Safety Performance Evaluation Model for Airline Flying Fleets

Yijie Sun[1(✉)], Min Luo[1], Yanqiu Chen[1], and Changhua Sun[2]

[1] China Academy of Civil Aviation Science and Technology, Beijing, China
{sunyj,luomin,chenyq}@mail.castc.org.cn
[2] Civil Aviation Administration of China, Beijing, China
ch_sun@caac.gov.cn

Abstract. The idea of Reason's model is applied to the establishment of civil aviation safety performance index system. The index types and the multiple lines of defense of the unsafe events are placed in one-to-one correspondence to set five dimensions of the index types, including safety result, operation quality, risk management, safety assurance and safety foundation. Taking into account the management complexity and operating characteristics of different flight fleets, the concept of management difficulty coefficient is introduced to improve efficacy coefficient method, and the safety performance evaluation model is established based on the improved efficacy coefficient method. The data of 5 flying fleets in an airline is used as an example of application to verify the feasibility and applicability of the evaluation model. The evaluation results show that the evaluation model can compare fleet safety performance from five dimensions quantitatively, as well as obtain the results of comprehensive evaluation of safety performance for each fleet.

Keywords: Civil aviation · Safety performance · Performance index · Efficacy coefficient method

1 Introduction

Safety performance is one of the important components of airline safety management system, and it is an important index which can reflect the safety management level of airlines. According to Safety Management Manual (ICAO DOC 9859) [1], safety performance is a State's or service provider's safety achievement as defined by its safety performance targets and indexes. Therefore, the evaluation of the safety performance of an airline flying fleet is a proactive and prior safety management approach, and is an important step to test the safety performance of the entire company and the effectiveness of the implementation of the safety management system.

Safety performance evaluation for airline flying fleets has great practical significance. First of all, airlines can understand the current overall safety level and find the existing problems of different flying fleets, and do vertical comparison of different time on entire company's or different flying fleets' safety level. Second, the results of performance evaluation can be used as the basis of incentive system. Based on the safety performance, the airline can encourage employees to become more engaged, and

help to promote performance-oriented safety culture. Thus, it can improve the safety level of the whole company. Third, the results of safety performance evaluation can be used as an important basis for the development and effectiveness test of safety measures.

In recent years, there are many researches on safety performance of civil aviation. Zhang applied fuzzy comprehensive evaluation method to evaluate the safety performance of airlines by using ANP method to determine the index weight [2]. Wang utilized the evidence theory to determine the index weight for the airline safety risk assessment model [3]. Shyur used the data of unsafe events caused by human errors to establish a quantitative model to assess the safety risk [4]. But in these studies, there is no theoretical support for the safety performance index system of airlines, and the index system is not comprehensive, since the safety status are only measured by the high consequence indicators such as accidents and incidents, but not paid enough attention to the low consequence indicators such as management indicators and process indicators. In addition, the existing safety performance evaluation methods do not take into account the complexity of the operation of different flying fleets and the difficulty of management, which lead to the mismatch between evaluation results and subjective perception.

In view of the above problems, based on Reason's model, the index system of airlines safety performance is established. And the concept of management difficulty coefficient is introduced and the efficacy coefficient method is improved to establish the safety performance evaluation model.

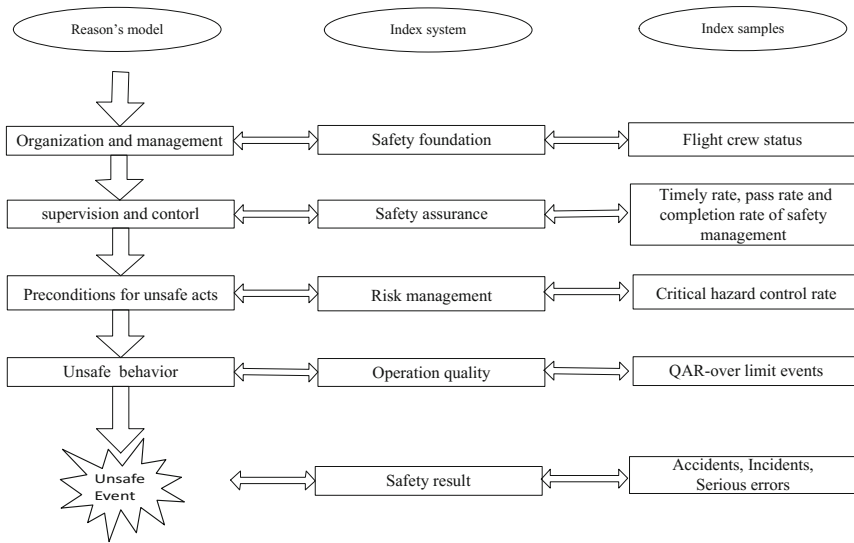## 2 Safety Performance Index System Based on Reason's Model

A systematic and scientific index system is required to assess an organization's safety status comprehensively and accurately, which can reflect the safety results and expose operational and management issues at the same time. Therefore, the index system should not only contain high-consequence indicators which are the safety results indexes, but also contain low-consequence indicators which are process indexes and management indexes.

