Junfeng ZHANG, Ph.D.[1]
(Corresponding author)
E-mail: zhangjunfeng@nuaa.edu.cn
Tong XIANG, Master candidate[1]
E-mail: yarck01@nuaa.edu.cn
Ming ZHOU, Master candidate[1]
E-mail: mmxyh@nuaa.edu.cn
Bin WANG, Senior engineer[2]
E-mail: teddy11@163.com
[1] College of Civil Aviation
Nanjing University of Aeronautics and Astronautics
No.29, Jiangjun Avenue, Jiangning District, Nanjing, Jiangsu, 211106, China
[2] Central and Southern Regional Air Traffic Management Bureau of the Civil Aviation Administration of China
No.3, Nanyun East Street, Baiyun District, Guangzhou, Guangdong, 510403, China

Traffic Engineering
Original Scientific Paper
Submitted: 1 Apr. 2022
Accepted: 13 July 2022

ESTABLISHING THE CORRELATION BETWEEN COMPLEXITY AND PERFORMANCE FOR ARRIVAL OPERATIONS

ABSTRACT

Air traffic complexity indicators play an essential role in measuring operational performance and controller workload. However, current studies mainly depend on the manual scoring method to scale performance or workload. This paper focuses on arrival operations and presents a data-driven strategy to establish the correlation between complexity and performance to avoid the subjectivity of the currently used manual scoring method. Firstly, we present twenty-six indicators for describing air traffic complexity and two indicators for arrival operational performance. Secondly, the clustering method distinguishes peak and off-peak situations for arrival operation. Moreover, clustering results are compared to investigate the correlation between complexity and performance initially. Thirdly, the classification method is adopted to determine such correlation further. In addition, we also identify the affecting factors which could influence operational performance. Finally, trajectories of arrival aircraft landing at Guangzhou Baiyun International Airport (ZGGG) are used for case validation. The results indicate that there is a strong correlation between complexity and performance. The accuracy and precision of classification are approximately 90%. Furthermore, the number of aircraft significantly impacts the arrival operational performance within TMA.

KEYWORDS

air traffic; air traffic complexity; complexity indicators; performance; correlation.

1. INTRODUCTION

Air traffic management (ATM) plays an essential role in air transportation by accelerating air traffic flow in a conflict-free environment. Unfortunately, the past two years witnessed a significant impact on air transportation due to the coronavirus (COVID-19) pandemic. As a consequence, air traffic growth slowed down. However, the civil aviation industry will regain its vigour with several control measures. At this time, the ever-increasing demand for air traffic is doomed to cause airspace congestion, substantial flight delays, excessive fuel consumption and consequential air pollutant emissions, which pose a significant challenge to the ATM system [1]. Therefore, measuring air traffic control (ATC) performance and identifying the affecting factors could help enhance the levels of safety, capacity, efficiency and environmental sustainability [2]. Air traffic complexity is a critical limiting factor to increasing performance in the current ATC environment [3].

Firstly, Laudeman et al. [4] presented a mathematical model of dynamic density, an aggregate complexity metric. Eight factors were combined to represent the controller's workload. Unfortunately, the choice of factors is often related to the specific environment. Therefore, it hindered the efforts to form a unified definition of air traffic complexity. As a result, research on dynamic density has been further expanded to collecting affecting factors or variables [5]. For example, Masalonis et al. [6] selected 12 metrics from a list of 41 potential complexity metrics and used multiple regression analysis to determine the importance of the different metrics. They also analysed the correlation between dynamic density and complexity (controller workload). Besides dynamic density, Netjasov et al. [7, 8, 9] developed a generic metric for measuring the complexity of Terminal Airspace (TMA). The metric was defined from static and dynamic perspectives and considered arrival and departure traffic.

Secondly, due to the subjective nature of the controller workload, some scholars did not believe there was a correlation between complexity factors and controller work [10]. Therefore, Delahaye et al. [11] tried to objectively evaluate the air traffic complexity based on the intrinsic metrics derived from the geometric properties of the aircraft pairs, including density, convergence/divergence and sensitivity. Furthermore, they extended the previous results and presented a new air traffic complexity metric based on a non-linear vector field model of air traffic [12]. In addition, they developed a novel macroscopic traffic assignment model to mitigate the congestion of intensive urban air mobility operations [13]. Besides, Wang et al. [14] proposed a complex networks based method to describe air traffic complexity and elaborate its evolutionary laws. Wee et al. [15] developed a dynamic tactical complexity model, Conflict Activity Level, by establishing an overall score for an entire region or sub-regions of interest.

Thirdly, some machine learning methods have been implemented since it was hard to model the correlation between complexity factors and controller work straightforwardly [16]. Gianazza and Guittet [17] summarised common complexity factors, used principal component analysis for factors reduction and implemented a neural network to find the correlation between complexity and workload. Xiao et al. [18] developed an integrated classifier, Zhu et al. [19] presented a new ensemble learning model, and Cao et al. [20] proposed a learning framework based on knowledge transfer to construct the correlation between complexity factors and complexity levels. Radišić et al. [21] put forward novel complexity indicators under trajectory-based operation, and Andraši [22] estimated the air traffic complexity using neural networks. Xie et al. [23] used images instead of complexity factors to evaluate the operational complexity by deep convolutional neural networks.

In most previous studies, it was the responsibility of air traffic controllers (ATCOs) to determine the complexity levels. However, the manual scoring process is time-consuming, leading to only small samples used for complexity analysis. Besides, the manual scoring process is a challenging task to avoid the problem of subjectivity. Therefore, this paper, on the one hand, substitutes complexity levels with the operational performance; on the other hand, it attempts to establish the correlation between air traffic complexity and performance. Such methodology could ensure objective data-driven assessment during the entire process. Unfortunately, the results provided by Standfuss and Rosenow [24] showed that there is no statistically significant correlation between complexity and productivity (operational performance). However, complexity and productivity indicators were derived from a macroscopic perspective, while our study will focus on the microscopic perspective. More specifically, our study attempts to establish the correlation between air traffic complexity and performance for arrival operation within TMA.

2. METHODOLOGY

2.1 Overall framework

This paper proposes a method to identify air traffic complexity, avoid manual scoring and further establish the correlation between air traffic complexity and control performance. The proposed method could be divided into three steps, as summarised in *Figure 1*.

Step A: Data Preparation. The original data is aircraft tracking messages received by the Surveillance Data Processing System (SDPS). After decoding these aircraft tracking messages, raw data ($\bar{p}^M$) is obtained, including a sequence of trajectories $\{tr_1, tr_2, \ldots, tr_i, \ldots, tr_m\}$. Each trajectory consists of a series of points: $\{p_1^i, p_2^i, \ldots, p_{L(i)}^i\}$, where $L_{(i)}$ represents the number of points in the trajectory p^M. Each point contains the target aircraft's timestamp, flight ID, position (latitude, longitude, altitude), heading and speed (ground and vertical speeds). Subsequently, the data pre-processing step is necessary to identify and modify the target trajectories' incomplete, incorrect or irrelevant parts. Moreover, the scope of the target trajectories should be defined according to arrival operation, which will be presented in Section 3.

