

## Research on the Evaluation of Dangerous Goods Sources in Air Transportation Based on Principle Component Analysis (Pca) of Matlab

Zhang Yan

Economy and Management Department, Civil Aviation Management Institute of China, Beijing, China

**Keywords:** Principle component analysis, Matlab, Air transportation, Dangerous goods sources, Evaluation and analysis

**Abstract:** With the vigorous development of civil aviation transportation industry, the transportation volume of dangerous goods is increasing. However, the aviation unsafe events caused by dangerous goods are also increasing. In order to find out the important sources of dangerous goods that affect air transportation, seven dangerous goods sources that affect dangerous goods unsafe events are selected basing on the data statistics of dangerous goods unsafe events in 2010-2017. By fully understanding the principal component analysis method, matlab is to be realized the calculation process to carry out systematic research. The first four principal dangerous goods sources are extracted through the principal component analysis method. According to the Principal component analysis of MATLAB, the results show that the main sources of dangerous goods in air transportation are classified into types of dangerous goods, types of airlines, transportation area and transportation mode. In addition, the corresponding preventive measures are given for the four important dangerous goods sources. These preventive measures including enhancing the management of consigning and agency, strengthening the training of relevant personnel and increasing the publicity in all aspects, can provide guidance for the safety management of dangerous goods in air transportation. It is of great value to effectively control the probability of dangerous goods unsafe events and improve the level of safety management.

### 1. Introduction

The vigorous development of civil aviation has promoted the development of economy, and at the same time it has also led to the increase in the volume of dangerous goods transported by air. Dangerous goods is a high-risk and high-profit transportation activity, accounting for a small part of air cargo transportation. However, due to the explosive, burning, toxic, corrosive, radioactive and other properties of dangerous goods, aviation accidents may be caused in air transportation, which obviously endangers human health, safety or causes damage to property. Therefore, under the emphasis of safety as the primary goal of all things in air transportation, reduce the risk of accidents in air transportation of dangerous goods, and reduce the possible consequences of accidents in air transportation. Therefore, it is of great practical significance to find out the dangerous sources of dangerous goods air transportation and avoid them effectively, so as to ensure the efficient and effective completion of the task of dangerous goods air transportation.

Principal component analysis is a kind of comprehensive evaluation method with strong objectivity. It can eliminate the relevant influence among evaluation indexes through quantitative calculation, find out the key factors affecting the system, comprehensively and systematically analyze problems, and thus provide a convenient way for system optimization. There are many quantitative analysis methods for factor analysis, such as fuzzy comprehensive evaluation method, clustering analysis method, chromatography analysis method, neural network analysis method [1-6]. However, many factors need to be considered in the application of these methods, and there are certain correlations among them. There are some repetitions in some information, which not only increases the amount of calculation, but also complicates the problem. Principal component analysis method can avoid these problems and can better meet the requirements of multi-index evaluation. Therefore, it has a good analysis effect in multivariate statistical analysis and project comprehensive effect evaluation. MATLAB is an interactive programming language, which is used to calculate the

calling method of functions of main components of matrix data [7]. The PCA algorithm of the principal component analysis method is combined with MATLAB programming to carry out principal component analysis on the detected high-dimensional data set, so as to achieve the purpose of reducing the data processing amount on the basis of not damaging the data information content [8]. Some scholars have also used this method to do related research. On the basis of principal component analysis, Li Xu and others used MATLAB to implement the calculation process and study the influencing factors of cultivated land area changes in Beijing [9]. Xia Jianguo et al. Carried out an exploratory study on the evaluation of cultivated land quality using principal component analysis [10]. He Weijun and others used PCA to analyze the influencing factors of electricity consumption in China [11]. Zhiyong Chan and others analyzed the key factors of port logistics development based on PCA [12]. In order to accurately assess the safety level of dangerous goods transportation companies to reduce transportation risks, Shen Xiaoyan and others used the principal component analysis BP neural network safety evaluation model [13]. Cui Jianguo et al. Used dynamic PC and improved SVM to efficiently diagnose aero engine faults [14]. However, the analysis of related research on air transport dangerous goods using MATLAB's principal component analysis has not been reported.

In this paper, according to the statistical data of dangerous goods unsafe events from 2010 to 2017, seven indexes are selected. Based on the principle of principal component analysis, MATLAB is used to realize the relevant calculation process, to evaluate the sources of dangerous goods transported by air, and to analyze the main factors affecting the sources of dangerous goods transported by air.