2.1 Principle of Setting Safety Performance Indexes

Reason's model [5] is proposed by Professor James Reason. It's a classical theoretical model used in aviation accident investigation and analysis. In this paper, the Reason's model is applied to the establishment of the safety performance index system. Multi-level performance indicators are built from the active failure defense and the latent failure defense to fully reflect the safety level of the airline fleets.

According to Reason's model, besides the happened occurrence has a response chain for itself, there is also a set of penetrated failure defenses. The contributed factors of unsafe events and the shortcomings (or safety risks) of the organization on each layer are long-standing, but do not cause significant disasters necessarily. Once multiple

levels of defenses have been broken by some contributed factors sequentially or at the same time, the unsafe event will be occur. Therefore, in addition to the analysis of unsafe acts, it needs to pay more attention to analyze the preconditions of unsafe acts, unsafe supervision and organization factors. By this way, these defects could be recognized comprehensively. In order to set scientific and comprehensive indicators of safety performance, it can follow the principle of Reason's model. If the safety performance indicators are corresponding to the multiple layers of defense which lead to the unsafe events, the index types can be divided into five categories, as shown in Fig. 1.

Reason's model	Index system	Index samples
Organization and management	Safety foundation	Flight crew status
supervision and contorl	Safety assurance	Timely rate, pass rate and completion rate of safety management
Preconditions for unsafe acts	Risk management	Critical hazard control rate
Unsafe behavior	Operation quality	QAR-over limit events
Unsafe Event	Safety result	Accidents, Incidents, Serious errors

Fig. 1. Principle of setting safety performance index

The first one is safety result. These indicators are corresponding to the occurrences themselves, which are high-consequence indicators used to assess the risk of unsafe events such as accidents, incidents, and serious errors in flight fleets.

The second one is operation quality. In accordance with the unsafe acts, these indicators are low consequence process indicators that can evaluate operational technical conditions and the risk of operational bias of the flight fleets.

The third one is risk management. Corresponding to the preconditions of unsafe acts, these indicators can assess the effectiveness of the fleet's control of all types of risks, especially the critical risks.

The fourth one is safety assurance. These indicators are corresponding to the first-line supervision, which can evaluate the development level of safety supervision.

The fifth one is the safety foundation. These indicators are corresponding to the organizational factors, which are used to reflect the appropriate degree of flight crew composition and the state of crew fatigue.

Following the classification method above, there are 430 safety performance indicators have been set up for a comprehensive evaluation of the airline fleet safety level.

## 2.2 Methods for Setting Various Types of Safety Performance Indexes

Setting of Safety Result Indexes. According to China Civil Aviation "civil aircraft incident standard" and "event sample", combined with airline's serious errors, the general errors and other unsafe events criteria, safety result indicators can be established. Usually, the safety results indicators are described as the rate of occurrence such as accident, serious incident, general incident, serious error, general error and other unsafe event. Some indicator samples are shown in Table 1. After recognizing the indicators, it needs to establish risk value calculation model for different types of indicators based on Heinrich's Law to build a safety performance evaluation model.

Table 1. Safety result indexes samples

Occurrence rank	Indicators sample
Accident	Aircraft in operation resulted in 10 or more serious injuries
Serious incident	Landing on wrong runway
General incident	Gravely deviates from the scheduled course
Serious error	Flight turns back or lands at an alternate airport caused by human error
General error	Wrong parking caused by human error
Other unsafe event	Aircraft fly with foreign objects

Setting of Operation Quality Indexes. According to the airline's safety management objectives and historical data on occurrences, 11 critical risks of the airline are reviewed, including loss of control, runway overrun/excursion, tail wiping and so on. Then, the possible direct or indirect causes can be derived from the critical risks by applying the method of qualitative fault tree analysis (FTA). In the cause analysis, SHEL model and Reason's model can be used synthetically to analyze the unsafe behavior or status of the critical risks, as shown in Fig. 2. In order to obtain quantifiable indicators, it's helpful to make the unsafe behavior or status corresponding to QAR monitoring items, and set them as operation quality indicators. Table 2 is an example of partial indicators.

