Air Traffic Scenario Evaluation Based on Metric Learning

Dong SUI[1], Qian LI[2], Tingting ZHOU[3], Kechen LIU[4]

[1] dong_sui@126.com, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics

[2] Corresponding author, lee_qian@nuaa.edu.cn, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics

[3] zhoutt@nuaa.edu.cn, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics

[4] liukechen@nuaa.edu.cn, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics

ABSTRACT

Air traffic scenario evaluation can support the optimisation of traffic flow and airspace configuration to improve the safety of air traffic control. Since the air traffic scenario is influenced by the interaction of multiple factors, and real labelled data are lacking, the feature index selection and scenario evaluation are challenging endeavours. In this study, indicators were selected from three dimensions: airspace structure, traffic characteristics and meteorological conditions. The evaluation indicators were quantitatively screened according to information importance and overlap. Utilising the flow control and traffic flow information, the authors defined the free and saturated states of the state interval and developed a metric-based learning method to calibrate the state samples. A multilayer perceptron regression model was employed to establish the mapping relationship between the feature indicators and air traffic scenario. The evaluation accuracy of the sample set from three sectors in Shanghai exceeded 80%, which verified the effectiveness of the scenario evaluation model. This contribution holds practical significance in enhancing the safety of airspace operations.

KEYWORDS

air traffic safety; air traffic scenario; evaluation indicator; metric learning.

1. INTRODUCTION

With the ongoing development of the world's civil aviation transportation industry, the volume of air traffic has continually increased. The sector serves as the fundamental unit in air traffic control. If the air traffic scenario within a sector results in a workload surpassing the tolerable level for air traffic controllers, it can lead to unsafe incidents. For effective risk management in air traffic control, it is crucial to identify traffic scenarios that surpass the controller's workload at the earliest possible stage. Various traffic management methods are employed to adjust the traffic scenario and mitigate the controller's workload. However, the approach of determining the workload level directly from the controller's actual experience has the limitation of being subjective. In contrast, air traffic scenario evaluation has the potential to utilise the airspace structure, traffic characteristics, inclement weather information and other factors to establish a mapping relationship with the workload level. In practical applications, the objective assessment of scenarios and the formulation of precise traffic management measures can be achieved by utilising air traffic and meteorological forecasting information, thereby enhancing air traffic safety. The prevailing approach for situational evaluation involves constructing an evaluation index system through a comprehensive analysis of the factors influencing the air traffic scenario. Linear regression and machine learning methods are employed to establish the mapping relationship between the index and the scenario to evaluate different situational characteristics. Laudeman et al. [1] selected nine traffic characteristic indicators and used a linear weighting method to build a dynamic density model to assess controller workload, while Djokic et al. [2] selected 24

traffic characteristic indicators. Due to the extensive selection of indicators, the author employed principal component analysis to extract eight principal components. This approach was adopted to mitigate the interaction between indicators and develop a multiple regression model, which serves to assess the workload of the controller. Yong et al. [3] employed gray cluster analysis to streamline the indicator system, resulting in the selection of three key indicators from the initial set of eight traffic characteristic indicators. Subsequently, they constructed a multiple linear regression model for assessing the workload of the controller.

Chatterji and Sridhar [4] showed that linear regression models exhibit limited expressive power in representing the mapping relationship between indicators and the scenario. They further argued that nonlinear neural network models are more suitable for this purpose. The authors employed a supervised back propagation (BP) neural network approach to assess the workload of air traffic controllers, considering three levels: high, medium and low. They carefully selected 16 indicators from two dimensions, namely airspace structure and traffic characteristics. Gianazza and Guittet [5] developed a neural network model based on back propagation (BP) to forecast the operational status of merged/armed/split sectors. The model employed a total of 28 indicators derived from two dimensions, namely airspace structure and traffic characteristics. Principal component analysis was conducted to extract the principal components, which were then utilised as inputs for the model. Xiao et al. [6] employed a genetic algorithm to identify seven key metrics from a pool of 28 factors. These metrics were then utilised to construct a neural network model based on back propagation (BP), which aimed to assess the complexity of airspace across three levels: low, medium and high. Andraši et al. [7] selected 20 traffic characteristic indicators and built a multilayer perceptron model to estimate the air traffic complexity. Antulov et al. [8] present the existing issues and new solutions for improved determination of air traffic complexity. The authors propose a novel methodology and plan of future research based on air traffic controller tasks, which is implemented by the machine-learning approaches. Due to the requirement of an extensive amount of labelled samples for training, supervised models face challenges when actual labelled samples are scarce and manual calibration is expensive. To address this, Zhu [9], Cao et al. [10], Zhang [11] and Zhang et al. [12] put forth alternative approaches for assessing airspace complexity, utilising techniques such as semi-supervised learning, transfer learning and unsupervised clustering. Zhu proposed a semi-supervised learning model capable of training with unlabelled samples. In each iteration, the model employs two strategies, controller expert labelling and automatic labelling by the model, to annotate unlabelled samples, constantly updating the training sample set and enhancing the accuracy of model evaluation. Cao trained the sector complexity evaluation model using available calibration samples and employed the transfer learning method to apply the trained model to other sectors lacking calibration samples. Zhang applied principal component analysis to extract the first two principal components from a set of 28 situational evaluation indicators to conduct cluster analysis. They proposed an unsupervised evaluation method for assessing airspace complexity. The resulting clusters categorised the airspace complexity into three levels: low, medium and high, which were then compared with the calibration results provided by control experts for validation. Zhang et al. focus on arrival operations and present twenty-six indicators for describing air traffic complexity and two indicators for arrival operational performance. The authors take the classification method to determine the correlation between complexity and performance. Trajectories of arrival aircraft landing at Guangzhou Baiyun International Airport (ZGGG) are used for case validation.

Researchers have chosen various indicators from the dimensions of airspace structure and traffic characteristics to construct the evaluation index system. However, a universally accepted index system has yet to be established. The present study integrates indicators selected from previous studies to formulate a comprehensive evaluation of the air traffic scenario. Given the numerous indicators, significant variations among different indicator systems and potential mutual influences, a systematic selection of indicators is necessary to enhance the evaluation indicator system. Authors of works from the analysed literature [1, 4, 7, 11] did not optimise the indicator system, whereas others [2, 5, 11] used principal component analysis to extract the principal components. However, this approach posed challenges in analysing the influential

relationship between specific indicators and airspace posture due to the lack of clear interpretability of the principal components. Studies described in the literature [3] used gray cluster analysis to identify key indicators. However, the studies did not adequately account for the information overlap between the identified indicators. The authors [6] used a genetic algorithm to directly search for the set of indicators with the optimal evaluation effect without conducting a prior screening of the evaluation indicator system. However, this approach is limited in ensuring the scientific rationality of the indicator system. Overall, the contribution of the index system optimisation method in the aforementioned research is still insufficient. For a scientific and comprehensive assessment of the air traffic situation, on the basis of the initial indicators of air traffic scenario evaluation, a two-stage evaluation index optimisation method based on the information importance and overlap degree is proposed in this paper. The method is expected to screen out the indices that have a significant impact on the evaluation results and reflect a low degree of information overlap.