## **2. The Basic Thinking of Principal Component Analysis**

Principal Component Analysis (PCA) is also called principal component analysis or matrix quantitative analysis. It belongs to multivariate statistical analysis technology and uses the idea of dimension reduction to transform multiple complex indexes into a few comprehensive indexes (i.e. principal components). Each principal component largely reflects most of the information of the original variable, and the information contained is not repeated [8]. In this method, the relevant variables are transformed into some uncorrelated comprehensive index variables by the method of variable transformation. When analyzing and evaluating the variable indexes, the relevant interference is avoided and the leading factors are found out. That is to say, the introduction of multi-faceted variables and the reduction of complex factors into several principal components make the problem simple, and the results are more scientific and effective when making accurate evaluation and selection.

MATLAB software carries out numerical analysis with matrix as basic data unit, while principal component analysis method is based on matrix composed of multiple variables and other related data. Therefore, compared with other statistical software such as SPSS and Eviews, matlab can directly or jointly call the contained functions, reduce the research difficulty under the same conditions, simplify the operation, and facilitate the user to improve the model more specifically [15].

This paper uses the statistical data of dangerous goods unsafe events from 2010 to 2017 to analyze the assessment and analysis of dangerous goods sources in air transportation of dangerous goods, and selects 7 parameters for correlation variable analysis. They are respectively dangerous goods category, airline type, person responsible for dangerous goods incident, transportation area, accident type, transportation link and unsafe event type event performance.

## **3. Assessment and Analysis of Sources of Dangerous Goods in Air Transportation**

According to statistics, the total number of dangerous goods unsafe incidents from 2010 to 2017 reached 121. Figure 1 shows that the number of dangerous goods unsafe incidents occurred in 2010 was 4, and 31 unsafe incidents occurred in 2017, with the number increasing year by year. Therefore, it is imperative to study the source analysis and evaluation of dangerous goods in unsafe

incidents.

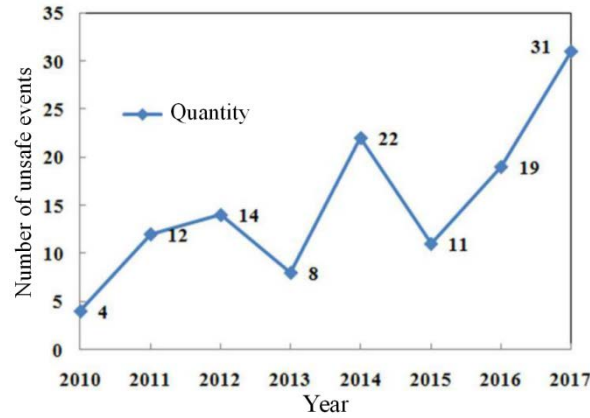


Fig.1 Years Distribution of Dangerous Goods Unsafe Events

In this paper, the statistical data of dangerous goods unsafe events from 2010 to 2017 are used. Due to the complex source factors of dangerous goods unsafe events, the study selects 7 indicators based on SMART principles (simplicity, measurability, availability, reliability and timeliness), which are summarized as categories of dangerous goods, types of airlines, persons responsible for dangerous goods events, transportation areas, types of accidents, transportation links, and performance of unsafe events, as shown in Table 1.

Table 1 Index Factors Of Dangerous Goods Sources

Dangerous goods source	
Type of dangerous goods	X1
Airline type	X2
Person in charge of dangerous goods incident	X3
Transportation area	X4
Accident type	X5
Transportation link	X6
Unsafe event type event performance	X7

Principal component analysis was performed with MATLAB, and the Corrccoef function was called to calculate the vector correlation coefficient matrix. The Pcacov function in MATLAB is called to solve the principal component from the correlation coefficient matrix. The equation's contribution rate of the principal component represents the degree of the principal component's response to sample information changes. The larger the contribution rate of the equation, the more statistical information the principal component reflects.  $\lambda$  is the characteristic value, and the formula of the main component contribution rate is:

$$Z_i = \lambda_i / \sum_{k=1}^m \lambda_k \quad (i=1,2,\dots,m) \quad (1)$$

Cumulative contribution rate formula:

$$\sum_{k=1}^i \lambda_k / \sum_{k=1}^m \lambda_k \quad (i=1,2,\dots,m) \quad (2)$$

When the cumulative contribution rate of principal components reaches 85%, according to the general principle of determining principal components, the first few principal components can be extracted, which can be recognized as representing the information reflected by the original multiple variables. From table 2, the contribution rate of the first principal component is 40.13%, that of the second principal component is 20.68%, that of the third principal component is 19.42%, that of the fourth principal component is 12.80%, and that of the first four principal components is 93.03%, which indicates that the first four principal components can reflect the information of 93.03% of the seven indicators. According to the general principle of determining the principal components, the cumulative contribution rate is  $\geq 85\%$ , the number of principal components determined by comprehensive measurement is 4, and the first four principal components are

extracted.

