Prediction of Departure Aircraft Taxi Time Based on Deep Learning

LI Nan[1*], *JIAO Qingyu*[1], *ZHU Xinhua*[2,3], *WANG Shaocong*[1]

1. College of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, P.R. China;
2. College of Economics and Management, Civil Aviation University of China, Tianjin 300300, P. R. China;
3. China Civil Aviation Environment and Sustainable Development Research Center, Tianjin 300300, P. R. China

(Received 20 February 2020; revised 12 April 2020; accepted 15 April 2020)

Abstract: With the continuous increase in the number of flights, the use of airport collaborative decision-making(A-CDM) systems has been more and more widely spread. The accuracy of the taxi time prediction has an important effect on the A-CDM calculation of the departure aircraft's take-off queue and the accurate time for the aircraft block-out. The spatial-temporal-environment deep learning (STEDL) model is presented to improve the prediction accuracy of departure aircraft taxi-out time. The model is composed of time-flow sub-model (airport capacity, number of taxiing aircraft, and different time periods), spatial sub-model (taxiing distance) and environmental sub-model (weather, air traffic control, runway configuration, and aircraft category). The STEDL model is used to predict the taxi time of departure aircraft at Hong Kong Airport and the results show that the STEDL method has a prediction accuracy of 95.4%. The proposed model also greatly reduces the prediction error rate compared with the other machine learning methods.

Key words: air transportation; taxi time; deep learning; surface movement; convolutional neural network (CNN)

CLC number: U8 **Document code**: A **Article ID**: 1005-1120(2020)02-0232-10

0 Introduction

With the development of civil aviation transportation industry, the number of take-off and landing flights in China continues to grow, but the punctuality rate of flights continues to decline with the increase of the number of flights. The decrease in punctuality rate is not only due to the limitation of airspace capacity, but also the impact of airport operations. In busy airports, departure and landing aircraft need to share some taxiways due to the complex layout of the airport, which may cause airports' high load operation for a long time. Some factors like runway configuration, boarding gate assignment, taxiing path planning and taxi time prediction will directly affect the operation efficiency of the airport. However, current airport collaborative decision-making (A-CDM) only uses airport average taxi time as the prediction taxi time of all aircraft in the airport. It neglects the factors such as stands, runway configuration, number of taxiing aircraft and weather, which leads to the low prediction accuracy of the aircraft taxi-out time and take-off time, resulting in flight delays and increased fuel-burn costs. Therefore, the accuracy of aircraft taxi time prediction plays an important role in optimizing flight pushback time and improving the efficiency of departure time sequence. At the same time, it can provide a theoretical reference for airlines to accurately calculate fuel-burn costs and reduce emissions.

The study of aircraft taxi time is based on the historical data of airport operation, using statistics and data mining algorithms to predict and analyze the taxi time of aircraft. Shumsky[1] and others analyzed queuing theory to predict the departure time

*Corresponding author, E-mail address: nanli@cauc.edu.cn.

How to cite this article: LI Nan, JIAO Qingyu, ZHU Xinhua, et al. Prediction of departure aircraft taxi time based on deep learning[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2020, 37(2): 232-241.

http://dx.doi.org/10.16356/j.1005-1120.2020.02.006

of aircraft. Idris et al.[2] analyzed the impact of the number of ground taxiing aircraft and the taxi distance on the taxi time. Floris et al.[3] predicted the key-related features that influence taxi-out time by using the neural network, regression tree, reinforcement method. Clewlow et al.[4] took Zurich Airport as the research object, and analyzed the relationship between the number of taxiing aircraft and the taxi-out time. Ravizza et al.[5-6] presented that the taxi distance should be divided into pushback section, turning section and straight taxiing section, and studied the taxiing angle and speed when the aircraft taxis through these three sections. Lee et al.[7] simulated the pushback time and taxi time of the apron by using the simulation software Linos. And they compared and analyzed the taxi time accuracy of computer simulation and software simulation by using the random forest algorithm in machine learning. Zhang et al.[8-9] analyzed the factors affecting the airport taxi time, used the econometric model to predict the unimpeded taxi time of non-normal distribution samples and calculated the airport taxi efficiency index. Diana[10] compared the aircraft taxi time prediction model with integrated machine learning, ordinary least squares and regular term algorithm in Seattle airport. Yin et al.[11] optimized the flight taxiing path under the prediction of departure taxi time by fragmenting BP neural network, and Yao[12] used long-term memory model and cyclic neural network algorithm to predict the aircraft taxiing path. Guan et al.[13] used least squares regression and queuing theory algorithms to calculate the pushback time and optimize the gate-hold rate, thereby reducing ground emissions and pollution.