Besides linear regression models, researchers have introduced a variety of evaluation methods for scenarios, including supervised, unsupervised, semi-supervised and transfer machine learning models. The accuracy of supervised models for evaluation relies on the precision of sample labelling. However, the authors [4–8] have relied on manual sample labelling by controllers to acquire a sample set. In this approach, however, it is difficult to avoid the subjective influence of manual sample labelling and it is costly, which makes sample labelling a difficult problem in the field of situational evaluation. The evaluation methods of semi-supervised learning and transfer learning, as proposed in the literature [9, 10], rely on a limited number of labelled samples for model training. However, when the sector structures are optimally adjusted, the previously labelled samples become obsolete and require re-labelling. This process incurs a substantial workload and cost. However, an unsupervised clustering method proposed in the literature [11] eliminates the need for sample calibration. Nonetheless, the accuracy of the model evaluation still necessitates manual verification by the controllers. In this paper we quantitatively characterise the air traffic scenario as a percentage of saturation degrees, and present the definition of the free and saturation states of the scenario interval, which is based on flow control and traffic flow information. Additionally, the authors propose a metric learning-based method for sample calibration in air traffic scenario evaluation, which eliminates the need for manual labelling by a controller and ensures an adequate supply of labelled sample data.

The main contributions of this paper are the following. (1) Based on the initial selection of indicators in three dimensions of airspace structure, traffic characteristics and meteorological conditions, the screening method of air traffic scenario evaluation indicators is studied from the perspectives of information importance and information overlap. A set of optimised evaluation indicators was established. (2) By fully applying the operational data, the authors propose using the saturation degree of the sector as the sample label and design a metric learning based method for sample calibration in air traffic scenario evaluation, which solves the problem of the labelled samples being difficult to obtain and enables the continuous evaluation of an airspace scenario from a free state to a saturation state. A multilayer perceptron regression model is developed, trained and tested with sample set data from three sectors in Shanghai to verify the effectiveness of the evaluation method.

The remainder of this paper is organised as follows. Section 2 describes the construction process of the air traffic scenario evaluation indicator system. Section 3 presents the sample label definition method based on metric learning. Section 4 describes the basic structure, parameter settings and experimental results of the situational evaluation model. Section 5 summarises the paper.

2. AIR TRAFFIC SCENARIO EVALUATION INDICATOR SYSTEM

Air traffic dynamics is influenced by the interaction of multiple factors. Authors in the literature [1–7] and [11] mainly selected indicators from two dimensions: airspace structure and traffic characteristics. However, in addition to the inherent structural characteristics of airspace and dynamic traffic characteristics, meteorological conditions as uncontrollable factors affecting air traffic have a key influence on the variation of dynamics. Inclement weather may prevent aircraft from flying as scheduled and the aircraft may be forced to reroute. Consequently, the traffic flow distribution and airspace availability may change. This study aims

to conduct a comprehensive analysis of the influencing factors of air traffic by selecting indicators from three dimensions: airspace structure, traffic characteristics and meteorological conditions. To achieve this, a synthesis of the impact factors of airspace complexity outlined in the literature [13–15] and the evaluation metrics proposed in the classical literature [4, 5, 16] was performed. As a result, a total of 37 evaluation indicators were identified and summarised in *Table 1*.

The indicators were selected from several papers. Consequently, in order to mitigate the mutual influence among these indicators, it was necessary to optimise the evaluation indicator system based on the initially

Table 1 – Indicators of air traffic scenario evaluation

Dimensions	Symbol	Indicators
Airspace structure indicators	X_1	Sector size [13–15]
	X_2	Geometric volume of a sector [13–15]
	X_3	Number of sector sides [13–15]
	X_4	Length of air route
	X_5	Number of intersections of flight paths [13–15]
	X_6	Number of intersecting flight paths [13–15]
	X_7	Proportion of special use airspace [13–15]
	X_8	Number of available altitude layers [13–15]
Traffic characteristics indicators	X_9	Number of aircraft [4, 5]
	X_{10}	Squared number of aircraft [5]
	X_{11}	Number of climbing aircraft [4, 5]
	X_{12}	Number of descending aircraft [4, 5]
	X_{13}	Future incoming flow in horizons of 5 min [5]
	X_{14}	Future incoming flow in horizons of 10 min [5]
	X_{15}	Future incoming flow in horizons of 30 min [5]
	X_{16}	Future incoming flow in horizons of 60 min [5]
	X_{17}	Aircraft density [5, 16]
	X_{18}	Horizontal proximity measure 1 [4, 5]
	X_{19}	Vertical proximity measure 1 [4, 5]
	X_{20}	Vertical proximity measure 2 [4, 5]
	X_{21}	Variance of ground speed of aircraft within a sector [4, 5]
	X_{22}	Ratio of standard deviation of speed to average speed [4, 5]
	X_{23}	Average vertical speed of aircraft within a sector [5]
	X_{24}	Number of potential crossings [5]
	X_{25}	The ratio of flight phase (stable/climbing/descending) [5]
	X_{26}	Variability in aircraft headings [5]
	X_{27}	Variability in aircraft speeds [5]
	X_{28}	Rate of divergences between aircraft pairs [5, 16]
	X_{29}	Rate of convergences between aircraft pairs [5, 16]
	X_{30}	Sensitivity of distance change between diverging aircraft [5, 16]
	X_{31}	Sensitivity of distance change between converging aircraft [5, 16]
	X_{32}	Insensitivity of distance change between diverging aircraft [5, 16]
	X_{33}	Insensitivity of distance change between converging aircraft [5, 16]
	X_{34}	Conflict perception between aircraft pairs maintaining vertical separation [5, 16]
	X_{35}	Conflict perception between aircraft pairs not maintaining vertical separation [5, 16]
Meteorological indicators	X_{36}	Severe weather coverage area [13, 15]
	X_{37}	Severe weather severity [13, 15]

selected indicators. The specific concepts are as follows. Firstly, key indicators are selected based on their information importance using principal component analysis. This allows for the elimination of the least important indicators in the system by considering the magnitude of the load value of the principal component factors. Secondly, indicators are screened based on their information overlap using the Pearson correlation coefficient method. This helps in eliminating indicators with high overlap in the system.

2.1 Screening of indicators based on the importance of information

Principal components obtained using principal compone t analysis methods in the literature [2, 5, 11] are essentially linear combinations of evaluation indicators. These have vague interpretative meanings and cannot be used to analyse the relationship between the influence of specific indicators and evaluation results. This paper employs a screening method for identifying key indicators based on the factor loading values of the initial evaluation indicators on the principal components. The factor loadings indicate the degree of influence that the indicators have on the evaluation results. A stronger correlation between the indicators and the principal components is observed when their absolute values are larger. Additionally, a higher degree of reflected information content corresponds to a more significant influence on the evaluation results [17].

This study focused on Shanghai sectors 11, 16 and 20. The initial evaluation indicators were calculated using Chinese flight tracks and meteorological radar data collected from 4 August to 10 August 2018. A total of 28,617 samples were extracted, with a time slice of 1 minute, and the Z-score method was employed to standardise the data. Principal component analysis was conducted using SPSS software to extract the principal components, and the results are summarised in *Table 2*. Ten principal components were obtained from the initial set of 37 indicators, accounting for a cumulative variance contribution rate of 82.567%.