Table 2 Eigenvalues and Contribution Rate of Principal Components

Component	Characteristic value	Contribution rate (%)	Cumulative contribution rate (%)
1	0.1244	40.13%	40.13%
2	0.0641	20.68%	60.81%
3	0.0602	19.42%	80.24%
4	0.0397	12.80%	93.03%

The correlation coefficient  $R(Z_k, X_i)$  of the  $k$ -th principal component  $Z_k$  and the original variable  $X_i$  is called the load of the  $k$ -th principal component on the  $i$ -th index, which reflects the closeness of the relationship between the original index data and the principal component. This correlation coefficient matrix is called the load matrix (see Table 2),  $\lambda$  is the eigenvalue,  $e_{ij}$  is the eigenvector element corresponding to the eigenvalue, and the principal component load matrix calculation formula is [15]:

$$K_{ij} = \sqrt{\lambda_i} e_{ij} \quad (i, j=1, 2, \dots, n) \quad (3)$$

Table 3 Principal Component Load Matrix

Original variable	Principal component			
	Z1	Z2	Z3	Z4
X1	0.6757	0.4596	0.3343	0.3858
X2	-0.6348	0.0428	0.3528	0.6553
X3	0.0243	0.1879	-0.2371	-0.0587
X4	-0.2495	0.6057	0.4828	-0.4647
X5	0.1658	-0.0446	-0.1041	0.3316
X6	-0.2043	0.6176	-0.6803	0.1795
X7	-0.0916	0.0364	-0.0282	-0.2453

According to Table 2, the absolute values of the first, third and fifth factor loads in the expression of first principal component are relatively large. These three indicators play a major role, and first principal component can be regarded as a comprehensive indicator reflecting the total amount of safety incidents, which is characterized by categories of dangerous goods, persons responsible for dangerous goods incidents and types of accidents. The comprehensive index accounts for 40.13% of all safety accidents. In the second principal component, the first, fourth and sixth indexes have great influence, and the fourth and sixth indexes have relatively great influence. Therefore, the second principal component can be regarded as a comprehensive index reflecting the total amount of safety incidents expressed by transportation areas and transportation links, i.e. a comprehensive index reflecting transportation instructions by the second principal component. The comprehensive index accounts for 20.68% of all safety accidents. In the third principal component, the first, second and fourth indexes have great influence, and the third principal component is regarded as a comprehensive index reflecting the total amount of safety incidents expressed by the categories of dangerous goods, types of airlines and transportation areas. The comprehensive index accounts for 19.42% of all safety accidents. In the fourth principal component, the first, second and fifth indexes are larger than the rest indexes, and the fourth principal component can be regarded as a comprehensive index reflecting the total amount of safety incidents expressed by dangerous goods category, airline type and accident type. The comprehensive index accounts for 12.80% of all safety accidents.

Based on the above indicators, the main factors affecting the sources of dangerous goods transported by air are the categories of dangerous goods, types of airlines, transportation areas and transportation links.

#### 4. Risk Control and Prevention of Hazard Sources in Air Transport of Dangerous Goods

Through the assessment and analysis of dangerous goods air transport hazards, it can be seen that

the categories of dangerous goods, types of airlines, transport areas and transport links are the four key hazards that need attention.

In view of the above hazards, suggestions are as follows: 1. Strengthen the management of shippers and their agents, transfer illegal acts to public security for punishment, cancel their agency qualifications, and establish blacklist management control system for shippers and their agents' statisticians and units who have repeatedly violated regulations. 2. Strengthen the training of dangerous goods knowledge for various personnel such as shippers, agents, and consignors, and establish and perfect the system of consignments and consignments inspection. 3. Increase the intensity of information disclosure: Post the disclosure in the reservoir area, safety supervision department, boarding gate security check link, and official website broadcast the disclosure. Enhance the publicity of dangerous goods to cargo owners, aircrew, passengers and other personnel [16], and remind shippers and agents not to conceal dangerous goods.