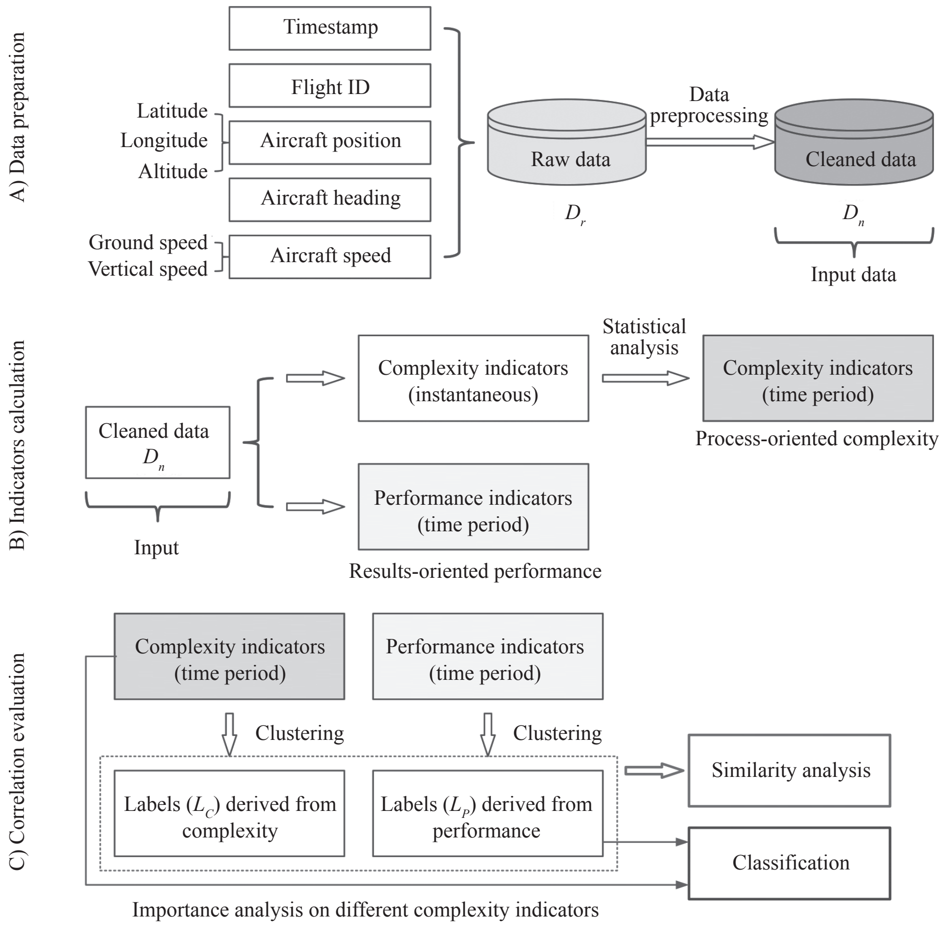


Figure 1 – Overall framework of the proposed method

Step B: Indicator Calculation. This step focuses on two types of indicator calculation: one is complexity indicator calculation, and the other is the performance indicator calculation. The former reflects the complexity of the control process, while the latter describes the performance of the control results. It should be noted that the process-oriented indicators are often defined as instantaneous metrics, and the results-oriented indicators are always defined as metrics over a period. Therefore, the process-oriented indicators should be transformed into metrics over a corresponding period using the statistical analysis method. A detailed description will be presented in subsection 2.2.

Step C: Correlation Evaluation. This step employs machine learning methods to establish the correlation between process-oriented complexity indicators and result-oriented performance indicators. Such methodology could avoid the subjectivity associated with the manual scoring.

On the one hand, the clustering method is employed on complexity indicators and performance indicators, respectively. Subsequently, the similarity between the two clustering results is analysed to investigate whether there is a correlation between complexity and performance or not.

On the other hand, the classification method is employed on a data set to further explore the correlation between complexity and performance. The data set comprises the complexity indicators, which are treated as features (input), and the clustering results of performance indicators, which are treated as labels (output). A detailed description will be presented in subsection 2.3.

2.2 Complexity and performance framework

Complexity indicators for arrival operation

Most attention has been paid to en-route operation in previous research on air traffic complexity. In this study, we focus on the complexity within TMA since operation in TMA plays a critical role in the whole ATM system [25]. Furthermore, the complexity of only arrival operations is considered for several reasons. First, arrival and departure operations are mutually independent in most airports around China by lateral or vertical segregated methods. Second, arrival operation requires more attention from controllers to maintain safe separation and establish an approach sequence. Third, the traffic situation of an arrival operation presents converging characteristics, which complexity indicators could describe.

Kopardekar [26] and Gianazza [17] laid the foundation for complexity indicator definition. Most subsequent studies derived complexity indicators from their initial efforts. In this paper, we presented the complexity indicators based on the combination of previous studies, as shown in *Table 1*.

It should be noted from *Table 1* that several complexity indicators are process-oriented and described as instantaneous metrics at each minute. Therefore, we need to transform these process-oriented instantaneous indicators into period metrics, i.e. interval indicators, to establish the correlation between complexity and performance. Therefore, the time interval in this paper is determined as 15 min due to the same time interval having been used in the previous studies [26].

It is also worth mentioning that the complexity indicators from previous work [17] were mainly used in the en-route phase. However, in the TMA domain, the complexity indicators, like density/convergency/sensitivity/collision risk, will have unexpected values when the aircraft are on Final (the runway extension for landing aircraft to conduct the final approach). Therefore, we only take the arrival operation from the entry fix to the Final interception in our work.

Performance indicators for arrival operation

After obtaining the complexity indicators, the performance indicators should be defined to establish the correlation between complexity and performance for the arrival operation. Therefore, the selection of performance indicators for arrival operation becomes crucial.

Key Performance Areas (KPAs) and Key Performance Indicators (KPIs) are usually adopted to measure air traffic control performance. For arrival operation, we need to focus on three KPAs, i.e. capacity, efficiency and environment [27]. Four KPIs in our previous studies were proposed which closely related to these KPAs, such as airport acceptance rate, final approach separation, flight time within TMA, and taxi-in time [28]. In this paper, taxi-in time is not considered since we only focus on the complexity indicators in the air. Airport acceptance rate and final approach separation could be transformed into the number of landings of arrival aircraft within a given period. Flight time within TMA is a critical indicator for arrival performance measurement related to flight efficiency and controller workload [29]. It is worth noting that flight distance and fuel consumption are significant ATC performance objectives. However, those indicators highly correlate with flight time [27]. Therefore, number of landings and flight time are selected as candidate indicators for arrival performance measurement.

Unfortunately, arrival flight time is affected by different runway uses, entering fixes, aircraft types and arrival routes. Given that, EUROCONTROL proposes an additional ASMA (Arrival Sequencing and Metering Area) time [30] to address the impact of these factors. Therefore, the additional ASMA time, a proxy for the average arrival runway queuing time on the arrival traffic flow, is selected as a substitute for arrival flight time. Please refer to [29] for a detailed definition and methodology of the additional ASMA time (hereafter referred to as additional time).

The number of landings N^{LD} within 15 min could be used directly for performance measurement. However, the additional time T^{Add} within 15 min needs to be further processed for performance measurement. Therefore, only the partial additional time of a specific aircraft should be allocated to the given target of 15 min. For example, as shown in *Figure 2*, the additional time within 15 min of arrival aircraft #AC3 is equal to the original additional time while the additional times within 15 min of #AC1, #AC2, #AC4, and #AC5 are 1/2, 4/7, 1/2, and 1/3 of the original additional time. These proportions are derived from the share of flight time within the given target of 15 min.