Table 2 – Contribution rate of principal component

Principal components	Eigenvalue	Variance contribution rate (%)	Cumulative contribution rate (%)
1	9.062	24.492	24.492
2	7.160	19.351	43.843
3	3.810	10.298	54.142
4	2.207	5.966	60.107
5	1.958	5.292	65.399
6	1.773	4.792	70.191
7	1.383	3.738	73.929
8	1.184	3.201	77.130
9	1.012	2.734	79.864
10	1.000	2.704	82.567

The factor loading values of the 37 indicators on the ten principal components mentioned above can be found in Appendix A. To ensure that the selected evaluation indicators have a significant impact on the evaluation results, the indicators with lower absolute values of factor loading on the principal components were removed. The screening criteria for the absolute values of factor loadings in some studies are between 0.4 and 0.9. The authors chose to retain the indicators with absolute values of factor loadings on principal components higher than 0.6. For the seventh and ninth principal components, the authors observed that the variance contribution of each indicator to these components is not greater than 0.6. Therefore, the authors retained the indicator with the largest absolute value of factor loadings in both components to ensure that all the screened indicators have significant effects on air traffic dynamics. The results of indicator screening are shown in Appendix A. In the screening process of information importance, 11 indicators, including X_{11}, X_{12}, X_{18}, X_{24}, X_{26}, X_{28}, X_{30}, X_{32}, X_{33}, X_{34} and X_{35}, were omitted, and 26 indicators were retained.

2.2   Indicator screening based on information overlap

To eliminate redundant information among the remaining 26 indicators, a second round of screening was performed based on the degree of information overlap. The Pearson correlation coefficient method was employed to identify and remove overlapping indicators. A higher correlation coefficient between two indicators signifies a stronger correlation and greater information overlap. To enhance the rationality of the screening process, the authors systematically reduced the level of information overlap between evaluation indicators by calculating their coefficient of variation. In cases where two indicators have a higher correlation coefficient, the indicator with a smaller coefficient of variation was eliminated. The coefficient of variation measures the degree of dispersion in the values of an indicator, with a higher value indicating greater importance [18].

The Pearson correlation coefficients and coefficients of variation between the indicators were calculated, and the results are presented in Appendix B and C. The critical value of the Pearson correlation coefficient was set as 0.9, that is, when the correlation coefficient between two indicators was greater than 0.9, the correlation between indicators was considered high, and the indicator with lower coefficient of variation was omitted to reduce the information overlap between indicators. In this round of the screening process, a total of 13 indicators, including X_1, X_2, X_3, X_4, X_6, X_8, X_9, X_{13}, X_{14}, X_{16}, X_{17}, X_{29} and X_{37}, were omitted, and the final air traffic scenario evaluation indicator system containing 13 evaluation indicators was determined, as summarised in *Table 3*.

Table 3 – Air traffic scenario evaluation indicator system

Dimensions	Symbol	Indicators
Airspace structure indicators	X_5	Number of intersections of flight paths
	X_7	Proportion of special use airspace
Traffic characteristic indicators	X_{10}	Squared number of aircraft
	X_{15}	Future incoming flow in horizons of 30 min
	X_{19}	Vertical proximity measure 1
	X_{20}	Vertical proximity measure 2
	X_{21}	Variance of ground speed of aircraft within a sector
	X_{22}	Ratio of standard deviation of speed to average speed
	X_{23}	Average vertical speed of aircraft within a sector
	X_{25}	The ratio of flight phase (stable/climbing/descending)
	X_{27}	Variability in aircraft speeds
	X_{31}	Sensitivity of distance change between converging aircraft
Meteorological indicators	X_{36}	Severe weather coverage area

3.   METRIC-LEARNING-BASED APPROACH TO SAMPLE LABEL DETERMINATION

3.1   Sample label definition

Air traffic situational change is a dynamic and continuous process. Currently, control experts manually calibrate samples to classify their states into high, medium and low complexity levels. However, this manual calibration method is costly, time-consuming and cannot continuously calibrate sample states. To overcome the challenge of obtaining sample labels, the authors propose using the saturation degree of the sector as the sample label. Furthermore, the authors provide objective definitions for three airspace operation states: the free state, saturation state and intermediate state. These definitions are based on sector traffic statistics and flow control information and are as follows.

1) *Free state.* The sector operation state is defined as the free state based on hourly flight statistics, indicating significantly low traffic volume. In this state, the low volume of flights results in low operating pressure, and the sector saturation value is set to 0. Samples corresponding to this state are labelled as "0". *Figure 1* presents the flight flow statistics of sectors 11, 16 and 20 in the Shanghai control area on 6 August 2018. The flight flow during the 00:00 ~ 06:00 period is the lowest and accounts for less than 5% of the total daily flight flow. Hence, the sectors are classified as being in the free state during this time, and the corresponding samples are labelled as "0".
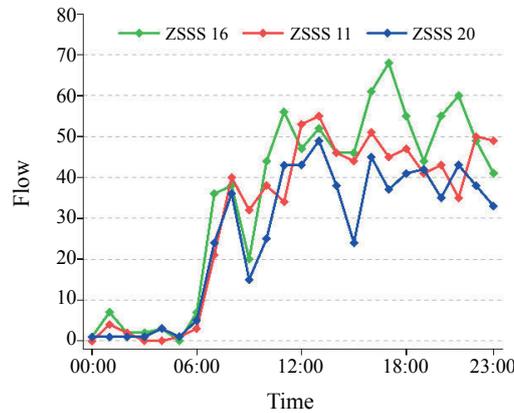

Figure 1 – 24-h flight flow in three sectors of Shanghai
(ZSSS11, ZSSS16 and ZSSS20 represent Shanghai sectors 11, 16 and 20, respectively)

2) *Saturated state.* When there is excessive flow or inclement weather that puts excessive pressure on controllers, air traffic control issues flow control instructions to restrict the number of flights entering the sector for a specific period, ensuring operational safety. Consequently, the sector is classified as being in the saturation state during the flow control period, with a saturation degree of 100%. Samples corresponding to this state are labelled as "1".

3) *Intermediate state.* The states other than the free and saturated states are classified as intermediate states, with their saturation degree values ranging between 0 and 1. While the saturation degree of free state and saturated state samples can be determined using flight flow and flow control information, determining the saturation degree of intermediate state samples lacks objective data. To address this, we propose a label determination method based on a distance metric to assess the similarity between samples by calculating the distance between them. In classification or clustering algorithms, the traditional Euclidean distance is frequently employed to measure the similarity between samples. However, the Euclidean distance treats differences between different attributes of the sample data equally, disregarding their importance in various data structures or distributions. Therefore, it is crucial to identify a suitable metric that can effectively capture the distance or similarity between samples. To address this, in this paper, the authors propose a metric learning approach that aims to learn a metric reflecting the distribution of the dynamic data and providing a more accurate representation of the distance or similarity between samples [19].

