IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Research on PTC Cable Materials Based on Principal Component Analysis and Quantitative Correspondence Factor Analysis Method in Big Data Technology

IOP Conf. Series: Materials Science and Engineering **493** (2019) 012084 doi:10.1088/1757-899X/493/1/012084

Research on PTC Cable Materials Based on Principal Component Analysis and Quantitative Correspondence Factor Analysis Method in Big Data Technology

Meimei Wang *

School of Tianjin University of Technology, Tianjin, China

*Corresponding author e-mail: 13752378231@163.com

Abstract. To analyse the huge production data scientifically, the records of enterprises were used. Principal component analysis techniques and quantitative corresponding factor analysis techniques in big data technology were applied. The primary and secondary factors affecting the design performance of PTC cable materials and the influence laws were found. Through analysis, the best process recipe conditions in the existing data were obtained. The results showed that the optimization of the PTC cable material formulation process effectively guided industrial production and met different practical needs. In summary, multi-factor and multi-level visual design and analysis methods, artificial neural network models and big data technology have good qualitative and quantitative analysis functions. A complex process optimization problem with four influencing factors and three indicators is solved.

1. Introduction

PTC material refers to a material with positive temperature coefficient effect. Because of its excellent mechanical properties and convenient molding process, it is widely used in industrial temperature control. China's research on PTC materials started relatively late. Most of them are theoretical studies of PTC materials. The basic formula, volt-ampere characteristics, and molding process have achieved certain research results [1]. In the late 1980s, patent applications were initiated in related fields, but there were no products that could be practically applied to production and large-scale applications. In the early 1990s, some universities and research institutes in China began to pay attention to the research of PTC materials, which set off a PTC material research boom. The research field is extensive, including theoretical research and various aspects of industrial production. More results have been obtained to promote the advancement of domestic PTC material technology. However, the domestic research and development situation still has a long way to go compared with foreign countries. From the domestic research literature on PTC conductive materials, the polyethylene/carbon black system is emphasized. The developed products mainly focus on self-controlled temperature heating cables with a heating temperature lower than 90 °C, while foreign countries have high temperature PTC heating cables with 110, 120-150 °C grades. In addition to polyethylene and polypropylene, the matrix material also includes fluorine-containing and chlorine-containing crystalline resins, polyurethane elastomers, epoxy resins and other amorphous polymers. The resistance of foreign PTC conductive components is stable and the service life can reach 15-20 years. The average lifetime of PTC conductive elements developed in China is only 3-5 years, and the reproducibility of the resistivity-temperature curve of domestic PTC materials is bad, which affects the actual use [2].

Polymer-based PTC composites have been widely used. In particular, the important product of polymer-based PTC composites - polyethylene self-limiting temperature heating cable belt has many excellent performances not available in other traditional heating elements [3], such as preventing overheating, saving electricity, convenience and automation. It is widely used in petroleum, chemical, home appliances, medical equipment and other fields. Research and development of self-limiting temperature heating tape with excellent performance has important practical significance [4]. However, self-limited temperature heating technology is relatively backward. Although the product has been put into production, the stability of the resistivity is bad and the life is short. The mechanical properties are poor, the heating temperature is low and the starting current is large. China's current self-limiting temperature heating cable is mainly a heating cable, and the temperature is generally below 90 °C [5]. Studies have shown that the matrix materials of high temperature heating cables contain fluorine. After being discarded, it will cause serious pollution to the environment and the price of fluorine-containing materials is expensive, which is not conducive to the widespread use of high-temperature heat-carrying electric belts in the civilian field [6]. Therefore, PTC self-limiting temperature heating cable materials with no fluorine, lower production cost and higher heating temperature are imperative.

To this end, for the production data of the new products of the enterprise, the principal component analysis in the big data technology and the innovative quantitative factor analysis method are adopted. The calculation program is developed to extract information that is implicit in the data. Using the visual graphical method, the rules for the development and production of PTC cable materials by enterprises are effectively summarized to guide the research and development and production of enterprises.

2. Related technologies and methods

2.1. Big data technology

There are usually three types of big data, that is, big data, big scale data, massive data. At present, the definition of unity has not yet been formed. The core is the large amount and variety of data. Most of the data comes from the operation of the objective world. In general, valuable information is hidden in large-scale data. The information reflects the law of the things in the data set.