Setting of Other Types of Indexes. Indexes of risk management, safety assurance and safety foundation are used to evaluate the progress and effectiveness of operation management and safety management. They can be collectively referred to as management indicators. These three types of indicators can be set up by using the brainstorming method combined with the safety management system elements method. In other words, it needs to organize aviation management experts to discuss weaknesses in the actual
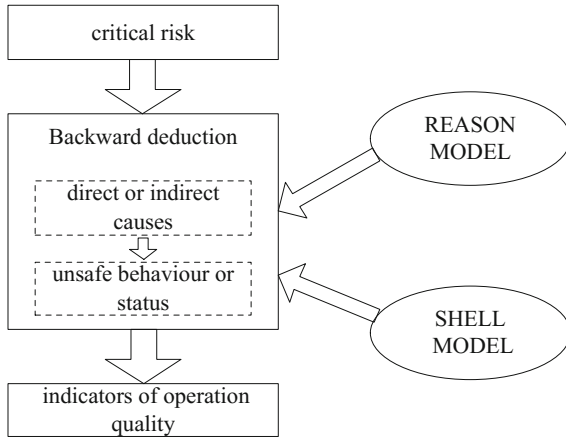
Fig. 2. Approach of operation quality index setting

Table 2. Operation quality indexes samples

Critical risks	Indicators sample
Tail wiping	Speed low at rotation
	Pitch high at liftoff
	Pitch rate high at rotation
	Touch and go
	Pitch high at landing
Hard landing	Drop rate high
	Pitch low at landing
	Touch down with throttle lever

Table 3. Risk management, safety assurance and safety foundation indexes samples

Category of other indicators	Indicators sample
Risk management	Effectiveness of critical risk control
	Timely rate of risk management work
Safety assurance	Execution rate of occurrence investigation
	Proportion of internal audit issues
	Proportion of repeat issues
Safety foundation	Crew fatigue status
	Proportion of safety management staffs

work under 12 elements of safety management system (SMS), and design the management category indicators for assessing the implementation and effectiveness of these SMS' elements. Table 3 shows the examples of these management category indicators.

3 Safety Performance Evaluation Model Based on Improved Efficacy Coefficient Method

According to the safety performance index system established in Sect. 2, we can see that the airline safety performance evaluation system is a multi-index and multi-level system. To evaluate safety performance of the airline flying fleet, it is required that the computation cost of the evaluation model should not be too big, and should not rely too much on expert experience. The evaluation model should be able to carry out multi-dimensional quantitative comparison, and the evaluation results should be consistent with the subjective cognition of the management.

The existing comprehensive evaluation methods for safety performance include analytic hierarchy process (AHP), fuzzy comprehensive evaluation, principal component analysis (PCA), etc. These algorithms have their own application scope and limitations. AHP can't solve the problem of decision making with high quantitative requirements alone, and requires decision-makers to have a deep and comprehensive understanding of the problems faced. The fuzzy comprehensive evaluation method is more complex and has a large amount of calculation. PCA has a high requirement on the quantity and quality of samples, and the meaning of the principal component is fuzzy. In addition, the existing comprehensive evaluation models of safety performance don't take into account the operation complexity and the management difficulty of different flying fleets. For airlines, it often leads to the deviation between the final evaluation results and the management's subjective cognition.

In view of the above problems and the actual needs of airlines, the concept of management difficulty coefficient is introduced in this paper, and the safety performance evaluation model is established based on the improved efficacy coefficient method.