3.2   Sample label determination

A metric serves as a function that defines the distance between elements in a given set. The metric learning method can autonomously learn a task-specific distance metric function based on different tasks. Since the Mahalanobis distance has a learnable metric matrix, M, the metric learning model is usually built on the basis of the Mahalanobis distance. For given samples, x_i and x_j, the expression of the Mahalanobis distance between them is shown in *Equation 1*, and the learned inter-sample Mahalanobis distance is the learned metric matrix, M.

$$D_M(x_i, x_j) = \sqrt{(x_i - x_j)^T M (x_i - x_j)} \tag{1}$$

Weinberger and Saul [20] proposed the large margin nearest neighbours (LMNN) metric learning method. Its learning strategy is to minimise the sum of the distances of samples of the same category in the

nearest-neighbour objective, while the samples of different categories are separated by the large margin; i.e. it can make the distance between samples of the same category closer and the distance between samples of different categories farther, as shown in *Figure 2*.
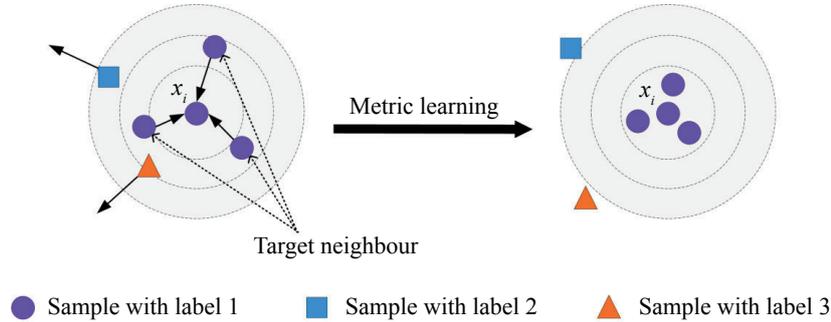

Figure 2 – Schematic view of metric learning strategy

LMNN learns a new metric for the original data using a supervised learning approach based on the labelled training set. In this paper, the authors utilised the calibrated samples from the free state and saturated state as the training set to learn a Mahalanobis distance metric. This metric is designed to effectively capture the distribution of state data, leveraging an existing metric learning model proposed in previous studies [20]. The model is as follows.

$$\min \sum_{ij} \eta_{ij} (x_i - x_j)^T M (x_i - x_j) + c \sum_{ij} \eta_{ij} (1 - y_{ik}) \xi_{ijk} \tag{2}$$
s.t.

$$(x_i - x_k)^T M (x_i - x_k) - (x_i - x_j)^T M (x_i - x_j) \geq 1 - \xi_{ijk} \tag{3}$$

$$\xi_{ijk} \geq 0 \tag{4}$$

$$M \geq 0 \tag{5}$$

Let $\{(x_i, y_i)\}_{i=1}^n$ be the model training set, where $x_i \in R^d$ denotes the sample data with feature dimension d, and y_i denotes the sample labels. Moreover, $\eta_{ij} \in \{0,1\}$ indicates whether samples x_i and x_j are nearest-neighbour samples, $\eta_{ij}=0$ indicates a non-nearest-neighbour relationship, and $\eta_{ij}=1$ indicates a nearest-neighbour relationship. $y_{ij} \in \{0,1\}$ signifies whether labels y_i and y_j are consistent, $y_{ij}=0$ indicates that x_i and x_j are different types of samples, and $y_{ij}=1$ indicates that x_i and x_j are the same type of sample. Furthermore, ξ_{ijk} is the slack variable and $M \in R^{d \times d}$ is the metric matrix to be learned.

The LMNN model is solved using the DR Toolbox implemented in MATLAB to obtain the metric matrix M. This metric matrix allows us to compute the new Mahalanobis distance metric function $D_M(x_i, x_j)$, which is used to measure the distance or similarity between different posture samples. The metric matrix of different sectors should be trained with the sample data of each sector, and the metric matrix of each sector is different. Taking 11 Shanghai sectors as an example, metric matrix M is obtained as shown in Appendix D.

Based on the distance metric, the labels of the intermediate state samples are determined by calculating the similarity degree between the intermediate state samples and the saturated state samples. First, based on metric expression $D_M(x_i, x_j)$, the distances from the intermediate state samples x_i to the cluster centre points C_0 and C_1, denoted as d_0 and d_1, respectively, are calculated as shown in *Equations 6 and 7*, which represent the distances between the intermediate state samples and the free and saturated state samples.

$$d_0 = \sqrt{(x_i - C_0)^T M (x_i - C_0)} \tag{6}$$

$$d_1 = \sqrt{(x_i - C_1)^T M (x_i - C_1)} \tag{7}$$

The distance ratio $dis_{ratio}=d_1/(d_0+d_1)$ is defined according to d_0 and d_1 to indicate the distance of the intermediate state sample from C_1. The larger the d_1 and dis_{ratio} are, the greater the distance of the sample from C_1 is. The lower the corresponding degree of similarity with C_1 is, the lower the degree of saturation is.

Conversely, the closer the intermediate state sample is to C_1, the smaller the dis_{ratio} is and the more similar the sample is to C_1, the higher the degree of saturation is. The calculation formula is shown in *Equation 8*.

$$similarity = 1 - \frac{d_1}{d_0 + d_1} \tag{8}$$

Based on the above label definition method, the labels of the intermediate state samples in sectors 11, 16 and 20 of Shanghai are determined to obtain the complete set of labelled samples. The label values in the final sample set range from 0 to 1, representing the continuous saturation degree of the sample airspace. To compare the measurement effect of the Mahalanobis distance and traditional Euclidean distance with Shanghai sector 11 as an example, the distance distribution between samples is calculated based on the Mahalanobis distance and Euclidean distance, respectively. Taking the cluster centroids of the free state samples and the saturated samples as C_0 and C_1, the distances of the free state samples from C_0 and the saturated samples from C_1, respectively, are calculated and denoted as D_{S0} and D_{S1}, indicating the distribution of distances between similar samples, as shown in *Figure 3*. Next, the distances of the free state samples from C_1 and the saturated state samples from C_0 denoted as D_{d0} and D_{d1}, respectively, are calculated to represent the distance distributions of different categories of samples, as shown in *Figure 4*.

a) D_{S0} distribution range

b) D_{S1} distribution range
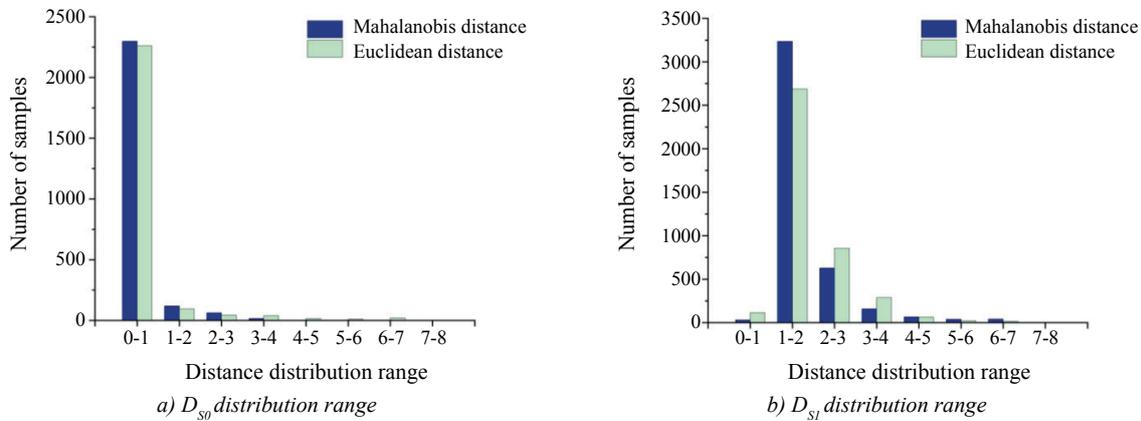
Figure 3 – Distance distribution of samples in the same categories

