is the smallest natural number leading all of the principal components to positive.

(2) Entropy

Entropy can display the order degree of information in variables or indexes. The amount of information is expressed in e_t ($t=1, 2, \dots, q$):

$$e_t = -\frac{1}{\ln n} \sum_{i=1}^q h_{it} \ln h_{it} \quad (9)$$

In the formula: $h_{ij} = f'_{ij} / \sum_{j=1}^q f'_{ij}$ where f'_{ij} is the data after nonnegative processing of principal components.

(3) Principal component weight

The disperse degree of the q Principal components is expressed in information avail value d_t ($t=1, 2, \dots, q$), the formula is:

$$d_t = 1 - e_t \quad (10)$$

The data in the q principal component more scattered, the corresponding data in the d_t ($t=1, 2, \dots, q$) lager, indicate that the importance degree of the index is higher, on the contrary, the lower importance degree of the index. Thus, the weight distribution vectors of principal components by normalizing d_t ($t=1, 2, \dots, q$) are acquired:

$$\omega = (\omega_1, \omega_2, \dots, \omega_q)^T, \omega_t = d_t / \sum_{t=1}^q d_t \quad (11)$$

2.5. Fuzzy Nonlinear Principal Component-based Comprehensive Decision Model

Every principal component weighted and synthesized, the comprehensive decision model is shown:

$$O_i = F_{i1}\omega_1 + F_{i2}\omega_2 + \dots + F_{iq}\omega_q \quad (i=1,2, \dots, n) \quad (12)$$

3. Planning Scheme Decisions by Fuzzy Nonlinear Principal Component Analysis to Distribution Network

3.1. Scheme Comprehensive Evaluation Index System of Network Planning

Network planning involves many factors. The scheme comprehensive evaluation index system of network planning is summarized from the [1-4] references, also known as 24 attribute comprehensive evaluation index set. The index system can overall reflect network planning scheme. For some repetitive elements in the evaluation index set, to specific issues, the appropriate comprehensive evaluation indexes are chosen.

3.2. Scheme Decision-making Process

After selecting, evaluation index system of distribution network planning scheme is formed. All of the indexes data in the evaluation system are converted to number between 0 and 1 using the fuzzy consistent matrix decision-making model. Thus the data between 0 and 1 is the membership degree of every index to schemes that establishes initial data matrix of nonlinear principal component analysis and the nonlinear principal components are obtained. After the data in principal component matrix is non-negative processed, entropy utility value of the principal component is gotten, and by the weight distribution after normalization and comprehensive decision can be made.

In this paper, the data in the initial data matrix is weighted nonlinear integration using Acheson logarithmic center transformation and then the principal component gotten, that is to say, the principal component is nonlinearity.

4. Examples

A distribution network planning example from the References [1], verify the correctness of the arithmetic described in this paper. Long-term planning schemes of high-voltage distribution network in the planning area proposed include: scheme 1 (35kV program), scheme 2 (110kV program) and scheme 3 (35kV and 110kV hybrid scheme). The evaluation index system is simplified, selected and shown in Figure 2. According to the planning schemes and comprehensive evaluation indexes system, the evaluation of every index to schemes is shown in Table 1.

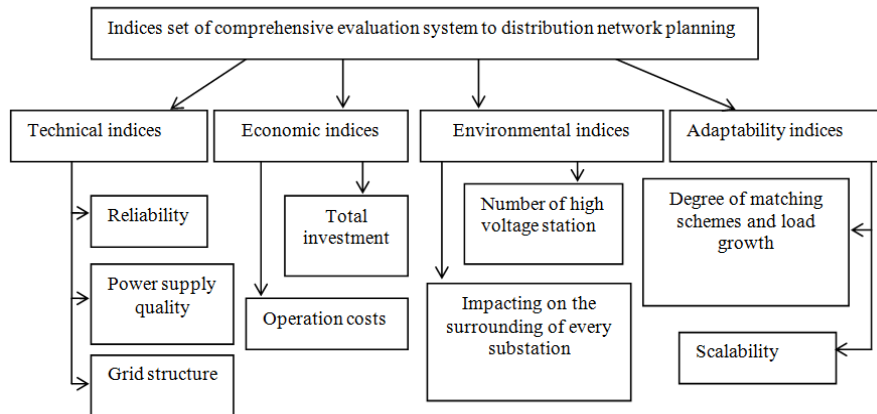


Figure 2. Indices set of Simplified Comprehensive Evaluation System to Distribution Network