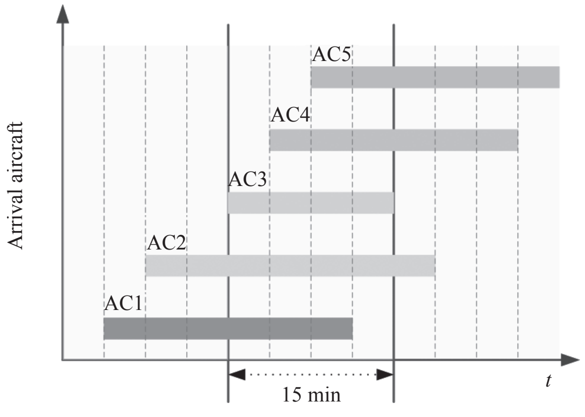


Figure 2 – Illustration of processed additional time

Table 1 – Definition of complexity Indicators

Item	Description	In our Study
Density D_{ij}	$D_{ij} = 0.5\cdot\left(e^{-\alpha(d_{ij})^2}+e^{-\beta d_{ij}}\right), d_{ij}=\sqrt{\langle \overrightarrow{x_i x_j}, \overrightarrow{x_i x_j}\rangle}$ where α=0.002, β=0.01, and x_i represents the position of aircraft i	-
Maximum density D^{Max}	$D^{Max}=\max\{D_{ij}\}$	$\bar{D}^{Max}$
Accumulative density D^{Sum}	$D^{Sum}=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\{D_{ij}\}$	$\bar{D}^{Sum}$
Mean density D^{Mean}	$D^{Mean}=\frac{D^{Sum}}{\sum_{i=1}^{n-1} i}$	$\bar{D}^{Mean}$
Convergency C_{ij}	$C_{ij}=\vec{v}_{ij}\cdot 1_{\mathrm{R}^+}\cdot\lVert\vec{v}_{ij}\rVert\cdot D_{ij}$, where v_i represents the velocity of aircraft i	-
Maximum convergency C^{Max}	$C^{Max}=\max\{C_{ij}\}$	$\bar{C}^{Max}$
Accu. convergency C^{Sum}	$C^{Sum}=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\{C_{ij}\}$	$\bar{C}^{Sum}$
Mean convergency C^{Mean}	$C^{Mean}=\frac{C^{Sum}}{\mathrm{Count}(C_{ij}\neq 0)}$	$\bar{C}^{Mean}$
Sensitivity S_{ij}	$S_{ij}=\vec{v}_{ij}\cdot 1_{\mathrm{R}^+}\cdot\lVert\overrightarrow{\nabla v}_{ij}\rVert\cdot D_{ij}$	-
Minimum sensitivity S^{Min}	$S^{Min}=\min\{S_{ij}\}$	$\bar{S}^{Min}$
Mean sensitivity S^{Mean}	$S^{Mean}=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\{S_{ij}\}/\mathrm{Count}(C_{ij}\neq 0)$	$\bar{S}^{Mean}$
Insensitivity I_{ij}	$I_{ij}=C_{ij}^2/S_{ij}$	-
Maximum insensitivity I^{Max}	$I^{Max}=\max\{I_{ij}\}$	$\bar{I}^{Max}$
Accu. insensitivity I^{Sum}	$I^{Sum}=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\{I_{ij}\}$	$\bar{I}^{Sum}$
Mean insensitivity I^{Mean}	$I^{Mean}=\frac{I^{Sum}}{\mathrm{Count}(I_{ij}\neq 0)}$	$\bar{I}^{Mean}$
Number of aircraft n	Aircraft number of arrivals within TMA	$\bar{n}$
Squared number of aircraft n^2	Square of aircraft number of arrivals within TMA	$\bar{n}^2$
Controlled number of aircraft N	Aircraft handled by controllers in current 15 minutes	N
Arrival traffic level TL	Mean number of aircraft in the last 15 minutes	TL
Arrival traffic demand TD	Number of aircraft entering the TMA 10-25 mins later	TD
Speed average μ^{GS}	Mean ground speed for arrivals within TMA	$\bar{\mu}^{GS}$
Speed variance σ^{GS}	The variance of ground speed for arrivals within TMA	$\bar{\sigma}^{GS}$
Altitude average μ^{Alt}	Mean altitude for arrivals within TMA	$\bar{\mu}^{Alt}$
Altitude variance σ^{Alt}	The variance of altitude for arrivals within TMA	$\bar{\sigma}^{Alt}$
Heading average μ^{H}	Mean heading for arrivals within TMA	$\bar{\mu}^{H}$
Heading variance σ^{H}	The variance of heading for arrivals within TMA	$\bar{\sigma}^{H}$
The ratio of medium class p^M	The percentage of medium-class aircraft	$\bar{p}^M$
The ratio of heavy class p^H	The percentage of heavy-class aircraft	$\bar{p}^H$
Approaching traffic $MD_{5\text{-}10}$	The number of pairs whose distance is between 5 and 10 NM	$MD_{5\text{-}10}$
Collision risk CR	$CR=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\frac{1}{\left(\frac{d_{ij}}{c}\right)^{\lambda}}$ where λ is a weighting factor, $\lambda=3$; c is a normalisation constant, c = 5nm	$\overline{CR}$

2.3 Correlation evaluation

The clustering method is used in this paper to figure out air traffic complexity and operational performance levels based on the complexity indicators and performance indicators. It is different from the manual scoring method, which is widely adopted [18–23] but influenced by the prior experience of different experts. The K-means is exploited as an alternative algorithm to conduct clustering because it is simple and easy to implement [31]. After clustering, the clustered labels could help identify the complexity and performance levels.

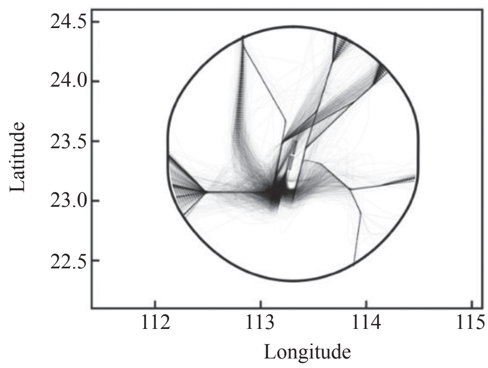
Subsequently, the correlation between complexity and performance is evaluated from two perspectives. On the one hand, the labels of complexity and performance are compared to analyse the correlation between process-oriented complexity and results-oriented operational performance. On the other hand, the complexity indicators are treated as inputs (x) and performance labels as outputs (y). Then, classification methods, like Support Vector Machine (SVM) [32], Logistic Regression (LR) [33] and Random Forest (RF) [34], are utilised to evaluate the correlation between complexity and performance. In addition, the RF method is also used for identifying the importance of different affecting factors.

3. CASE STUDY

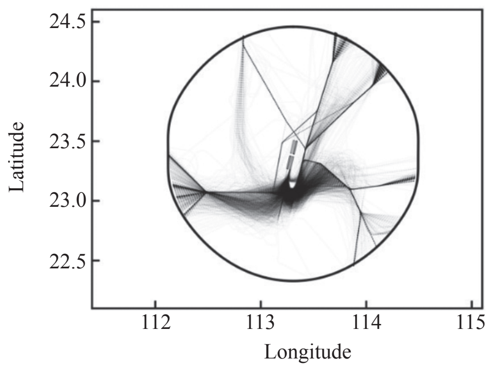
3.1 Data preparation

Guangzhou Baiyun International Airport (ZGGG) has been selected for the case study in this paper. ZGGG is one of the busiest airports in China, with three runways: Runway 01/19, Runway 02R/20L and Runway 02L/20R. The former two runways are used for landing, in which Runway 01/19 is for mixed arrival and departure operation while Runway 02R/20L is only for arrival operation. To make a more explicit definition of the range of TMA in this study, we focus on the area within a black circle as shown in *Figure 3*, with a radius of 65NM from the Aerodrome Reference Point.

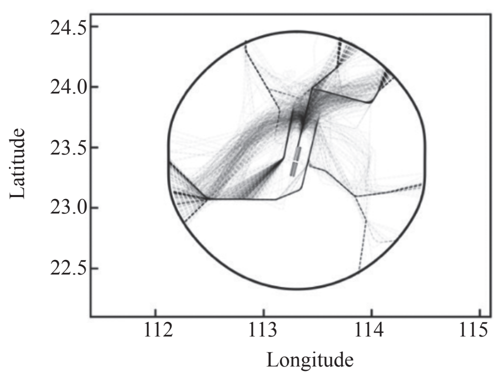
The arrival trajectories in *Figure 3* are obtained by pre-processing the original data received from the Central and Southern Regional Air Traffic Management Bureau, which is from 00:00 on 1 December 2019 to 11:59 on 31 December 2019. A



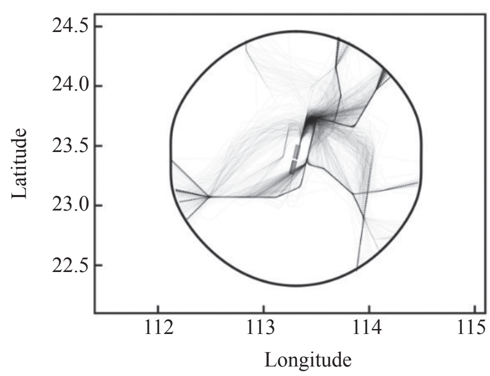
a) Arrival trajectories on Runway 01



b) Arrival trajectories on Runway 02R

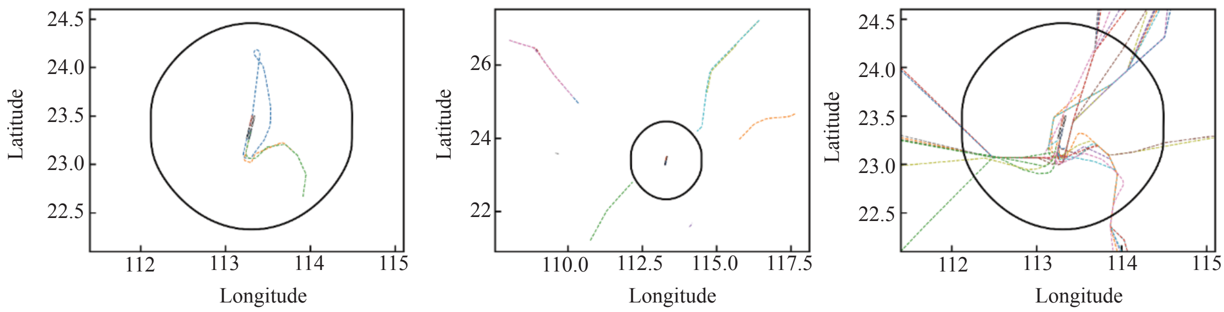


c) Arrival trajectories on Runway 19



d) Arrival trajectories on Runway 20L

Figure 3 – Arrival trajectories of ZGGG airport



a) Scenario 1 *b) Scenario 2* *c) Scenario 3*

Figure 4 – Outliers of arrival trajectories of ZGGG airport

complete arrival trajectory contains the following information: flight ID, runway in use, timestamp, latitude, longitude, altitude, ground speed, and vertical speed.