a) D_{d0} distribution range

b) D_{d1} distribution range
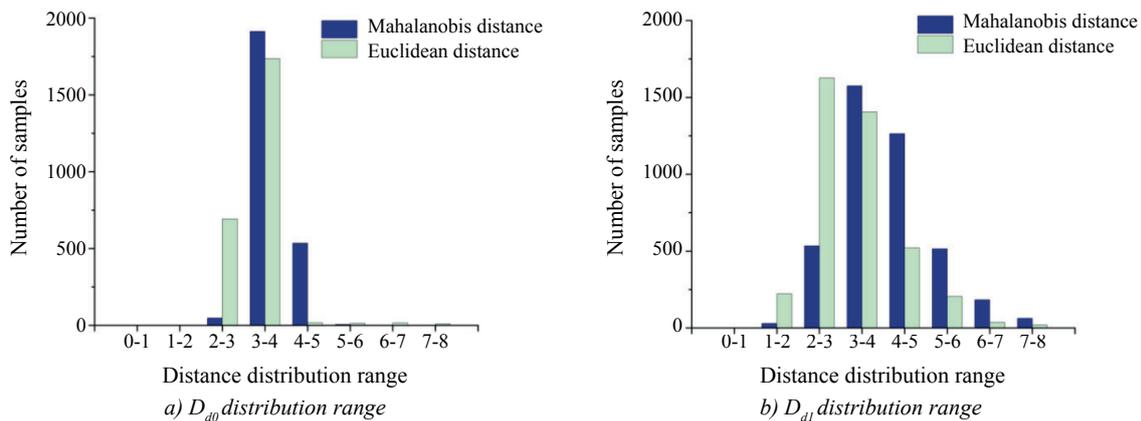
Figure 4 – Distance distribution of samples in the different categories

In *Figure 3 and 4*, the horizontal axis represents the distance distribution range between samples, while the vertical axis indicates the number of samples within each range. By analysing the distance distribution of samples belonging to the same category in *Figure 3*, it can be observed that the main distance distribution ranges for D_{S0} are [0, 3] and [0, 4] under the Mahalanobis and Euclidean distance, respectively. While the main distance distribution ranges for D_{S1} are [1, 3] and [1, 4] under these two distances metrics, respectively.

By analysing the distance distribution of the samples belonging to different categories in *Figure 4*, it is observed that the main distance distribution ranges for D_{d0} are [3, 5] and [2, 4] under the Mahalanobis and Euclid-

ean distance metric, respectively. While the main distance distribution ranges for D_{d1} are [2, 7] and [1, 6] under these two distances, respectively. Therefore, compared with the Euclidean distance, the Mahalanobis distance metric has a smaller distribution of distances between samples of the same category and a larger distribution of distances between samples of different categories. The results obtained using the Mahalanobis distance metric are consistent with the learning strategy of metric learning.

4. MLP-BASED AIR TRAFFIC SCENARIO EVALUATION MODEL

The air traffic scenario evaluation model aims to help air traffic management departments acquire the real-time airspace operation status. In addition, based on aviation flight track prediction and meteorological forecast data, it can evaluate the trend of future scenario changes, achieve monitoring and early warning of the airspace operation status, avoid the over-saturated operation status and ensure airspace operation safety.

4.1 Multilayer perceptron regression model

Supervised multilayer perceptron models are widely employed for addressing classification and regression problems in the field of machine learning. Considering that the indicator data in this study are not complicated, the neural network model can better analyse the nonlinear mapping relationship between indicators and scenario. With the challenge of sample calibration resolved and an abundance of labelled sample data available for model training, a regression model based on the multilayer perceptron is constructed to evaluate the air traffic scenario. The model network structure is shown in *Figure 5* and consists of three parts: an input layer, a hidden layer and an output layer, which are fully connected.
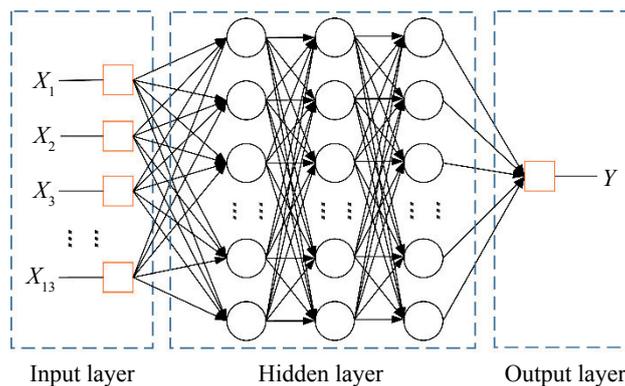
Figure 5 – Network structure of MLP model

1) *Input layer.* The layer consists of 13 nodes, representing the 13 screened air traffic scenario evaluation indicators used as inputs for the model.
2) *Output layer.* The layer consists of one node and its output represents the air traffic scenario, specifically the saturation level value of the airspace obtained from the model.
3) *Hidden layer.* The hidden layer follows the input layer. The number of layers and the number of nodes are important parameters. The number of layers is generally set to one to three layers. The number of nodes can be selected according to the specific application. A conservative approach suggests setting the number of nodes in the hidden layer between the number of output neurons and the number of input neurons [21], specifically ranging from 1 to 13 in this study. In addition, the ReLU activation function is employed due to its ability to mitigate the issue of the "vanishing gradient" problem and its widespread adoption compared to the sigmoid and tanh functions.

Root mean square error (RMSE), mean absolute error (MAE) and accuracy are used as model evaluation metrics, as shown in *Equations 9–11*, where y_t denotes the true label value, y_t represents the predicted value and n is the number of predicted samples.

$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (y_t - \hat{y}_t)^2} \qquad (9)$$

$$MAE = 1/n \sum_{i=1}^{n} \left| y_t - \hat{y}_t \right| \tag{10}$$

$$Accuracy = 1 - \frac{\| y_t - \hat{y}_t \|_F}{\| y \|_F} \tag{11}$$

Parametric experiments were conducted to determine the optimal number of hidden layers and nodes in the model. The experiments aimed to evaluate the model's performance by varying the numbers of hidden layers and nodes. The results are shown in *Figure 6*.
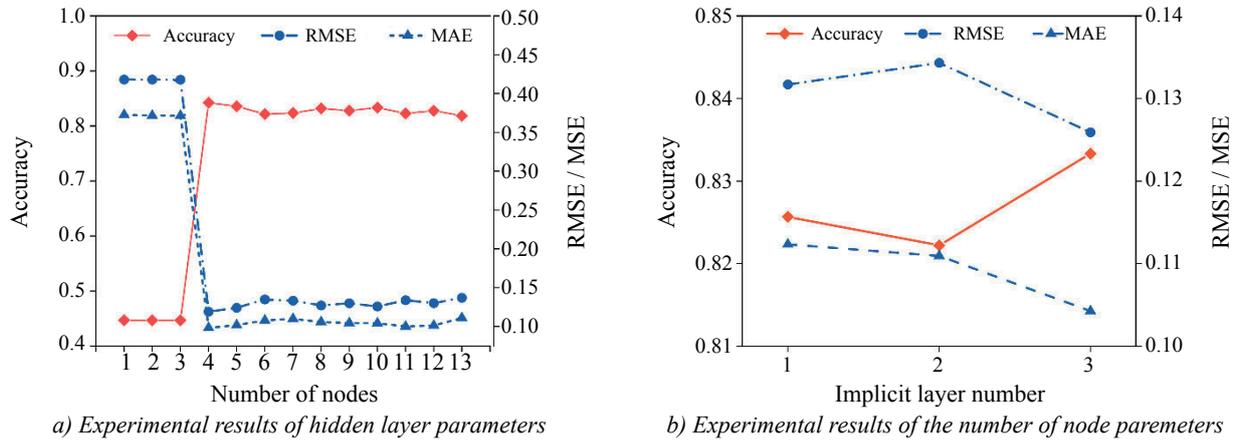

a) Experimental results of hidden layer parameters   *b) Experimental results of the number of node paremeters*

Figure 6 – Hidden layer parameters and implicit layer number parameter experimental results

According to *Figure 6* , the model achieves the highest prediction accuracy with three hidden layers. The prediction results of the model exhibit stability when the number of nodes in the hidden layers is set to ten. Consequently, the model is configured with three hidden layers and ten nodes.