There is currently no uniform definition of big data technology. It should include two major categories of database design and management and data manipulation and analysis. The most important technology here is data mining. Data mining is an interdisciplinary subject, including database, statistics, pattern recognition, machine learning, domain knowledge, etc. [7]. For the production data of a company's research and development of new products for many years, big data technology is used to mine the information hidden in the data, discover the law, optimize the process interval, and in turn guide the enterprise R&D and production. There are many branches for big data analysis processing. The most valuable principal component analysis and innovative quantitative factor analysis methods were studied. The computer program was developed. Data information is analyzed and mined. The company's research and development and production of PTC cable materials are effectively summarized to guide the company's research and development and production.

2.2. Principal component analysis

The four experimental factors of PTC cable material experimental data are representative and systematic. Experimental conditions should be covered throughout the study. Four-dimensional space cannot be directly observed. Therefore, it is mapped to a two-dimensional space. The simple method is to perform a linear transformation, and the new two-dimensional variable is a linear combination of the original four-dimensional variables. Let p random variables of p-dimensional population x be x_1, x_2,..., x_p, and their linear combination constitutes a new comprehensive variable, namely:

$$f_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1p}x_p = a_1 x \tag{1}$$

In order to replace the original variable with as few integrated variables as possible, the more the variance var(f_1) of the integrated variable is, the more information the variable f_1 contains.

$$\text{var}(cf_1) = c^2 \,\text{var}(f_1) \tag{2}$$

That is, a_1 is multiplied by a constant to increase the variance var(f_1) arbitrarily. To make the variance var(f_1) comparable, the coefficients of the linear combination should satisfy the normalization conditions:

$$a_1^T a_1 = a_{11}^2 + a_{12}^2 + \ldots + a_{1p}^2 = 1 \tag{3}$$

In order to find the principal component, the formula of the variance var($a_i^T x$) is analyzed first. The nature of the multivariate random variable can be obtained as var($a_i^T x$):

$$\text{var}(a_i^T x) = a_i^T D(x) a_i = a_i^T \sum a_i \tag{4}$$

When $a^T a=1$, equation (7) is obtained:

$$\text{var}(a^T x) = \sum_{i=1}^{p} \lambda_i (a^T u_i)^2 \le \lambda_i \sum_{i=1}^{p} (a^T u_i u_i^T u_1)^2 = \lambda \tag{5}$$

Therefore, it can be proved that $f_1=u_1^T x$ is the first principal component of the population x. Similarly, it can be proved that $f_2=u_2^T x$ is the second principal component of the population x. Further, $f_i=u_i^T x(i=3,4,\ldots,p)$ can be obtained as the i-th principal component of the population x.

According to the above principle, the data processing program Mat-Des-A0x is written. Big data from experiments and production practices are analyzed and processed.

3. Analysis of results

3.1. Principal component analysis of ptc cable materials

Production data in PTC cable material modification 1458 groups were studied. First, standardization pretreatment was performed on 1485 experimental data. The self-programming program is then used to perform a two-dimensional spatial projection operation on the four variables of the test data. The data is distributed very uniformly in two dimensions, indicating that the experimental design conditions of the enterprise are systematic. Continuity was discovered, indicating that the test was not neglected and the results were systematic. Two dimensions of the 1458 experimental set (resistance-starting current-heating temperature) were projected in two dimensions.

The three indicators can be divided into two broad categories. Material properties can be divided into two broad categories, which require further analysis and analysis using other big data. This type of analysis method combining experimental conditions with experimental results has not been reported in the literature. The results of projecting a seven-dimensional space into a two-dimensional space are performed. The two-factor projection of four factors and three indicators can be divided into five categories. The experimental record manual is reviewed. The first to the 328th trials were Class I, the 329th to 648th trials were Class II, the 649th to 942th trials were Class III, the 943th to 1191th trials were Class IV, and the 1192th to 1458th trials were Class V. The first comprehensive variable has the largest eigenvalue and the largest contribution, which is the U_1 coordinate of the results. The remaining comprehensive variables are listed below.