After data pre-processing, 20,846 arrival trajectories remain, as shown in *Figure 3*. *Figure 4* provides three scenarios of data pre-processing. The first one is due to the starting point of arrival trajectory located within the TMA, for example, a test flight. The second one is due to data missing, in which the ending point of arrival trajectory is out of the TMA. The third one is due to the ending point of the arrival trajectory being far away from the runway but still on Final. Therefore, the trajectories in Scenario 1 and Scenario 2 are discarded, and the trajectories in Scenario 3 are used in this research.

3.2 Indicator calculation

Figure 5 provides the distribution of daily traffic for the whole month. The number of daily number of arrival aircraft does not fluctuate much over the 31 days in December.

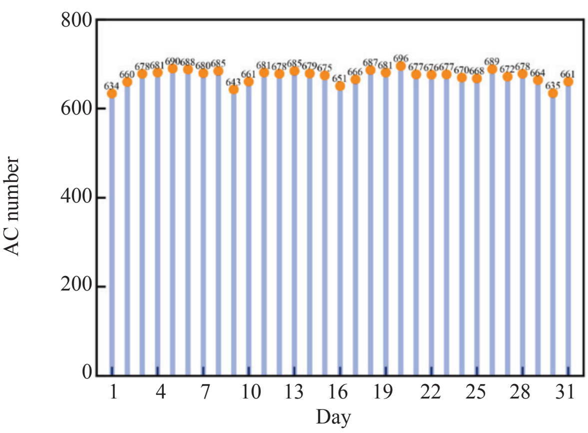


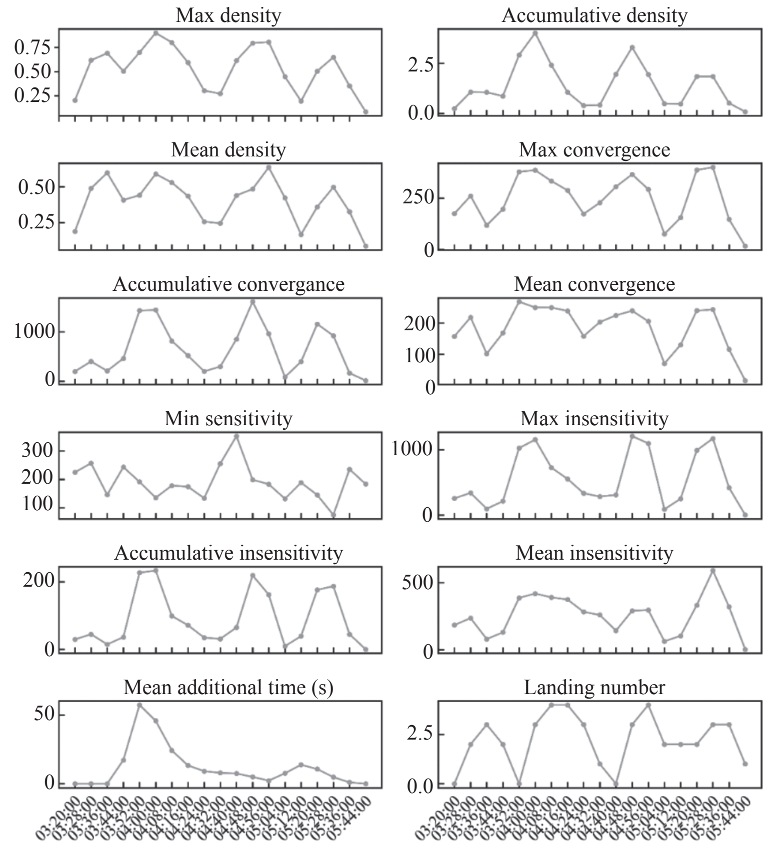
Figure 5 – Distribution of daily traffic

The indicators, defined in subsection 2.2, could be obtained based on the pre-processed data. Firstly, the process-oriented complexity indicators at each minute are calculated. Secondly, these process-oriented instantaneous indicators are transformed into interval indicators within 15 minutes. Thirdly, those interval indicators are presented from two perspectives. One is from the perspective of time series, while the other is from statistical analysis.

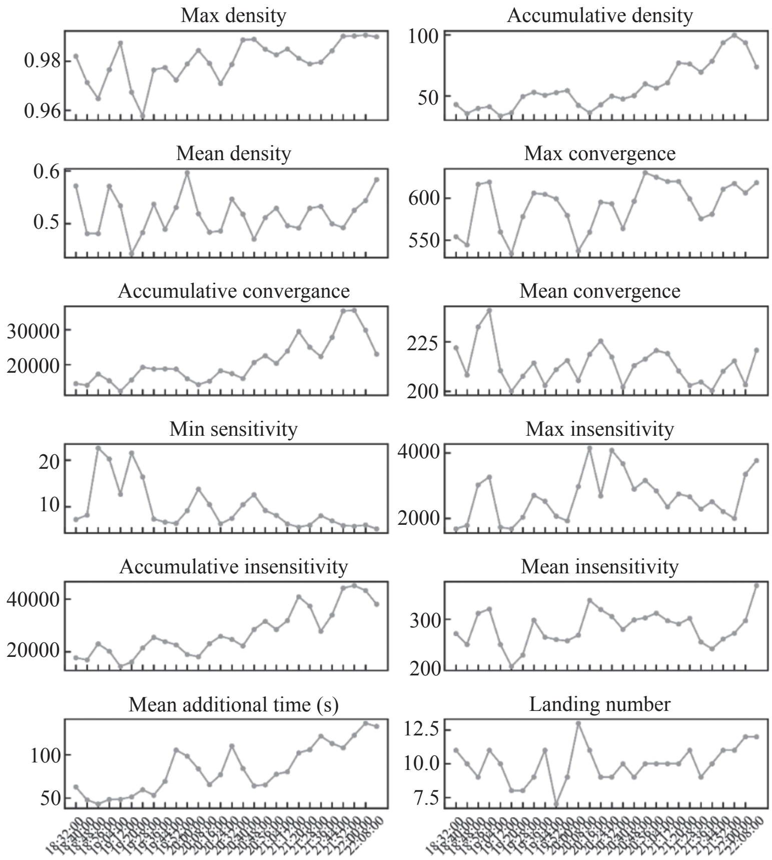
Figure 6 provides interval indicators from the perspective of time series, in which there are two typical scenarios. One is from 03:20 to 05:44 on 5 December, a typical scenario of free time, while the other is from 18:32 to 22:08 on 5 December, a typical scenario of rush hour. From *Figure 6*, we could find that the complexity indicators vary with time and depend on the traffic volume. More traffic volume leads to more significant complexity indicators. Furthermore, the rise and fall of traffic demand result in the fluctuation of the complexity indicators.

Figure 7 provides the current number of arrival aircraft distribution, which could help present the interval indicators from the statistical analysis perspective. For example, from *Figure 7*, we could find that the maximum number of aircraft within the ZGGG TMA is 24, while the most common situation is that there are 9 to 16 arrival aircraft within the TMA.

Figure 8 provides the boxplots of interval indicators from the perspective of statistical analysis, including the complexity indicators of density, convergence and sensitivity. It could be concluded from *Figure 8* that: (1) the accumulative density and convergence rise with the current number of arrival aircraft increase; (2) the mean density, convergence and sensitivity remain unchanged. However, the mean density, convergence and sensitivity distribution become more compact with the number of arrival aircraft increase. (3) The extremum density, convergence and sensitivity rise or fall with the



a) Typical off-peak scenario



b) Typical peak scenario

Figure 6 – Interval indicators from the perspective of time series

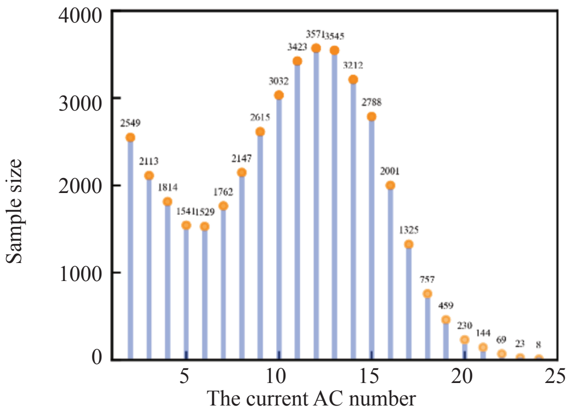


Figure 7 – Distribution of the current aircraft number

current number of arrival aircraft increase. Furthermore, the extremum density, convergence, and sensitivity tend to be unchanged when the number of arrival aircraft is more than 10.