4.2 Experimental results

The experimental data contain a total of 28,617 samples from the sectors 11, 16 and 20 in Shanghai, as summarised in *Table 4*. The training set and test set of each sector contain free state, saturated state and intermediate state samples, of which 80% of the data are used as the training set and 20% of the data are used as the test set. The data in the training and test sets are not duplicated.

Table 4 – Amount of experimental data in each sector in different scenarios

Shanghai sector	Number of free state samples	Number of saturation state samples	Number of intermediate state samples	Total
11	2502	4203	2940	9645
16	2070	3623	3612	9305
20	2148	4555	2964	9667

To compare the effects of different metrics on the model evaluation results, this section includes a comparative experiment that utilises the traditional Euclidean distance to calculate the labels for intermediate state samples. The models are trained and tested on datasets from the sectors 11, 16 and 20 in Shanghai, respectively, and the model evaluation results are obtained as shown in *Table 5*.

Table 5 – Model evaluation results for each sector based on different metric distances

Shanghai sector	Mahalanobis distance			Euclidean distance		
	RMSE	MAE	Accuracy	RMSE	MAE	Accuracy
11	0.085	0.114	85.31%	0.168	0.137	77.14%
16	0.112	0.147	80.41%	0.192	0.136	73.82%
20	0.118	0.086	84.94%	0.182	0.145	76.12%
Average	0.105	0.116	83.55%	0.181	0.139	75.69%

From the experimental results, it is evident that the mean value of model evaluation accuracy reaches 83.55% when calculating the labels for intermediate state samples using the Mahalanobis distance. In contrast, when utilising the Euclidean distance to calculate the labels for intermediate state samples, the mean accuracy of the model evaluation is 75.69%, showing a 7.86% decrease compared to the Mahalanobis distance metric.

Through the metric learning approach, the model successfully learns suitable distance metric expressions to determine the labels of intermediate state samples, resulting in a model evaluation accuracy of 80% and above. This represents a substantial improvement when compared to the conventional distance metric approach for model evaluation. The metric learning based method for sample calibration which achieves the continuous evaluation of an airspace scenario from a free state to a saturation state, and solves the problem of the labelled samples being difficult to obtain, is proposed in this paper. This method can achieve higher accuracy of air traffic situation assessment and provide effective technical support for ATC operations and traffic management.

5. CONCLUSION

In this study, an air traffic scenario evaluation model was designed. Scenario evaluation indicators from three dimensions: airspace structure, traffic characteristics and meteorological conditions were used. The evaluation indicators were quantitatively screened using sample data, and a set of optimised evaluation indicators was established. By combining flow control and traffic flow information, a metric-learning-based sample label determination method was designed to solve the problem of the labelled samples being difficult to obtain. The experimental results from three sectors in Shanghai show that the metric-learned Mahalanobis distance can better measure the similarity between samples, and the model evaluation accuracy is improved by 7.86% compared with the traditional Euclidean distance metric. In addition, unlike the current discrete evaluation method, which divides the evaluation results into high, medium and low levels, the sample calibration method proposed in this paper enables the continuous evaluation of an airspace scenario from a free state to a saturation state. Moreover, the evaluation results are highly accurate and make it more convenient for the air traffic control department to monitor and manage the airspace scenario in real time. This contribution holds practical significance in ensuring the safety of airspace operations. In this paper, the experiments employed a linear metric learning method, which limited the exploitation of the nonlinear relationships among the samples. For future research, the application of the nonlinear metric learning method GB-LMNN can be explored, as it offers advantages in terms of robustness and speed.

REFERENCES
[1]   Laudeman LV, et al. *Dynamic density: An air traffic management metric*. Washington, D. C.: NASA; 1998.

[2]   Djokic J, Lorenz B, Fricke H. Air traffic control complexity as workload driver. *Transportation Research Part C.* 2010;18(6):930-936. DOI: 10.1016/j.trc.2010.03.005.

[3]   Yong H, Chen Z, Hui L. Evaluation method of ATC controllers' workload based on the metric of traffic complexity. *Information Technology Journal.* 2013;12(23):7215-7221. DOI: 10.3923/itj.2013.7215.7221.

[4]   Chatterji GB, Sridhar B. Measures for air traffic controller workload prediction. *Proceedings of the First AIAA Aircraft Technology, Integration, and Operations Forum,16-18 Oct. 2001, Los Angeles, CA.* DOI: 10.2514/6.2001-5242.

[5]   Gianazza D, Guittet K. Selection and evaluation of air traffic complexity metrics. *25th Digital Avionics Systems Conference.* Portland, OR, USA: IEEE Press; 2006. DOI: 10.1109/DASC.2006.313710.

[6]   Xiao MM, et al. ATCEM: A synthetic model for evaluating air traffic complexity. *Journal of Advanced Transportation.* 2016;50:315-325. DOI: 10.1002/atr.1321.

[7]   Andraši P, et al. Subjective air traffic complexity estimation using artificial neural networks. *Promet –*

Traffic&Transportation. 2019;31(4):377-386. DOI: 10.7307/ptt.v31i4.3018.

[8]   Antulov-Fantulin B, et al. Determining air traffic complexity: Challenges and future development. *Promet – Traffic&Transportation*. 2020;32(4):475-485. DOI: 10.7307/ptt.v32i4.3401.

[9]   Zhu X. A semi-supervised learning method for air traffic complexity evaluation. *2017 Integrated Communications, Navigation and Surveillance Conference (ICNS) 2017, 18-20 Apr. 2017, Herndon, VA, USA*. IEEE Press; 2017. DOI: 10.1109/icnsurv.2017.8011950.

[10] Cao XB, et al. A knowledge-transfer-based learning framework for airspace operation complexity evaluation. *Transportation Research Part C*. 2018;95:61-81. DOI: 10.1016/j.trc.2018.07.008.

[11] Zhang ZX, et al. Unsupervised evaluation of airspace complexity based on kernel principal component analysis. *Acta Aeronautica et Astronautica Sinica*. 2019;40(08):322969. DOI: 10.7527/S1000-3893.2019.22969.

[12] Zhang J, et al. Establishing the correlation between complexity and performance for arrival operations. *Promet – Traffic&Transportation*. 2022;34(6):927-942. DOI: 10.7307/ptt.v34i6.4137.

[13] Mogford RH, et al. *The complexity construct in air traffic control: A review and synthesis of the literature.* Atlantic City, NJ: FAA William Hughes Technical Center; 1995.

[14] Kopardekar P. *Dynamic density: A review of proposed variables.* Washington, D. C.: FAA WJHTC Internal Document; 2000.

[15] Hilburn B. *Cognitive complexity in air traffic control: A literature review.* Bretigny-sur-Orge: Eurocontrol Experimental Research Centre, EEC; 2004.

[16] Delahaye D, Puechmorel S. Air traffic complexity: Towards intrinsic metrics. *Proceedings of 3rd FAA/ Eurocontrol Air Traffic Management R&D Seminar.* Washington, D. C.: FAA; 2000.