## 5. Conclusion

The assessment and control of dangerous goods sources in air transport of dangerous goods plays an important role in the safety of dangerous goods. This paper uses the relevant data of 7 independent variables from 2010 to 2017, and analyzes the influencing factors of dangerous goods sources in air transport based on MATLAB software using principal component analysis. The results show that the source of dangerous goods in civil aviation transportation is mainly affected by four factors. These four aspects are: dangerous goods category, airline type, transportation area and transportation link. According to the analysis of the main influencing factors, the corresponding preventive measures are put forward, which has important practical significance and application value for effectively controlling and improving the safety of air transportation of dangerous goods.

## Acknowledgment

Fund Project: Civil aviation safety capacity building project fund (14000900100018J019)

## References

- [1] Leong C H. A simulation model of an air cargo terminal [D]. Kent Ridge: National University of Singapore, 2004.
- [2] LAI Yong; YANG Fuqiang, "Risk evaluation on self explosion of explosives in sulphide mine based on variable fuzzy sets theo", JOURNAL OF SAFETY SCIENCE AND TECHNOLOGY, 2019, 15(2):94-98
- [3] GAO Yang; ZHENG Jiang ping; CUI Zhen xin, Study on calculating model of probability for civil aviation route bird strike based on fuzzy mathematics [J]. JOURNAL OF SAFETY SCIENCE AND TECHNOLOGY, 2014, 10(1):155-160.
- [4] ZHANG Li, BAI Ping, WANG Zhong-yu, et al. Assessment of accident emergency plan based on analytic hierarchy process and fuzzy comprehensive evaluation[J]. JOURNAL OF SAFETY SCIENCE AND TECHNOLOGY, 2015, 11(9):126-131.
- [5] WANG Haibiao, XU Wenjing..Evaluation on emergency management of earthquake safety community based on AHP entropy matter element[J]. JOURNAL OF SAFETY SCIENCE AND TECHNOLOGY, 2019, 15(6):55-60.
- [6] WANG Xiaoli, WEI Zhibing, PENG Shitao, et al. Comprehensive assessment model of liquid pipeline leakage consequences based on principle component analysis[J]. JOURNAL OF SAFETY SCIENCE AND TECHNOLOGY, 2014, 10(5):85-89.
- [7] Zhang Min, Li Tao-shen, Zhong Shu-ying. "Implementation of principal components analysis method based on Matlab"[J]. Journal of Guangxi University (Nat Sci Ed), 2005, S2: 74-77.

- [8] Xu Cheng. "An Analysis of Factors Affecting the Development of Airport Air Logistics Based on Principal Component Analysis(PCA):A Case Study of Dubai Airport"[J]. TIANJIN SCIENCE & TECHNOLOGY, 2016, 43 (11): 65-70.
- [9] LI Xu, MA Hui-lan. "Study on Influencing Factors for Cultivated Land Loss in Beijing City with PCA MATLAB"[J]. Tianjin Agricultural Science, 2011,17(6):70-74.
- [10] Xia Jian-Guo, Li Yan-Xuan, Deng Liang-Ji. "The application of the principal component analysis method in quality evaluation of cultivated land"[J]. Southwest China Journal of Agricultural Sciences, 2000,13(2):51-55.
- [11] HE Wei-jun, KONG Yang, CUI Yong, et al. "A nalysis of Influencing Factors of Electricity Consum ption in China Based on Principal Component Analysis"[J]. M ATHEMATICS IN PRACTICE AND THEORY, 2018, 48 (1):67-74.
- [12] CHAN Zhi-yong, XU Zhang-yi,"Analysis of Key Factors about Port Logistics Development based on PCA"[J]. LOGISTICS ENGINEERING AND MANAGEMENT, 2012, 1 (34):60-62.
- [13] SHEN Xiao-yan,LIU Hao-xue, XIE Pei. "A Safety Assessment Model for Hazardous Material Transportation Enterprises Based on Principle Component Analysis"[J]. China Safety Science Journal, 2012,22 (1):124-130.
- [14] Cui Jian-guo, Yan Xue, Pu Xue-ping, et al."Aeroengine fault diagnosis based on dynamic PCA and improved SVM"[J]. Journal of Vibration.Measurement& Diagnosis, 2015, 35(1): 94-99.
- [15] Wang Xiao-hong, Yan Yan-li. "Study on the factors influencing the development of energy conservation and environmental protection industry in Gansu Province -- Based on the principal component analysis of MATLAB"[J]. EC0NOMIC RESEARCH GUIDE, 2015, 27 (281):25-27.
- [16] Zhang Yan, Zhao Hong-li. "Safety risk analysis of lithium battery air transport based on Bow-tie model"[J]. Inner Mongolia science and technology magazine, 2019, 5 (38):90-91.