Finally, the boxplot of the additional time with the current number of arrival aircraft is shown in *Figure 9*. The additional time rises with the current number of arrival aircraft increase. It indicates that more arrival aircraft within the TMA will lead to more possibilities for controllers to vector the arrival aircraft for establishing sequence and maintaining separation.

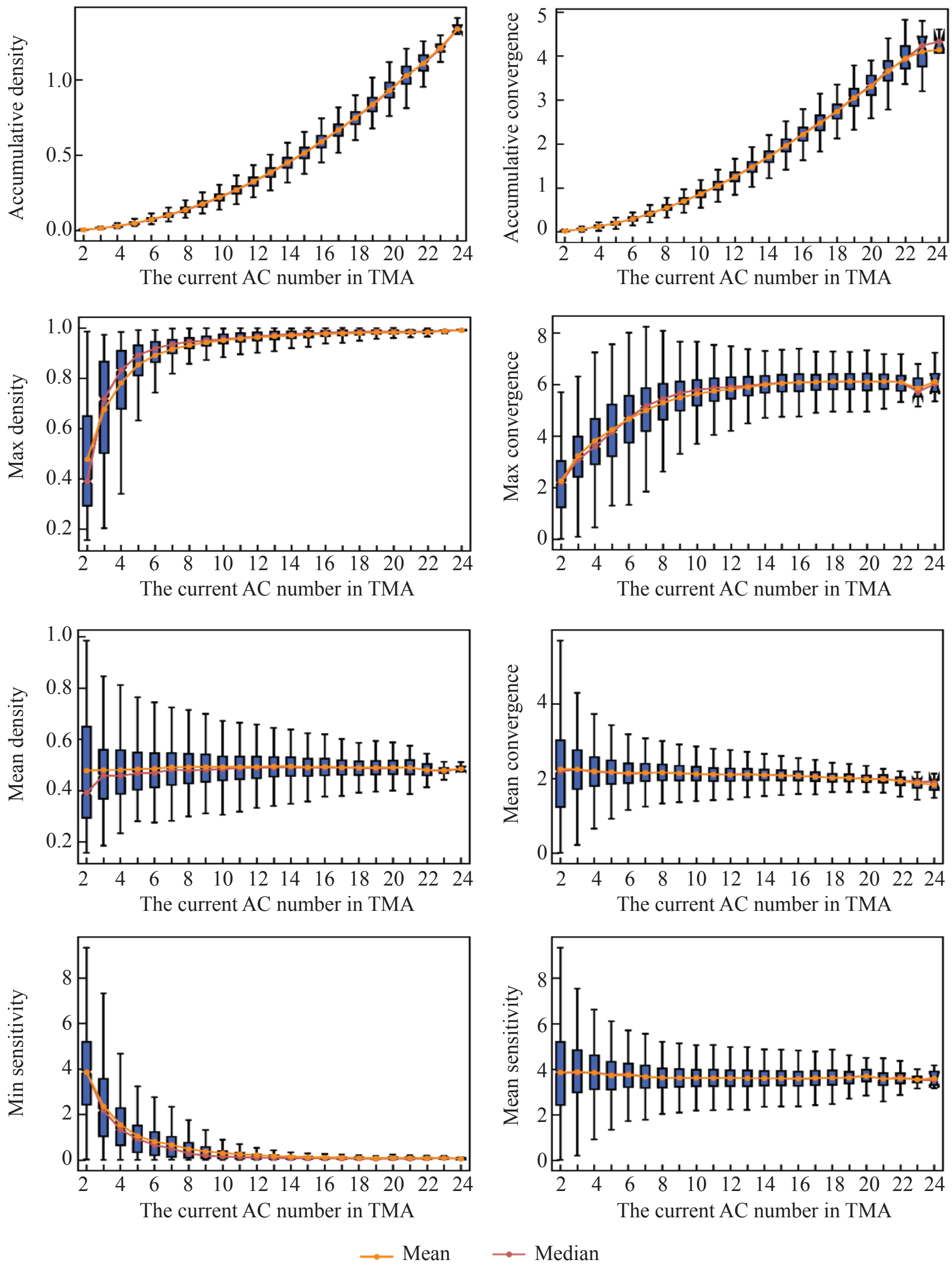


Figure 8 – Interval indicators from the perspective of statistical analysis

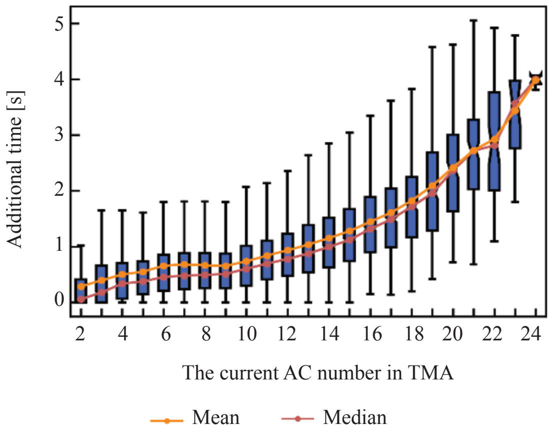


Figure 9 – Performance indicator from the perspective of statistical analysis

4. RESULTS AND DISCUSSION

4.1 Clustering on performance and complexity indicators

The clustering is firstly conducted on the performance indicators to avoid manual scoring. It is worth mentioning that we are interested in the busy time of arrival operation. For example, when the landing aircraft number is lower than a certain threshold, the air traffic is considered easy to control from the perspective of ATCOs. Therefore, this paper only focused on samples whose landing aircraft number is higher than a certain threshold. Furthermore, the number of clusters is set to two, representing off-peak and peak situations of arrival traffic during the clustering process on the performance indicators.

The K-means algorithm is used to cluster the performance indicators (number of landings and additional time), in which K=2. *Figure 10* provides the clustering results on the performance indicators when the thresholds of landing aircraft numbers are defined as 4, 5, 6 and 7.

As shown in *Figure 10*, the distinction between the two clusters, off-peak and peak, is dependent more on number of landings and less on additional time. We speculate mainly because arrival aircraft within TMA do not need to be vectored to produce additional times in the off-peak situation. However, as landing demand increases, arrival aircraft within TMA gradually need to be vectored to match the runway throughput, which results in unavoidable additional time.

Such speculation is also verified by the distribution of number of landings and additional time of the clustering results, as shown in *Figure 11*.

From *Figure 11*, we can find that the median of additional time is less than 50 seconds and the median of number of landings is less than eight under