[17] Chi GT, Cao TT, Zhang K. The establishment of human all-around development evaluation indicators system based on correlation-principle component analysis. *Systems Engineering: Theory & Practice*. 2012;32(1):111-119. DOI: 10.12011/1000-6788(2012)1-111.

[18] Chen, HH. Method of screening evaluation indicators based on circle ill condition index analysis. *Chinese Journal of Management Science*. 2019;27(1):184-193. DOI:10.16381/j.cnki.issn1003-207x.2019.01.018.

[19] Li X, Bai YQ. Convergence analysis of an intrinsic steepest descent method on semi-supervised metric learning. *Operations Research Transactions*. 2017;21(03):1-13. DOI: 10.15960/j.cnki.issn.1007-6093.2017.03.001.

[20] Weinberger KQ, Saul LK. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*. 2009;10(9):207-244. DOI: 10.5555/1577069.1577078.

[21] Shi YN, Zheng GL. A review of three neural network methods for manufacturing feature recognition. *Acta Aeronautica et Astronautica Sinica*. 2019;40(09):182-198. DOI: 10.7527/S1000-6893.2019.22840.

隋东，李倩，周婷婷，刘珂琛

基于度量学习的空中交通态势评估研究

摘要

空中交通态势评估可以支持交通流和空域配置优化，提升空中交通管制安全水平。由于空中交通态势受到多种因素的交互影响，且缺乏真实标签数据，使得特征指标选取和态势评价成为难点。论文从空域结构、交通特征和气象条件三个维度初选指标，并基于信息重要度和重叠度定量筛选评价指标。依据流量控制和交通流信息，定义了态势区间的自由态和饱和态，提出了基于度量学习的态势样本标定方法。以多层感知机回归模型建立特征指标与空中交通态势的映射关系，对上海3个扇区样本集的评估准确率在80%以上，验证了态势评估模型的有效性。本文方法对保障空域运行安全具有现实意义。

关键词

空中交通安全；空中交通态势；评价指标；度量学习

Appendix A. Principal component factor loading matrix and screening results

Indicators	Principal components										Screening results
	1	2	3	4	5	6	7	8	9	10	
X_1	-0.355	0.506	0.076	-0.005	-0.270	0.402	**0.543**	0.258	0.020	-0.005	●
X_2	0.472	**-0.856**	-0.046	-0.060	0.067	-0.174	0.031	-0.030	-0.011	-0.005	●
X_3	-0.398	**0.809**	0.013	0.082	0.083	-0.027	-0.387	-0.124	0.002	0.010	●
X_4	0.416	**-0.829**	-0.018	-0.080	-0.060	-0.005	0.335	0.101	-0.004	-0.009	●
X_5	0.450	**-0.859**	-0.031	-0.072	-0.004	-0.081	0.205	0.044	-0.007	-0.008	●
X_6	0.411	**-0.824**	-0.017	-0.080	-0.066	0.004	0.350	0.108	-0.003	-0.009	●
X_7	-0.438	**0.705**	0.069	0.023	-0.203	0.336	0.336	0.179	0.018	-0.001	●
X_8	0.470	**-0.808**	-0.059	-0.044	0.136	-0.258	-0.147	-0.104	-0.015	-0.003	●
X_9	**0.771**	0.555	-0.066	-0.141	0.010	-0.076	0.017	-0.134	-0.030	-0.006	●
X_{10}	**0.668**	0.566	-0.077	-0.218	0.038	-0.153	-0.043	-0.068	-0.030	-0.017	●
X_{11}	0.510	-0.192	-0.078	-0.087	0.139	0.325	-0.280	0.032	0.027	-0.003	×
X_{12}	0.562	-0.085	-0.054	-0.158	0.033	0.512	-0.175	0.177	0.011	0.001	×
X_{13}	**0.796**	0.491	-0.066	-0.113	-0.005	-0.036	0.106	-0.197	-0.025	-0.001	●
X_{14}	**0.828**	0.406	-0.063	-0.096	-0.018	0.023	0.183	-0.216	-0.019	0.003	●
X_{15}	**0.836**	0.343	-0.060	-0.080	-0.029	0.061	0.226	-0.219	-0.019	0.009	●
X_{16}	**0.821**	0.279	-0.049	-0.059	-0.040	0.093	0.250	-0.214	-0.016	0.012	●
X_{17}	**0.750**	0.567	-0.063	-0.134	0.011	-0.087	0.003	-0.137	-0.031	-0.007	●
X_{18}	0.281	-0.280	0.020	0.151	-0.012	0.144	-0.066	0.042	0.017	0.110	×
X_{19}	-0.014	0.006	0.003	0.023	-0.011	-0.036	0.042	-0.052	**0.579**	**0.676**	●
X_{20}	0.010	0.002	-0.004	0.007	0.001	-0.010	0.016	-0.022	0.575	**-0.719**	●
X_{21}	0.414	-0.012	0.378	**0.664**	-0.176	-0.056	-0.036	0.151	-0.003	-0.039	●
X_{22}	**0.695**	0.059	0.263	0.575	-0.137	-0.038	-0.038	0.040	-0.014	-0.016	●
X_{23}	0.458	-0.242	-0.054	-0.076	0.084	**0.608**	-0.274	0.172	0.043	0.015	●
X_{24}	0.125	0.137	-0.005	-0.059	-0.006	-0.105	0.005	0.194	-0.090	0.061	×
X_{25}	**0.723**	-0.116	-0.083	-0.173	0.092	0.494	-0.243	0.125	0.019	-0.002	●
X_{26}	-0.062	0.048	-0.121	0.045	0.030	-0.012	0.002	0.170	-0.554	-0.037	×
X_{27}	0.117	0.004	**0.970**	-0.117	0.060	0.002	-0.010	-0.013	-0.009	0.000	●
X_{28}	0.453	0.086	-0.595	0.571	-0.163	-0.030	-0.030	0.065	0.015	-0.006	×
X_{29}	0.405	0.054	**0.865**	0.182	-0.019	-0.015	-0.030	0.021	-0.003	-0.003	●
X_{30}	0.390	0.145	-0.555	0.547	-0.144	-0.029	-0.053	0.016	0.003	-0.020	×
X_{31}	0.346	0.111	**0.834**	0.202	-0.027	-0.015	-0.043	0.008	0.007	-0.012	●
X_{32}	0.256	0.035	-0.357	0.365	-0.130	-0.079	-0.025	0.198	0.077	0.060	×
X_{33}	0.299	0.002	0.577	0.071	0.014	-0.048	-0.037	0.104	-0.026	0.044	×
X_{34}	0.465	0.424	-0.063	-0.321	-0.005	-0.396	-0.074	0.533	0.083	0.004	×
X_{35}	0.465	0.424	-0.063	-0.321	-0.005	-0.396	-0.074	0.533	0.083	0.004	×
X_{36}	-0.014	0.197	-0.041	0.257	**0.911**	0.023	0.211	0.096	0.022	0.001	●
X_{37}	0.008	0.222	-0.047	0.258	**0.903**	0.023	0.221	0.088	0.017	0.002	●