Table 1. Eigenvalue and variable contribution

Number	Eigenvalues	Proportion of the sum of the eigenvalues	Cumulative contribution
1	0.10096	0.49498	0.49498
2	0.04438	0.21756	0.71254
3	0.02586	0.12680	0.83935
4	0.01732	0.08490	0.92425
5	0.01124	0.05509	0.97934
6	0.00397	0.01947	0.99881
7	0.00024	0.00119	1.00000

The eigenvectors reflect the characteristics of the composite variable. Analysis of the largest contribution of the comprehensive variable (the first comprehensive variable) U1 shows that the amount of coupling agent is the main factor (characteristic vector is 0.79026), followed by the resistance (-0.44299). The reason why the experimental conditions and results are classified into five categories is the difference in the amount of the coupling agent.

The first to 328th trials were Class I, and the 328 test coupling agents were all 27 g. The 329th to 648th tests were class II, and the 320 test coupling agents were all 53g. The 649th to 942th tests were Class III, and the amount of 294 test coupling agents was 78g. The 943th to 1191th trials were class IV, and the 249 test coupling agents were all 104g. The 1192th to 1458th tests were Class V, and the 267 test coupling agents were all 129g. The results indicate that the difference in the amount of coupling agent results in different product properties. The most significant difference between the 1458 experimental conditions was the coupling agent. The adjustment of the dose can control the performance of the product more than other experimental conditions. Analysis of the first comprehensive variable U_1: Among the four influencing factors, the amount of coupling agent has the greatest influence, resulting in a clear classification of products.

Analysis of the second comprehensive variable U_2: the amount of coupling agent (0.57982) still ranked first, and the heating temperature was the second (0.48376). It is indicated that in the direction of U_2 ordinate, the position of the point is determined by the amount of coupling agent and the heating temperature. The general rule is that the larger the amount of coupling agent, the higher the heating temperature. The position of the product point is more inclined to the upper right, which is far from the coordinate origin. Guidance for production: If the production of high temperature PTC cable materials is being developed, consideration should be given to increasing the amount of coupling agent.

3.2. Corresponding factor analysis of PTC cable materials

3.2.1. Corresponding factor analysis. Big data technology is further used to study the relationship between factors and indicators in experimental data. According to the R and Q factor analysis principle and the R-Q type factor joint analysis principle, the dyyzfx2 program is compiled independently.

3.2.2. Comprehensive analysis of experimental conditions and results. The 1458 experimental data was standardized and processed using dyyzfx2.

Table 2. Eigenvalue and variable contribution of corresponding factor analysis

Number	Eigenvalues	Proportion of the sum of the eigenvalues	Cumulative contribution
1	5.09298E-02	5.60945E-01	5.60945E-01
2	1.63452E-02	1.80027E-01	7.40972E-01
3	1.04874E-02	1.15509E-01	8.56481E-01
4	7.74419E-03	8.52952E-02	9.41776E-01
5	4.36637E-04	4.80916E-02	9.89867E-01
6	9.19973E-04	1.01327E-02	1.00000E+00
7	4.65661E-10	5.12883E-09	1.00000E+00

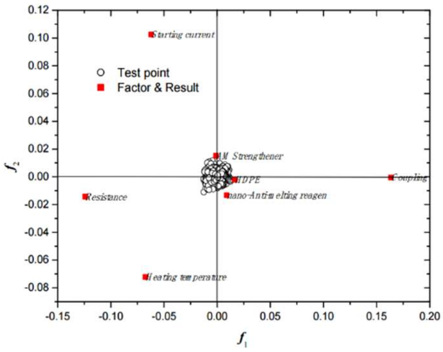

Figure 1. Corresponding factor analysis of experimental conditions and results

3.2.3. Corresponding factor analysis of experimental indicators. By analysing the corresponding factors of the experimental indicators, the group characteristics of the PTC cable materials can be obtained, and the corresponding process recipe can be found for the products producing the specified experimental indexes.

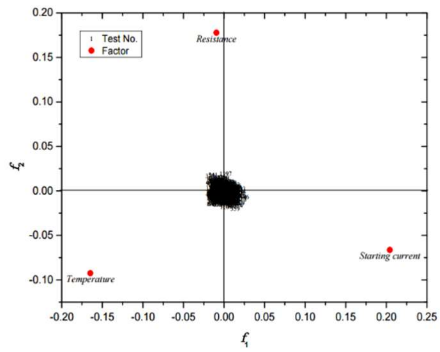

Figure 2. Corresponding factor analysis of experimental results

The abscissa axis f_1 is analyzed from the feature vector as the starting current-temperature axis (the starting current is increasing positively and the heating temperature is negatively increasing). The vertical axis f_2 is the resistance-temperature axis (the forward resistance increases and the negative heating temperature increases). Enterprises and markets require a relatively small starting current. The material has a relatively high heating temperature and the third quadrant meets the material design requirements. The data characteristics of the third quadrant are analyzed.