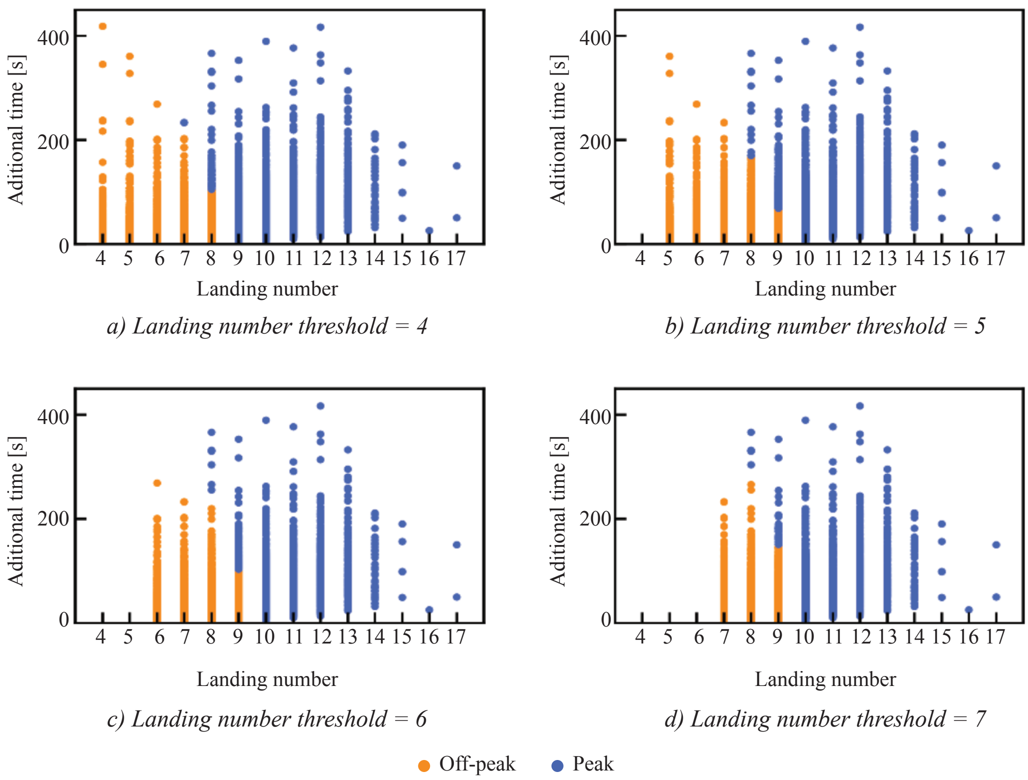
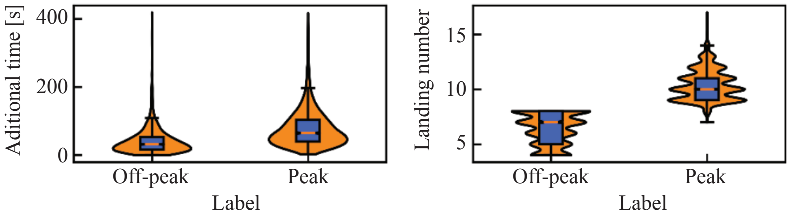
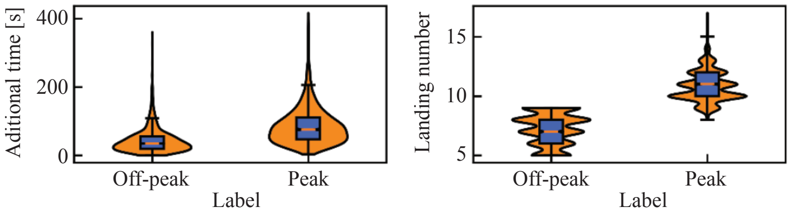


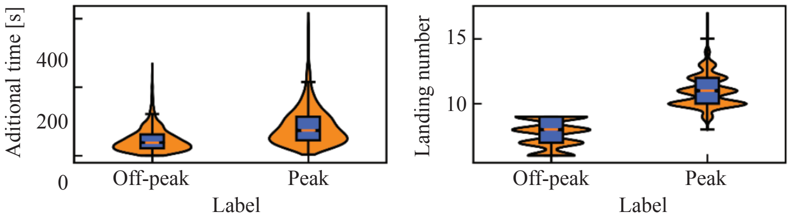
Figure 10 – The clustering results of performance indicators under different number of landings thresholds



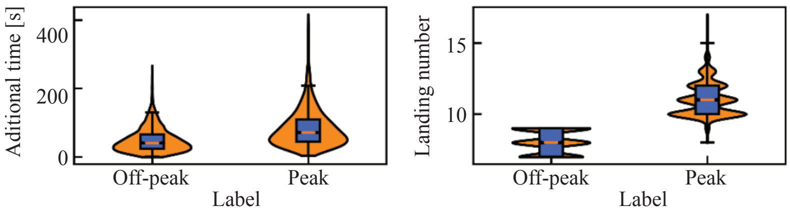
a) Landing number threshold = 4



b) Landing number threshold = 5



c) Landing number threshold = 6



d) Landing number threshold = 7

Figure 11 – Distribution of performance clustering results under different number of landings thresholds

all thresholds during the off-peak time. However, during the peak time, the median of additional time is nearly 100 seconds, and the median of number of landings is more than ten under all thresholds. These findings indicate that the clustering method based on performance indicators could be a viable alternative for performance evaluation instead of manual scoring.

In this paper, the Silhouette Coefficient is also used to determine the best number of landings threshold by measuring the goodness of the clustering algorithm. As shown in *Figure 12*, the best threshold in this study is five since the best clustering result is obtained under this situation.

The clustering is also conducted on the complexity indicators to establish the correlation between complexity and performance. The K-means algorithm is used to cluster the complexity indicators. The number of clusters is also set to two, representing off-peak and peak situations of arrival traffic, as clustering on performance indicators. *Figure 13* provides the clustering results on the complexity indicators when the number of landings threshold is defined as five.

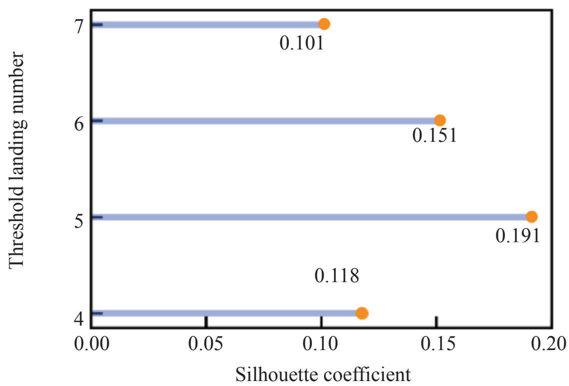


Figure 12 – Silhouette Coefficient in different landing number thresholds

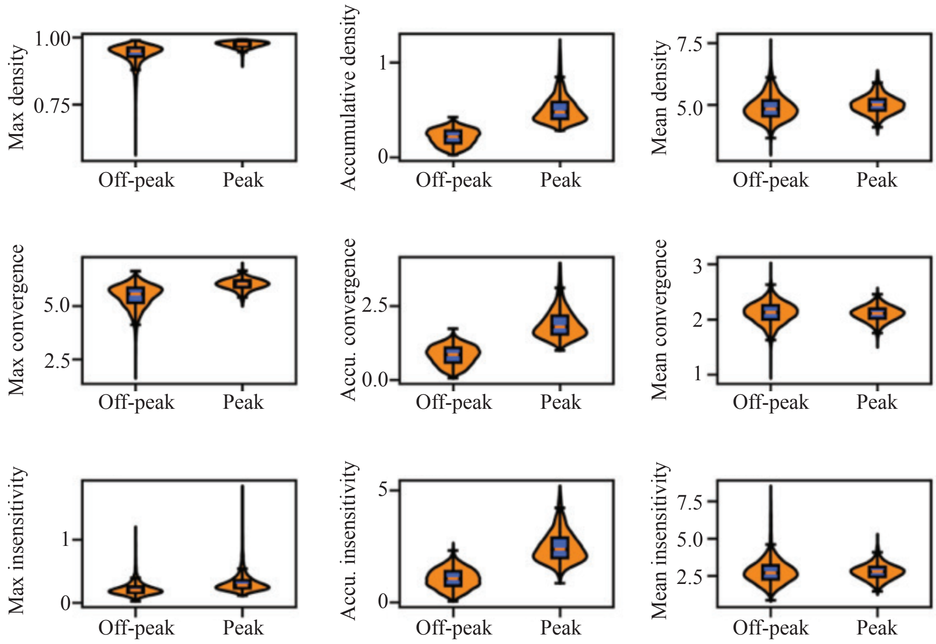


Figure 13 – Distribution of complexity clustering results - landing number threshold = 5

4.2 Clustering on performance and complexity indicators

This subsection offers three perspectives on the correlation between complexity and performance, including time series, comparison and classification.

Firstly, we provide a time series perspective based on the clustering results. For example, *Figure 14* presents the clustering results as a time series during a period on a specific day. In addition, the landing number threshold is defined as five. The majority of performance clusters coincide with the complexity clusters. However, there are two exceptions that can be seen from *Figure 14*. One is that the performance clusters sometimes change rapidly. The other is that sometimes there is a delay between the change of complexity clusters and the performance clusters, and sometimes there is not.

Secondly, we provide a comparison perspective based on the clustering results. Four metrics are adopted to measure the correlation between the clustering results of performance and complexity indicators, including precision, recall, accuracy and F_1 score.