Table 1. Evaluation of Every Index to Schemes

Attribute elements	Technical indexes		Economic indexes		Environmental indexes	Adaptability indexes	Scalability indexes
	Reliability	Power supply quality	Total investment (hundred million yuan)	Network loss rate(%)	Number of high voltage station	Degree of matching schemes and load growth	Extensible allowance
Scheme 1	Better	Better	13.02	0.73	More	Middling	The most
Scheme 2	Better	Better	12.73	0.72	Generic	Better	More
Scheme 3	Worse	The best	13.21	0.68	The least	Worse	Middling
Data type	Qualitative	Qualitative	quantitative	quantitative	Qualitative	Qualitative	Qualitative

By the formula (1) to (4), the initial data matrix by the obtained data from the evaluation of every index to schemes using fuzzy consistent matrix model is shown as following:

$$S = \begin{bmatrix} 0.5 & 0.16667 & 0.33333 & 0.08333 & 0.08333 & 0.33333 & 0.58333 \\ 0.5 & 0.16667 & 0.58333 & 0.33333 & 0.33333 & 0.58333 & 0.33333 \\ 0.08333 & 0.66667 & 0.08333 & 0.58333 & 0.58333 & 0.08333 & 0.08333 \end{bmatrix}$$

By the formula (5), the new matrix is:

$$Y = \begin{bmatrix} 0.762714 & -0.33588 & 0.357239 & -1.02909 & -1.02909 & 0.35724 & 0.916859 \\ 0.286675 & -0.81192 & 0.44082 & -0.1188 & -0.1188 & 0.44082 & -0.1188 \\ -0.85305 & 1.22643 & -0.85305 & 1.09289 & 1.09289 & -0.85305 & -0.85305 \end{bmatrix}$$

By the formula (6), the covariance matrix of that new matrix is $Z_{ij} = (z_{ij})_{m \times m}$, shown:

$$Z = \begin{bmatrix} 0.6958 & -0.7676 & 0.5633 & -0.8756 & -0.8756 & 0.5633 & 0.6965 \\ -0.7676 & 1.1381 & -0.7621 & 0.8912 & 0.8912 & -0.7621 & -0.6289 \\ 0.5633 & -0.7621 & 0.5248 & -0.6761 & -0.6761 & 0.5248 & 0.5014 \\ -0.8756 & 0.8912 & -0.6761 & 1.1338 & 1.1338 & -0.6761 & -0.9309 \\ -0.8756 & 0.8912 & -0.6761 & 1.1338 & 1.1338 & -0.6761 & -0.9309 \\ 0.5633 & -0.7621 & 0.5248 & -0.6761 & -0.6761 & 0.5248 & 0.5014 \\ 0.6965 & -0.6289 & 0.5014 & -0.9309 & -0.9309 & 0.5014 & 0.7912 \end{bmatrix}$$

The positive eigenvalues of matrix $Z_{ij} = (z_{ij})_{m \times m}$ are 0.008, 0.5233 and 5.4111, that is to say $q=3$, then by the formula (7), the principal component matrix is:

$$F = \begin{bmatrix} -0.0053 & -0.0387 & -0.0238 \\ 0.1607 & -0.2481 & 0.0927 \\ 0.4406 & 0.2764 & -0.6874 \end{bmatrix}$$

By the formula (8), the natural number $g=1$, after non-negative data in the principal component matrix, and by the formula (9), the entropy e_i of the three principal component is $e_i = (0.999911 \ 0.984817 \ 0.866554)$. By the formula (10), the entropy utility is $d_i = (0.0001 \ 0.0152 \ 0.1334)$, and by the formula (11), the weight vector is $\omega_i = (0.000672 \ 0.102219 \ 0.897108)$.