Note: ● represents retained indicators, × represents deleted indicators

Appendix B. Pearson correlation coefficient table

Indicators	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{29}	X_{31}	X_{36}	X_{37}
X_1	1	-.684**	.267**	-.338**	-.501**	-.318**	**.946****	-.833**	-.044**	-.045**	-.046**	-.047**	-.048**	-.049**	-.044**	-.062**	-.022**	0.004	.014*
X_2	-.684**	1	-.886**	**.918****	**.974****	**.909****	-.883**	**.973****	-.078**	-.115**	-.020**	.056**	.105**	.144**	-.095**	.097**	.023**	-.128**	-.138**
X_3	.267**	-.886**	1	**-.997****	**-.968****	**-.999****	.565**	-.756**	.131**	.181**	.055**	-.044**	-.108**	-.159**	.154**	-.089**	-.016**	.167**	.174**
X_4	-.338**	**.918****	**-.997****	1	**.984****	**1.000****	-.625**	.803**	-.124**	-.173**	-.050**	.047**	.109**	.159**	-.147**	.091**	.017**	-.163**	-.171**
X_5	-.501**	**.974****	**-.968****	**.984****	1	**.980****	-.754**	.896**	-.106**	-.151**	-.038**	.052**	.110**	.156**	-.127**	.096**	.020**	-.151**	-.160**
X_6	-.318**	**.909****	**-.999****	**1.000****	**.980****	1	-.608**	.790**	-.126**	-.176**	-.052**	.046**	.109**	.159**	-.149**	.091**	.017**	-.164**	-.172**
X_7	**.946****	-.883**	.565**	-.625**	-.754**	-.608**	1	**-.967****	0.006	.022**	-.021**	-.055**	-.078**	-.096**	.014*	-.083**	-.024**	.059**	.070**
X_8	-.833**	**.973****	-.756**	.803**	.896**	.790**	**-.967****	1	-.045**	-.073**	-0.001	.057**	.095**	.125**	-.058**	.093**	.024**	-.098**	-.109**
X_9	-.044**	-.078**	.131**	-.124**	-.106**	-.126**	0.006	-.045**	1	**.946****	**.934****	.882**	.835**	.776**	**.985****	.259**	.246**	.063**	.097**
X_{10}	-.045**	-.115**	.181**	-.173**	-.151**	-.176**	.022**	-.073**	**.946****	1	.844**	.761**	.698**	.625**	**.932****	.198**	.196**	.067**	.095**
X_{13}	-.046**	-.020**	.055**	-.050**	-.038**	-.052**	-.021**	-0.001	**.934****	.844**	1	**.945****	**.901****	.842**	**.919****	.260**	.239**	.056**	.089**
X_{14}	-.047**	.056**	-.044**	.047**	.052**	.046**	-.055**	.057**	.882**	.761**	**.945****	1	**.965****	**.915****	.864**	.268**	.240**	.046**	.078**
X_{15}	-.048**	.105**	-.108**	.109**	.110**	.109**	-.078**	.095**	.835**	.698**	**.901****	**.965****	1	**.963****	.816**	.271**	.238**	.037**	.067**
X_{16}	-.049**	.144**	-.159**	.159**	.156**	.159**	-.096**	.125**	.776**	.625**	.842**	**.915****	**.963****	1	.756**	.273**	.237**	.026**	.053**
X_{17}	-.044**	-.095**	.154**	-.147**	-.127**	-.149**	.014*	-.058**	**.985****	.932**	**.919****	.864**	.816**	.756**	1	.256**	.244**	.064**	.098**
X_{29}	-.062**	.097**	-.089**	.091**	.096**	.091**	-.083**	.093**	.259**	.198**	.260**	.268**	.271**	.273**	.256**	1	**.949****	-0.006	-0.001
X_{31}	-.022**	.023**	-.016**	.017**	.020**	.017**	-.024**	.024**	.246**	.196**	.239**	.240**	.238**	.237**	.244**	**.949****	1	0.002	0.009
X_{36}	0.004	-.128**	.167**	-.163**	-.151**	-.164**	.059**	-.098**	.063**	.067**	.056**	.046**	.037**	.026**	.064**	-0.006	0.002	1	**.984****
X_{37}	.014*	-.138**	.174**	-.171**	-.160**	-.172**	.070**	-.109**	.097**	.095**	.089**	.078**	.067**	.053**	.098**	-0.001	0.009	.984**	1

Appendix C. Table of coefficients of variation

Indicators	Coefficients of variation	Indicators	Coefficients of variation
X_1	0.0565	X_{16}	0.6050
X_2	0.1212	X_{17}	0.7193
X_3	0.1701	X_{19}	-313.4333
X_4	0.1729	X_{20}	188.3209
X_5	0.5021	X_{21}	2.4002
X_6	0.4720	X_{22}	1.0095
X_7	0.3840	X_{23}	1.7598
X_8	0.1768	X_{25}	1.4678
X_9	0.7071	X_{27}	9.3964
X_{10}	1.0144	X_{29}	1.9811
X_{13}	0.6739	X_{31}	2.1916
X_{14}	0.6413	X_{36}	1.9649
X_{15}	0.6236	X_{37}	1.9465

Appendix D. Metric matrix

$$M = \begin{pmatrix}
1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.3010 & 0.6871 & 0.0264 & -0.0012 & -0.0123 & 0.0964 & -0.0343 & 0.1185 & -0.0076 & 0.0366 & 0.1816 \\
0.0000 & 0.0000 & 0.6871 & 1.6211 & 0.0678 & -0.0029 & 0.0041 & 0.1502 & -0.0126 & 0.3553 & -0.0280 & 0.0742 & 0.4217 \\
0.0000 & 0.0000 & 0.0264 & 0.0678 & 0.0047 & -0.0001 & -0.0010 & 0.0126 & 0.0063 & 0.0292 & 0.0127 & 0.0038 & 0.0186 \\
0.0000 & 0.0000 & -0.0012 & -0.0029 & -0.0001 & 0.0000 & -0.0008 & 0.0020 & 0.0003 & -0.0007 & -0.0004 & -0.0003 & -0.0007 \\
0.0000 & 0.0000 & -0.0123 & 0.0041 & -0.0010 & -0.0008 & 0.1281 & -0.3510 & -0.0192 & -0.0044 & -0.0025 & 0.0068 & -0.0084 \\
0.0000 & 0.0000 & 0.0964 & 0.1502 & 0.0126 & 0.0020 & -0.3510 & 0.9840 & 0.0698 & 0.0765 & 0.0213 & -0.0120 & 0.0656 \\
0.0000 & 0.0000 & -0.0343 & -0.0126 & 0.0063 & 0.0003 & -0.0192 & 0.0698 & 0.1281 & 0.1093 & -0.0568 & -0.0309 & -0.0130 \\
0.0000 & 0.0000 & 0.1185 & 0.3553 & 0.0292 & -0.0007 & -0.0044 & 0.0765 & 0.1093 & 0.2260 & 0.0412 & 0.0033 & 0.0916 \\
0.0000 & 0.0000 & -0.0076 & -0.0280 & 0.0127 & -0.0004 & -0.0025 & 0.0213 & -0.0568 & 0.0412 & 0.2237 & 0.0389 & 0.0153 \\
0.0000 & 0.0000 & 0.0366 & 0.0742 & 0.0038 & -0.0003 & 0.0068 & -0.0120 & -0.0309 & 0.0033 & 0.0389 & 0.0143 & 0.0238 \\
0.0000 & 0.0000 & 0.1816 & 0.4217 & 0.0186 & -0.0007 & -0.0084 & 0.0656 & -0.0130 & 0.0916 & 0.0153 & 0.0238 & 0.1129
\end{pmatrix}$$