3.1 Safety Performance Evaluation Model Establishment Scheme

Five types of safety performance indexes are established in this paper, the data sources and characteristics of the indexes are different, so the evaluation models for different types of indexes are not the same. The safety performance evaluation model establishment scheme is shown in Fig. 3. For the indexes of safety result and operation quality, risk value model is established based on the principle of risk evaluation and Heinrich's Law. For the indexes of risk management, safety assurance and safety foundation, the index value is calculated according to the airline's internal assessment methods and AHP. In order to make different types of index values obtained by different calculation models comparable, the improved efficiency coefficient method with management difficulty coefficient is introduced to standardize the various indexes. The results of the evaluation model can not only show the different dimensions of the evaluation objects by radar chart, but also evaluate the comprehensive evaluation results of different evaluation objects.
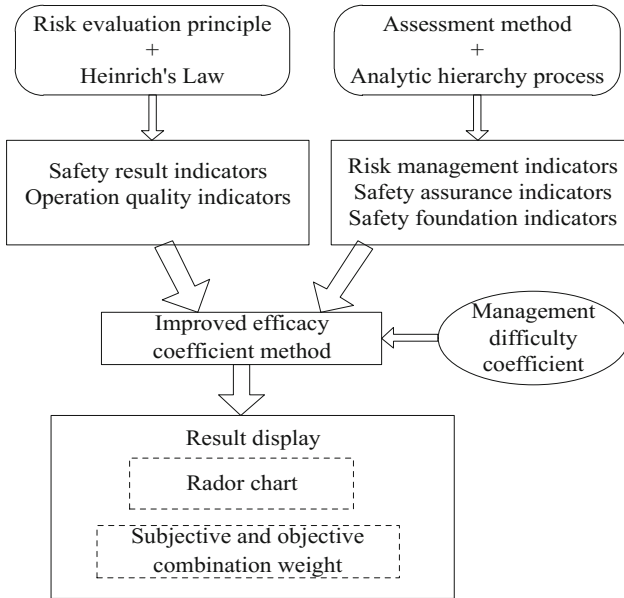
Fig. 3. Safety performance evaluation model establishment scheme

3.2   Procedures and Methods of the Evaluation Model

3.2.1   Risk Value Model for the Indexes of Safety Result and Operation Quality

Risk is the combination of the likelihood and consequences of a particular hazardous situation, which characterizes the probability and severity of a hazardous event. According to the principal of risk assessment, the risk value calculation model needs to be integrated with the probability and severity of risk events. For the indexes of safety result and operation quality in this paper, probability is the frequency of the occurrence of unsafe events, while the severity needs to develop uniform quantitative standards. The severity of different levels of unsafe events is assessed according to Heinrich's Law.

Heinrich's Law is the rules about aviation safety. According to Heinrich's Law, behind every serious accident, there are bound to be 29 minor accidents, 300 accident precursors and 1000 potential hazards. The occurrence of explicit high consequence unsafe events is the result of the accumulation of hidden low consequence events, and the explicit events and the hidden events are regularly proportional. When the severity of different levels of unsafe events is assigned, the severity coefficient can be defined as reciprocal of occurrence frequency according to Heinrich's Law. For different airlines, the proportion of explicit and hidden events is not exactly the same. Based on the historical data related to unsafe events of the airlines, the proportion of different levels of unsafe events in Heinrich's law needs to be adjusted to establish the severity coefficient in line with the actual operation of the airline.

Through collecting and sorting out the historical data of the airline in the past three years, the frequency of occurrence of unsafe events at each level is calculated, and the reciprocal is taken to get the severity of different levels of safety result indexes, as listed in Table 4.

Table 4. Severity of safety result indexes

The level of unsafe events	Severity
Accident/serious incident	3000
Incident	1500
Serious error	300
General error	150
Other unsafe event	20

On the basis of the calculation method of the possibility and severity of each index, the risk value R of safety result indexes of a flying fleet in a given evaluation period is as follows.

$$R = \frac{\sum(\text{severity of each index corrending to unsafe event happened})}{\text{flight movements}} \quad (1)$$

The calculation methods of the severity and risk value for operation quality indexes are similar to safety result indexes.

3.2.2 Calculation for the Indexes of Risk Management, Safety Assurance and Safety Foundation

These three types of indexes are management related process control indexes. The scores of each management related index is defined in airline's internal assessment methods. Take safety assurance indexes as an example, the weight of different indexes can be calculated by AHP, and the index value of safety assurance index can be obtained using the method of weighted arithmetic average.

3.2.3 Data Standardization Based on the Improved Efficacy Coefficient Method

The calculation methods and the units of measurement for different types of indexes are different, and data range is too large for comparison. In order to make the various types of indexes have uniform measurement, and reflect the operation complexity of different flying fleets in the evaluation of safety performance, the management difficulty coefficient is introduced to improve the efficacy coefficient method.

For the flying fleets in airlines, the factors that affect the operation complexity and management difficulty include models, machine age, route structure, professional system and fly missions, etc. The importance of the factors is sorted by the method of expert investigation, and the historical data of the fleet is combined with to determine the management complexity coefficient α of each flying fleet, as listed in Table 5.