By the formula (12), the comprehensive evaluation conclusion is:

$$O_i = (0.4119 \ 1.0579 \ 0.9747)$$

The scheme-order is scheme 2, scheme 3, scheme 1, that is to say, the scheme 2 is the optimal scheme. The conclusion consistent with the reference [1], proves the correctness of the method described in this paper.

5. Conclusion

1) According to the research results, index system of comprehensive evaluation to distribution network planning scheme is established. By prioritization of every scheme to indexes, fuzzy preferred relation matrix is established and then the fuzzy consistent matrix. By the relative membership degree matrix of every index to schemes, the initial data matrix in nonlinear principal component analysis is gotten.

2) After Acheson logarithmic center conversion to the initial data matrix, the covariance matrix is obtained. According to the corresponding eigenvectors of positive eigenvalues by the covariance matrix and initial data matrix, principal component matrix is gotten. The principal component is nonlinear combinations of the initial data matrix-vector, it is more reasonable than the commonly linear principal component analysis.

3) The weight distribution vector of principal component using normalized entropy utility not cumulative variance contribution rate, not only unifies the selected standard of principal components number, but also shows the amount of variable information carried in principal component from initial data matrix.

4) In comprehensive evaluation, the indexes having little influence on comprehensive evaluation need be reduced, the strong fuzziness of the comprehensive evaluation system needs to be considered, and the comprehensive evaluation conclusions are made more convincing as much as possible, for the many indexes described only using vague language in distribution network planning evaluation system. The judgment evaluation conclusion using nonlinear principal component analysis brought in fuzzy decision-making and entropy theory is easy to obtain. The calculation process is simple, orderliness, and conclusion is correct.

References

- [1] Nian Liu, Li Ma, Tieming Zhu. Synthetical assessment on distribution network Planning scheme considering Anti-Disaster ability and egional characteristics. *Power System Technology*. 2012; 36(5): 219-225.
- [2] Feng Shi, Peili Liu. Decision making method for urban power distribution grid planning using entropy weight. *Engineering Journal of Wuhan University*. 2012; 45(2): 211-215, 224.
- [3] Gang Wei, Jia Liu, Xin Zhang, et al. *Comprehensive decision-making of AHP/PROMETHEE in distribution network planning*. Proceedings of the CSU-EPASA. 2009; 21(3): 36-40.
- [4] Yu Yi. Comprehensive evaluation and decision-Making for distribution network planning. Dissertation for Master's Degree. 2012.

-
- [5] Lujie Liu, Rong Hu, Yang Fu, et al. Comprehensive evaluation of resource economy based distribution network planning Scheme. *Power System Technology*. 2008; 32(16): 66-70.
- [6] Hongzhan Nie, Song Nie, Yi Qian, et al. Comprehensive decision-making of alternative transmission network planning based on principal component analysis. *Power System Technology*. 2010; 34(6): 134-138.
- [7] Bei Liu. Application of non-Linear principal component analysis method to rock mass quality classification. *Water Resources and Power*. 2011; 29(12): 78-80.
- [8] Ping Li. Classification of expansive soil based on non-linear principal component analysis and cluster analysis. *Yangtze River*. 2012; 43(7): 40-43.
- [9] Zhengyong Zhou, Xiaohui Yuan, Yong Zhou. The prediction research about the strength of cement based on the non-linear principal component neural network. *Mathematics in Practice and Theory*. 2013; 43(3): 83-91.
- [10] Rong Zhan, Deshan Tang, Zizeng Liu. Optimization of rubber dam's styles based on fuzzy consistent matrix. *Water Resources and Power*. 2009; 27(6): 107-109.
- [11] Xiang Qiu. Improved fuzzy consistent matrix and its application in Engineering Evaluation. *Mathematics in Practice and Theory*. 2011; 41(17): 44-47.
- [12] Jikeng Lin, Tongfei Li, Ziming Zhao, et a. Assessment on power system black-start schemes based on entropy-weighted fuzzy comprehensive evaluation model. *Power System Technology*. 2012; 36(2): 115-120.
- [13] Herui Cui, Lihua Liang, Lihong Wang. *Reliability evaluation index of distribution system based on entropy-weight TOPSIS method*. Transactions of the CSAE. 2011; 27(supplement1): 172-175.