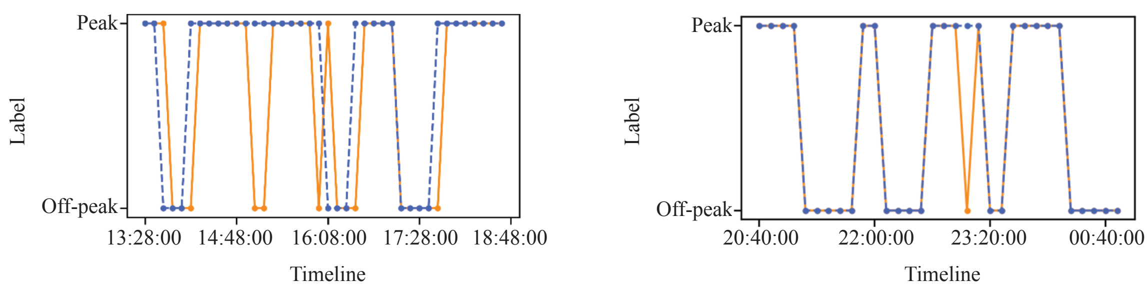
$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

In these metrics, True-Positive (TP) and False-Positive (FP) are the numbers of peak situations correctly and incorrectly clustered. At the



a) With a delay *b) Without a delay*

Figure 14 – The time series perspective based on the clustering results

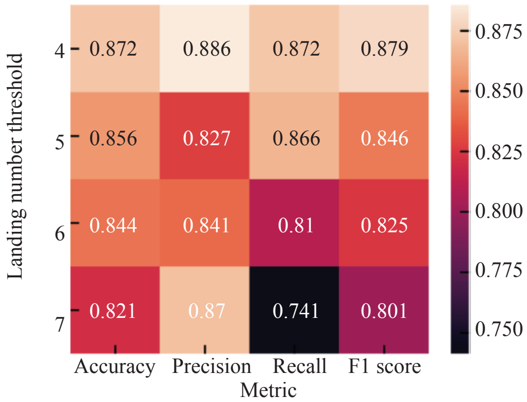


Figure 15 – The comparison perspective based on the clustering results

same time, True-Negative (*TN*) and False-Negative (*FN*) denote the numbers of off-peak situations correctly and incorrectly clustered.

These four metrics under different thresholds are shown in *Figure 15*. The values of these four metrics in *Figure 15* indicate that the complexity and performance indicators are highly correlated. That is to say, there exists a correlation between complexity and performance.

Thirdly, we provide a classification perspective for analysing the correlation between complexity and performance. As shown in *Figure 1*, the data set for classification is defined by choosing

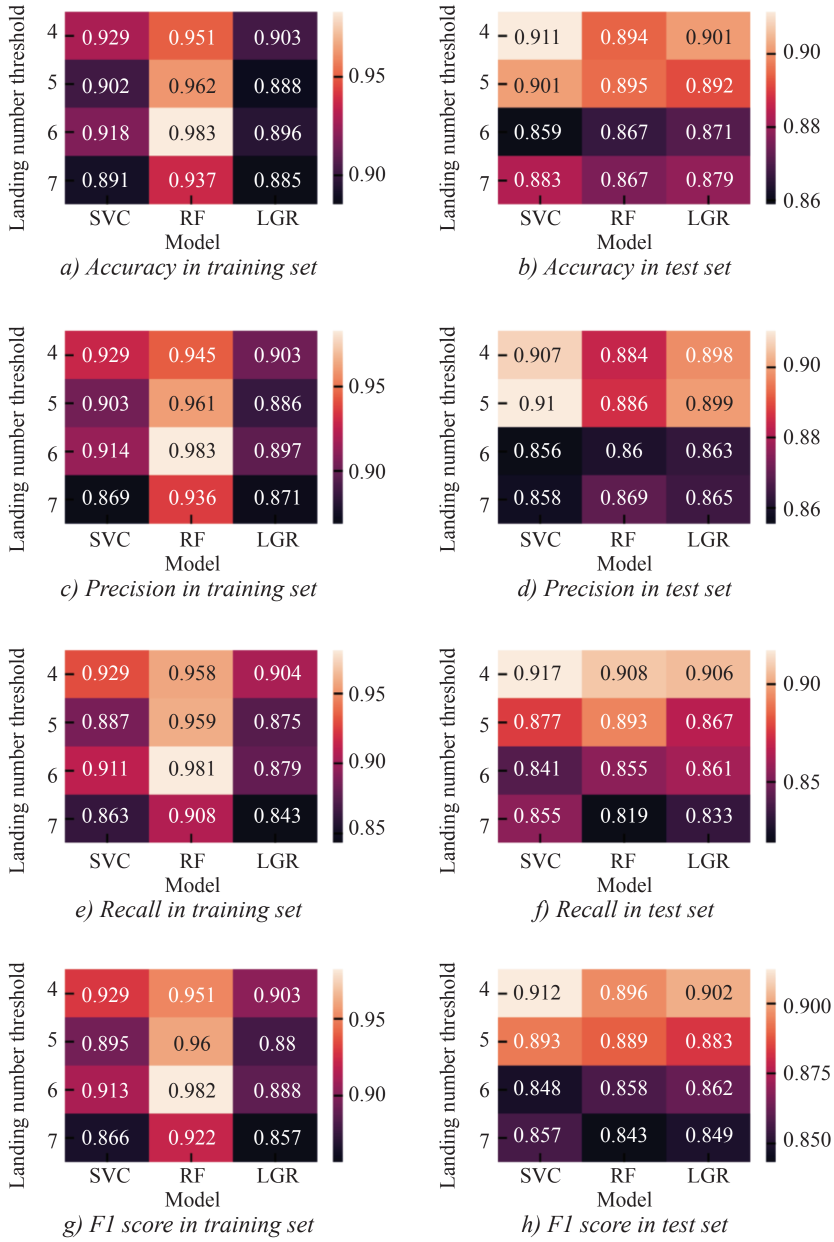


Figure 16 – The results of classification perspective

complexity indicators as inputs and the clustering results of performance indicators as outputs. Three classification methods (SVM, LR, RF) and four classification metrics (precision, recall, accuracy and F_1 score) are used for the correlation evaluation. 80% of the samples are chosen for training, and the remaining 20% for testing. The hyperparameters of the classification models (SVC and RF) are tuned via cross-validation. *Figure 16* presents the classification results based on three different models.

From the classification results on the test set, as shown in *Figure 16*, it can be seen firstly that the four metrics show an overall decreasing trend with the number of landings threshold increasing. Secondly, the four evaluation metrics are higher overall than the evaluation metrics of the clustering results, which indicates a non-linear correlation between the complexity and performance, which the classification models capture. For example, the accuracy can reach a maximum of 0.91 and a minimum of 0.86. In addition, the F1 score can reach a maximum of 0.91 and a minimum of 0.84. Thirdly, comparing the different models, the performance of the three models is similar, with SVC being slightly better.

Finally, the RF model is used to determine the importance of different complexity indicators that pose main effects on the peak and off-peak situations for arrival operation. As a result, *Figure 17* provides the feature importance distinguished by the RF model in which the landing number threshold is five. Here we have selected features with feature importance greater than five percent for display which includes $\bar{D}^{Sum}$, $\bar{C}^{Sum}$, $\bar{I}^{Sum}$, $\bar{n}$, $\bar{n}^2$, $MD_{5\text{–}10}$. As shown in *Figure 17*, the most significant factor is $\bar{D}^{Sum}$ which accounts for 22.02%. These significant factors are all related to the number of aircraft. Furthermore, controllers need to pay more attention to maintaining the separation and establishing the approach sequence.

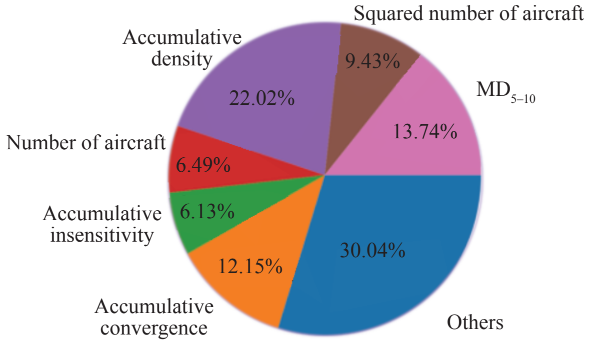


Figure 17 – Indicator importance for distinguishing peak and off-peak situations

5. CONCLUSION

This paper proposed a framework to establish the correlation between process-oriented complexity and result-oriented performance indicators for arrival operation. The performance labels generated by the K-means algorithm were used as an alternative to the manual scoring method for avoiding subjectivity.

1) There is a significant correlation between air traffic complexity and performance on arrival operations. However, delays exist between complexity and performance level from a time-series perspective.
2) Six complexity indicators that significantly impact performance are closely related to the number of aircraft within TMA. In other words, scheduling arrival aircraft better could promote ATC performance, especially in high inbound traffic demand.
3) For future work, we will focus on two areas. One is that we should try better ways to acquire more instructive performance labels. The other is that we could predict the traffic situation based on the correlation between complexity and performance, which could help controllers make better decisions for balancing workload and operational performance.