Table 5. Management complexity coefficient of each fleet

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
α	1.02	1.02	1.01	1.01	1.00

The management complexity coefficient α is introduced to efficacy coefficient method for data standardization.

$$x'_{ij} = c + \frac{x_{ij} - m'_j}{M'_j - m'_j} \times (100 \times \alpha - c) \tag{2}$$

In Eq. (2), α is the management difficulty coefficient, c is minimum value of desired data range, M'_j is the satisfaction value and m'_j is the not-allowed value.

3.2.4 Safety Performance Comprehensive Evaluation Results

After the data standardization, there are two ways to show the results of comprehensive evaluation.

1. Five dimensions of the flying fleet safety performance evaluation results can be displayed in the form of radar chart.
2. Comprehensive evaluation of each fleet can be calculated by using of subjective and objective combined weights method [6]. Subjective weighs can be calculated by analytic hierarchy process, objective weighs can be calculated by entropy value method, and combined weights can be calculated based on optimality theory. The comprehensive evaluation results of the fleet safety performance can be obtained through the weighted arithmetic mean method.

4 Application Examples

The airline's operational data of 2015 are used to verify the safety performance evaluation model established. And the safety performances of five flying fleets are evaluated.

4.1 Calculation of Safety Result Indexes

According to the risk value model for safety result indexes in Sect. 3.2.1, the number of unsafe events occurred in each fleet is collected, and the severity is calculated based on the event level. Risk value R of safety result indexes of each flying fleet at every month is calculated according to Eq. (1). And part data of the risk value of safety result indexes in 2015 are shown in Table 6.

According to Eq. (2), the data in Table 6 are standardized. Management difficulty coefficient of each fleet is set according to Table 5, c is set to 50, and M'_j and m'_j are the maximum value and minimum value of the risk value of safety result indexes respectively. The data standardization results are listed in Table 7.

Table 6. Risk value of safety result indexes

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
Jan	0.95	170.30	0.00	0.00	0.00
Feb	0.00	0.00	0.00	0.00	338.75
Mar	0.90	1.43	0.00	0.00	0.00
…	…	…	…	…	…
Dec	20.82	324.09	0.00	0.00	0.00
2015	3.29	25.94	13.27	3.83	23.00

Table 7. Data standardization results of safety result indexes

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
Jan	101.85	75.86	101.00	101.00	100.00
Feb	102.00	102.00	101.00	101.00	50.00
Mar	101.86	101.78	101.00	101.00	100.00
…	…	…	…	…	…
Dec	98.80	52.25	101.00	101.00	100.00
2015	101.49	98.02	99.00	100.42	96.61

4.2 Calculation of Operation Quality Indexes

The calculation methods of risk value and data standardization for operation quality indexes are similar to safety result indexes. And the results are listed in Table 8.

Table 8. Data standardization results of operation quality indexes

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
Jan	87.71	92.62	93.00	94.86	100.00
Feb	87.11	92.09	89.48	91.85	99.90
Mar	85.04	92.27	92.46	90.71	98.31
…	…	…	…	…	…
Dec	85.36	96.61	95.85	94.13	96.97
2015	83.60	93.99	93.39	92.37	97.04

## 4.3 Calculation of the Indexes of Risk Management, Safety Assurance and Safety Foundation

According to the calculation method in Sect. 3.2.2, the index value of risk management, safety assurance and safety foundation in 2015 are obtained, as shown in Table 9.

Table 9. Index value of risk management, safety assurance and safety foundation

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
Risk management	90.00	98.68	97.43	96.09	98.68
Safety assurance	97.27	90.36	94.90	96.80	90.00
Safety foundation	100.00	95.71	97.14	100.00	100.00

4.4   Safety Performance Evaluation Results

Radar Chart. Based on the calculation results above, five dimensions of safety result, operation quality, risk management, safety assurance and safety foundation for each flying fleet can be presented in the form of radar chart. Take the data in 2015 as an example to illustrate, as shown in Fig. 4. This way of result presentation can directly see the safety performance of each fleet in different dimensions, and it is convenient to analyze the differences and reasons of the safety performance.
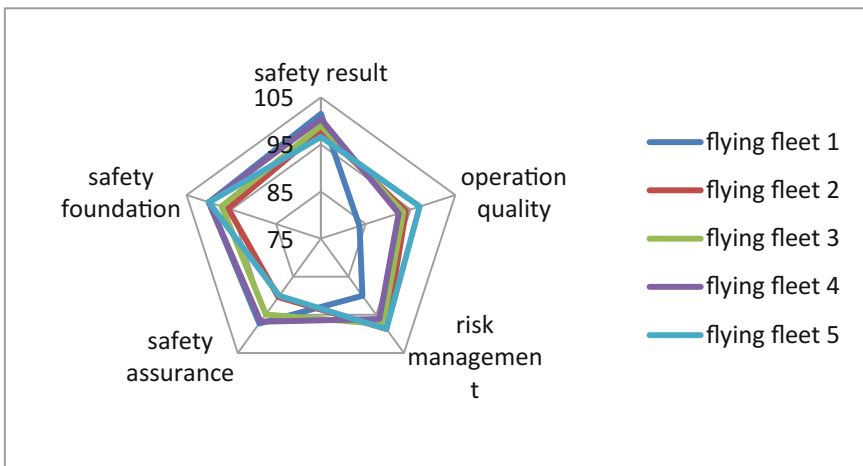
Fig. 4. Safety performance evaluation in 2015