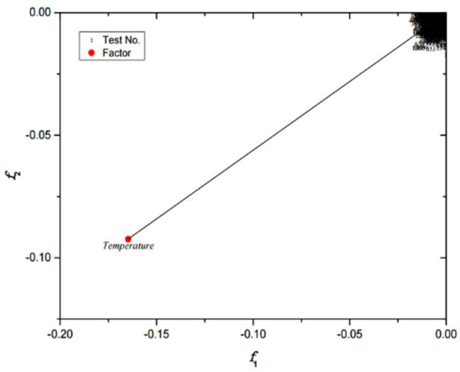

Figure 3. Corresponding factor analysis of experimental results (Triple quadrant)

The heating temperature is made from the center to the vector. Part of the test points are projected on the vector and observed. Test1200: The starting current is 0.1435A, and the heating temperature is 192.8°C. Test649: The starting current is 0.123A and the heating temperature is 182.3 °C. Test650: The starting current is 0.147A and the heating temperature is 186.6 °C. Test1198: The starting current is 0.179A and the heating temperature is 176.9 °C. After tracking the production record, it is convenient to find a material formula with a small starting current and a higher temperature: Test1200: The starting current of HDPE is 0.1435A, and the heating temperature is 192.8°C. Test649: The starting current is 0.123A and the heating temperature is 182.3 °C. Test650: The starting current is 0.147A and the heating temperature is 186.6 °C. Test1198: The starting current is 0.179A and the heating temperature is 176.9 °C.

Table 3. The best process recipe for big data

Test number	Coupling agent dosage (g)	HDPE dosage (g)	Resist agent dosage (g)	Glass fiber consumption (g)
$Test_{1200}$	134	1220	308	129
$Test_{649}$	34	1220	342	78
$Test_{650}$	759	1220	342	78

Detailed research can be done in the vicinity of the above formula data, and it is possible to find a process recipe with better material properties and lower cost.

4. Conclusion

Big data technology can scientifically analyze the huge production data of the enterprises. Primary and secondary factors and rules affecting product design performance are identified. Through analysis, the best process recipe conditions in the existing data were found. Based on this, conditions have been developed to further approximate the optimization results. Traditional multiple regression methods are used to find the law. Here, the amount of data is large and the types of variables are complex. Due to the increasing data, the traditional regression analysis method has a long modeling time and low efficiency. With complex and noisy production data, traditional methods are difficult to solve R&D problems.

References

[1] Guo L, Wang C, Chi C, et al. Exhaled breath volatile biomarker analysis for thyroid cancer. Translational Research, 2015, 166(2), pp. 188-195.

[2] Chen M, Shen M, Li Y, et al. GC-MS-based metabolomic analysis of human papillary thyroid carcinoma tissue. International Journal of Molecular Medicine, 2015, 36(6), pp. 1607-14.

[3] Sumar I, Cook R, Ayers P W, et al. AIM LDM: A Program to Generate and Analyze Electron Localization-Delocalization Matrices (LDMs). Computational & Theoretical Chemistry, 2015, 1070, pp. 55-67.

[4] Kang J W, Singh S P, Nguyen F T, et al. Investigating Effects of Proteasome Inhibitor on Multiple Myeloma Cells Using Confocal Raman Microscopy. Sensors, 2016, 16(12), pp. 11.

[5] Shen C T, Zhang Y, Liu Y M, et al. A distinct serum metabolic signature of distant metastatic papillary thyroid carcinoma. Clinical Endocrinology, 2017, 87(1), pp. 844-852.

[6] Iñigo Landa, Ibrahimpasic T, Boucai L, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. Journal of Clinical Investigation, 2016, 126(3), pp. 1052.

[7] Madhyasthatara M, Askrenmary K, BoordPeter, et al. Dynamic Connectivity at Rest Predicts Attention Task Performance. Brain Connectivity, 2015, 5(1), pp. 45-59.