From the literatures above, the following problems were discovered:

(1) The above studies rarely involve the impact of weather and runway configuration changes on taxi time prediction, leading to the problem of reduced prediction accuracy of the model in the event of changes in the external environment such as weather and runway configuration. Airport ground operations are susceptible to severe weather, major activities, and other factors, which can cause a sudden drop in ground operation efficiency. However, many studies or simulation models focus on ground operation efficiency under normal operation, and lack of research under abnormal operating conditions.

(2) Traditional machine learning models are not suitable for the model fitting problem with too many feature samples, which may decrease the converge speed of the model and calculation efficiency.

Therefore, according to the above two problems, we propose a spatial-temporal-environment deep learning (STEDL) model that overcomes the drawbacks of the existing machine learning methods. The proposed model includes the actual capacity of airport surface, the number of taxiing aircraft, weather, taxi distance, and other factors, and joints the convolutional neural network and deep neural network model to predict the departure aircraft taxi time. It not only balances the generalization and representation abilities in one model but also improves the convergence speed of the model and the calculation efficiency.

1 Definition of Taxi Time and Variables

Combining the research purpose of this article and the actual operation of Chinese airports, we define the number of taxiing aircraft, airport capacity, airport runway configuration, and air traffic control information.

1.1 Definition of taxi time

The A-CDM's definition of taxi time is as follows:

(1) Taxi-out time ($T_{\text{taxi-out}}$)

The duration between block-out time ($T_{\text{block-out}}$) and take-off time ($T_{\text{departure}}$)

$$T_{\text{taxi-out}} = T_{\text{departure}} - T_{\text{block-out}} \tag{1}$$

(2) Taxi-in time ($T_{\text{taxi-in}}$)

The duration between landing time (T_{arrive}) and block-in time ($T_{\text{block-out}}$)

$$T_{\text{taxi-in}} = T_{\text{arrive}} - T_{\text{block-in}} \tag{2}$$

1.2 Temporal-traffic flow variables

(1) The number of taxiing aircraft

Due to the complex layout of the taxiway and runways in busy airports, the landing and take-off aircraft have to share some taxiways, which induces the conflict and affects the taxi time of aircrafts[14-15]. The number of taxiing aircraft at different time periods reflects the traffic flow which is a key parameter affecting the taxi time. The parameter definition refers to the research of Idris[2]. The taxiing aircraft are divided into landing taxiing aircraft and departure taxiing aircraft.

① The value ($A(i)$) counts the number of landing taxiing aircraft if the landing time of the aircraft j is after the departure aircraft i's block-out time and before the time of aircraft i's departure time, as shown in Eq.(3).

② The value ($D(i)$) counts the number of departure taxiing aircraft if the block-out time of the departure aircraft j is later than the block-out time of the aircraft i and earlier than the departure time of the aircraft i, as shown in Eq.(4).

$$A(i) = \text{count}(j)$$
$$t_{\text{block-out}}(i) < t_{\text{arrive}}(j) < t_{\text{departure}}(i) \tag{3}$$

$$D(i) = \sum_j \text{count}(j)$$
$$t_{\text{block-out}}(i) < t_{\text{block-out}}(j) < t_{\text{departure}}(i) \tag{4}$$

Correlation analysis is conducted out on traffic flow (average) and taxi time (average), and the correlation coefficient is 0.43. At the same time, the correlation between the number of taxiing aircraft and the taxi time is analyzed, and the correlation is 0.62, indicating that although the traffic flow has an impact on the taxi time, the effect is not as large as the number of surface taxiing aircraft. Therefore, the number of taxiing aircraft is used to express the traffic flow.

(2) Airport capacity

Airport capacity is defined as the sum of the number of aircraft take-off and landing within n min before the estimate take-off time ($t^i_{\text{estimate\_depart}}$) of the departure aircraft i, as shown in Eq.(5). The value j_1 counts the number of departure aircraft if take-off time ($t^{j_1^i}_{\text{depart}}$) of the aircraft j_1^i is n min earlier than the estimate take-off time of the aircraft i. The value j_2 counts the number of landing aircraft if the arrival time ($t^{j_2^i}_{\text{arrive}}$) of the aircraft j_2^i is n min earlier than the estimate take-off time of the aircraft i. The n is derived from the average taxi time at the airport.

$$\text{airport capacity}(j) = j_1 + j_2 \tag{5}$$

where

$$j_1 = \sum_{j_1^i} \text{count}(j_1^i) \quad 0 \leqslant t^i_{\text{estimate\_depart}} - t^{j_1^i}_{\text{depart}} \leqslant n \text{ min}$$
$$j_2 = \sum_{j_2^i} \text{count}(j_2^i) \quad 0 \leqslant t^i_{\text{estimate\_depart}} - t^{j_2^i}_{\text{arrive}} \leqslant n \text{ min}$$

(3) Different time periods

Traffic flow varies with time, the taxi time will be different in different time periods. With reference to Clewlow's method[4] of classifying airport operation time periods, combined with the actual operation of Hong Kong Airport, the time period classification is revised. The new revised classification are Ⅰ (0:00—8:00), Ⅱ (8:01—16:00), and Ⅲ (16:01—24:00). The results are shown in Table 1.