- Flying fleet 1: although the safety result indexes show good results, but because of the risk management indexes are at low level, which reflect the critical risk control effect is not very good, and affecting the operation quality indexes. The follow-up work should strengthen the critical risk control.
- Flying fleet 2: The safety foundation is poor, and safety assurance indexes are not satisfied, which reflect the implementation and effect of the safety management work is not good. The follow-up work should strengthen the implementation and effect of safety management.

- Flying fleet 3 and 4: all types of safety performance indexes are in the middle level, and operation quality can be further enhanced. The follow-up work can enhance the operation technical conditions.
- Flying fleet 5: the performance of safety assurance index is poor, which directly affects the safety result indexes. The follow-up work should continue to improve safety management.

Comprehensive Evaluation Results. In addition to comparing fleet safety performance from five dimensions respectively, airline management level also need to understand the results of comprehensive evaluation of safety performance. Subjective and objective combined weights method mentioned in Sect. 3.2.4 is used to calculate the weights of each type of safety performance indexes, as listed in Table 10. And the comprehensive evaluation of safety performance for each flying fleet is shown in Table 11.

Table 10. Weights of safety performance indexes

	Safety result	Operation quality	Risk management	Safety assurance	Safety foundation
Weights	0.25	0.35	0.15	0.15	0.1

Table 11. Comprehensive evaluation results

	Flying fleet 1	Flying fleet 2	Flying fleet 3	Flying fleet 4	Flying fleet 5
Evaluation results	92.72	95.33	96.00	96.37	96.42

As can be seen from the tables, the company is more concerned about the process indexes and management indexes. Therefore, despite the safety results of flying fleet 1 is the best performance, due to its poor operation quality and unsatisfactory risk management, the overall evaluation of its safety performance ranks last.

5   Conclusion

Based on the idea of Reason's model, safety performance index system which can integrally reflect airline's safety state is established. And based on improved efficacy coefficient method, the airline safety performance evaluation model is presented. Some conclusions are given as follows.

- The safety performance index system based on Reason's model can reflect the airline's safety state from 5 dimensions of safety result, operation quality, risk management, safety assurance and safety foundation. The establishment of the index system has a theoretical basis, and is more comprehensive.
- Compared with the commonly used safety performance comprehensive evaluation methods, the evaluation model established in this paper has smaller computation

cost, doesn't rely much on expert experience, and the evaluation results is consistent with the subjective cognition of the management level.

- Safety performance evaluation model established in this paper could compare fleet safety performance from five dimensions respectively, as well as obtain the results of comprehensive evaluation of safety performance for each fleet. So it helps to further analyze the safety performance level and causes, and make work improvement plan.
- Safety management should pay attention to the process, not just focus on the results. Therefore the flying fleet which has worse process management also has lower evaluation result of safety performance.

References

1. International Civil Aviation Organization: ICAO Doc 9859 Safety Management Manual, 3rd edn. International Civil Aviation Organization, Montreal (2012)
2. Zhang, P.P.: Studies on Safety Performance Evaluation for Airlines. Civil Aviation University of China, Tianjin (2008)
3. Wang, Y.G., Wang, Y.W.: Goal oriented decomposition study of the airport safety and security based on the risk management. J. Saf. Environ. **12**(6) (2012)
4. Shyur, H.J.: A quantitative model for aviation safety risk assessment. Comput. Ind. Eng. **54**, 34–44 (2008)
5. Reason, J.: Human Error. Cambridge University Press, New York (1990)
6. Huang, X.X., Wang, R.H., Zhou, Y.R., et al.: Application of method of combined weights in comprehensive evaluation on inhabitants' health status. J. Math. Med. **19**(4) (2006)