ACKNOWLEDGEMENT

This work was supported by the Joint Fund of National Natural Science Foundation of China and the Civil Aviation Administration of China (Grant No. U1933117).

张军峰，博士[1], (通讯作者)
邮箱: zhangjunfeng@nuaa.edu.cn
项童, 硕士研究生[1]
邮箱: yarck01@nuaa.edu.cn
周铭, 硕士研究生[1]
邮箱: mmxyh@nuaa.edu.cn
王斌, 高级工程师[2]
邮箱: teddy11@163.com
[1] 南京航空航天大学，民航学院
中国江苏省南京市江宁区将军大道29号, 邮编: 211106
[2] 民用航空中南地区空中交通管理局
中国广东省广州市白云区南云东街3号, 邮编: 510403

构建进场运行复杂度与管制效能的关联

摘要

空中交通复杂度指标在衡量运行效能和管制工作负荷方面发挥着重要作用。然而，目前的研究主要依靠人工打分方法来衡量管制运行效能或工作负荷。本文聚焦于进场运行，提出了一种数据驱动的方法建立复杂度和绩效之间的关联，以期避免现有

研究中人工打分方法的主观性。论文首先梳理了描述空中交通复杂度的*26*个指标和刻画进场运行效能的*2*个指标。其次，引入聚类方法，区分进场运行的高峰与非高峰状况。此外，对聚类结果进行比较分析，初步研究了复杂度和效能之间的相关性。第三，采用分类方法进一步确定两者之间的关联性，同时开展了影响运行效能的因素分析。最后，采用广州白云国际机场的进场轨迹实现案例验证。结果表明，复杂度与管制效能之间有强关联性，分类的准确性和精确度约为*90%*，而且终端空域航空器的数量对进场运行效能有显著影响。

关键词

空中交通；空中交通复杂度；复杂度指标；绩效；关联性

REFERENCE

[1] Antulov-Fantulin B, Juričić B, Radišić T, Çetek C. Determining air traffic complexity–challenges and future development. *Promet – Traffic&Transportation*. 2020;32(4): 475-485. doi: 10.7307/ptt.v32i4.3401.

[2] International Civil Aviation Organization. *Manual on global performance of the air navigation system.* Report Number: Doc 9883, 2007.

[3] Prandini M, Piroddi L, Puechmorel S, Brázdilová SL. Toward air traffic complexity assessment in new generation air traffic management systems. *IEEE Transactions on Intelligent Transportation Systems*. 2011;12(3): 809-818. doi: 10.1109/TITS.2011.2113175.

[4] Laudeman IV, Shelden SG, Branstrom R, Brasil CL. *Dynamic density: An air traffic management metric.* NASA. Report Number: NASA/TM -1998-112226, 1998.

[5] Smith K, Scallen SF, Knecht W, Hancock PA. An index of dynamic density. *Human Factors.* 1998;40(1): 69-78. doi: 10.1518/001872098779480604.

[6] Masalonis AJ, Callaham MB, Wanke CR. Dynamic density and complexity metrics for realtime traffic flow management. *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar*. 2003. p. 139.

[7] Netjasov F. Terminal airspace traffic complexity. *Proceedings of 1st International Conference on Research in Air Transportation ICRAT.* 2004.

[8] Netjasov F, Janic M, Tosic V. The future air transport system: Looking for generic metrics of complexity for terminal airspace. *Proceedings of the 88th TRB annual meeting*; 2009. p. 11-15.

[9] Netjasov F, Janić M, Tošić V. Developing a generic metric of terminal airspace traffic complexity. *Transportmetrica*. 2011;7(5): 369-394. doi: 10.1080/18128602.2010.505590.

[10] Zhang J, Hu M, Zhang C. Complexity research in air traffic management [空中交通管理中的复杂性研究]. *Acta Aeronautica et Astronautica Sinica*. 2009;30(11): 2132-2142.

[11] Delahaye D, Puechmorel S. Air traffic complexity: Towards intrinsic metrics. *3rd USA/Europe Air Traffic Management R&D Seminar Napoli.* June 2000. p. 1–11.

[12] Delahaye D, et al. Air traffic complexity map based on non-linear dynamical systems. *Air Traffic Control Quarterly.* 2004;12(4): 367-388. doi: 10.2514/atcq.12.4.367.

[13] Wang Z, Delahaye D, Farges JL, Alam S. Air traffic assignment for intensive urban air mobility operations. *Journal of Aerospace Information Systems*. 2021;18(11): 860-875. doi: 10.2514/1.I010954.

[14] Wang H, Song Z, Wen R, Zhao Y. Study on evolution characteristics of air traffic situation complexity based on complex network theory. *Aerospace Science and Technology*. 2016;58: 518-528. doi: 10.1016/j.ast.2016.09.016.

[15] Wee HJ, Lye SW, Pinheiro JP. A spatial, temporal complexity metric for tactical air traffic control. *The Journal of Navigation*. 2018;71(5): 1040-1054. doi: 10.1017/S0373463318000255.

[16] Chatterji G, Sridhar B. Measures for air traffic controller workload prediction. *1st AIAA, Aircraft, Technology Integration, and Operations Forum.* 2001. p. 5242.

[17] Gianazza D, Guittet K. Selection and evaluation of air traffic complexity metrics. *2006 IEEE/AIAA25TH Digital Avionics Systems Conference.* 2006. p. 1-12.

[18] Xiao M, Zhang J, Cai K, Cao X. ATCEM: A synthetic model for evaluating air traffic complexity. *Journal of Advanced Transportation*. 2016;50(3): 315-325. doi: 10.1002/atr.1321.

[19] Zhu X, Cao X, Cai K. Measuring air traffic complexity based on small samples. *Chinese Journal of Aeronautics.* 2017;30(4): 1493-1505. doi: 10.1016/j.cja.2017.04.018.

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

[21] Radišić T, Novak D, Juričić B. Reduction of air traffic complexity using trajectory-based operations and validation of novel complexity indicators. *IEEE Transactions on Intelligent Transportation Systems.* 2017;18(11): 3038-3048. doi: 10.1109/TITS.2017.2666087.

[22] Andraši P, Radišić T, Novak D, Juričić B. Subjective air traffic complexity estimation using artificial neural networks. *Promet – Traffic&Transportation*. 2019;31(4): 377-386. doi: 10.7307/ptt.v31i4.3018.

[23] Xie H, Zhang M, Ge J, Dong X, et al. Learning air traffic as images: A deep convolutional neural network for airspace operation complexity evaluation. *Complexity*. 2021. doi: 10.1155/2021/6457246.

[24] Standfuss T, Rosenow J. Applicability of current complexity metrics in ATM performance benchmarking and potential benefits of considering weather conditions. *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*. 2020. p. 1-9.

[25] Condé Rocha Murça M, Hansman RJ. *Data-Driven modeling of air traffic flows for advanced air traffic management.* PhD thesis. Massachusetts Institute of Technology; 2019.

[26] Kopardekar P. *Dynamic density: A review of proposed variables*. FAA WJHTC internal document, Overall conclusions and recommendations. Federal Aviation Administration; 2000.

[27] Liu S, et al. Identifying operational benefits of the arrival management system–A KPI-based experimental method by evaluating radar trajectories. *Promet – Traffic&Transportation*. 2021;33(5):633-645. doi: 10.7307/

ptt.v33i5.3786.

[28] Peng Z, et al. Benefits derived from arrival management and wake turbulence re-categorization in China. *Transportation Research Record*. 2021;2675(11): 373-383. doi: 10.1177/03611981211017898.

[29] Zhang J, et al. Criteria selection and multi-objective optimization of aircraft landing problem. *Journal of Air Transport Management*. 2020;82: 101734. doi: 10.1016/j.jairtraman.2019.101734.

[30] EUROCONTROL. *Unimpeded ASMA Time - Technical Note*. 2021. https://ansperformance.eu/methodology/unimpeded-asma-time/.

[31] Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 2010;31(8): 651-666. doi: 10.1016/j.patrec.2009.09.011.

[32] Chang CC, Lin CJ. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems And Technology (TIST)* 2011;2(3): 1-27. doi: 10.1145/1961189.1961199.

[33] Menard S. *Applied logistic regression analysis*. Sage; 2002.

[34] Schonlau M, Zou RY. The random forest algorithm for statistical learning. *The Stata Journal.* 2020;20(1): 3-29. doi: 10.1177/1536867X20909688.