Table 1 Operation time classification

Operation time period	Type
0:00—8:00	Ⅰ
8:01—16:00	Ⅱ
16:01—24:00	Ⅲ

1.3 Spatial variables

In the actual operation of the airport, the taxi path is not the shortest path between the stand and the runway entrance. First, the automatic dependent surveillance-broadcast (ADS-B) monitoring data was analyzed to obtain the taxi path of the aircraft from the stand to the take-off runway. Then, statistics were made to take the most frequently used taxi path between each stand and the runway as the mainstream taxi path[16]. Finally, according to the mainstream taxi path, the taxi distance from each stand to the take-off runway was calculated.

1.4 Environmental variables

(1) Weather

According to the definition and classification of severe weather in the civil aviation meteorological forecast specifications and the operation of Hong Kong Airport, thunderstorms, tropical cyclones, typhoons, advection fog, and heavy precipitation are classified as bad weather in this paper. At the same time, the visibility of the airport, the wind direction, and wind speed directly affect the aircraft's operating speed and waiting time outside the runway, and the taxi time. The meteorological report of aerodrome conditions (METAR) every hour by the Air Traffic Management Bureau Meteorological Center is as the data source. The numerical variables are wind direction (WD), wind speed (WS), visibility (VIS), and cloud ceiling (CC). The dummy variables are used to describe the overall weather conditions. If it is severe weather, it is set to 1; otherwise, it is 0.

(2) Air traffic flow control

Air traffic flow control is a type of external environment restriction. If the amount of flights entering or leaving the airspace sector is too large or the air route cannot meet the required flow due to weather conditions, the air traffic control management center will release the flow control information. The type of flow control information release is mainly the airspace flow control. If the flow control information is released, the aircraft flying in this airspace will be delayed. The dummy variable is used to describe the flow control information. If the aircraft is affected by flow control, it is set to 1; otherwise, it is 0.

(3) Runway configuration

The runway configuration is mainly determined by the wind direction and air traffic flow of the airport at the current time. The change of the runway configuration includes the change of runway operation direction and multi-runway combination take-off and landing modes. Different runway configurations may result in different taxi time. Runway configuration is described in the form of "A1, A2 | D1, D2", where A1 and A2 are landing runways, and D1 and D2 are take-off runways. For example, "07R | 07L" indicates that the runway 07R is used as the landing runway and 07L runway is used as the take-off runway during this period. In actual operation, three kinds of runway configuration are used. The dummy variable is used to describe the runway configuration. If the corresponding runway configuration is used during takeoff, it is set to 1; otherwise, it is 0.

(4) Aircraft category

Although the taxi distance of some flights is the same, the taxiing speed and holding short of runway time may be different due to the influence of aircraft departure wake[17-18]. The Federal Aviation Administration (FAA)'s NextGen plan proposes a new wake reclassification standard to increase the runway capacity and reduce the minimum take-off interval under the premise of ensuring safe operation. The aircraft are reclassified into six categories: A, B, C, D, E, and F. The reclassification is called RECAT. According to the standard, the Hong Kong Civil Aviation Department conducted operational verification at the Hong Kong Airport[19-20]. The average taxi-out time by aircraft category is summarized in Table 2. In the same taxiing distance and time period, the average taxi time of E190 is 5 min longer than that of A330. The dummy variables are described the aircraft categories. If the aircraft type belongs to the corresponding category, it will be 1; otherwise, it will be 0.

Table 2 Taxi-out time by aircraft category

Aircraft category	Average taxi-out time/min
A	22
B	20.1
C	19
D	21
E	
F	26

2 Spatial-Temporal-Environment Deep Learning Model

STEDL model is divided into three sub-models, which are composed of a spatial-temporal model and environment model based on two convolutional neural networks, and a fully-connected spatial model. Based on these three sub-models, the predicted weights are obtained, and the prediction results of the three models are fused to obtain the predicted value of the final departure aircraft taxi time.

2.1 Sub-model design

(1) Time-flow sub-model

The input variables are composed of the number of taxiing aircraft ($A(i)$, $D(i)$), airport capacity, and traffic flow of different time (t). The main structure of the model is convolutional neural network (CNN). The model framework is as follows: $X^t(j)=(A(i), D(i), \text{airport capacity}(j), t)$. The model definition is $f(X^t(j))\rightarrow j^{n\times 4}$.

(2) Spatial sub-model

The input variable is the taxiing distance Dis(j). The model structure is fully-connection. The model framework is as follows: $f(\text{Dis}(j))\rightarrow j^{n\times 1}$.

(3) Environmental sub-model

The input variable are composed of airport weather, runway configuration, aircraft category, and air traffic control. The main structure of the model is CNN. The model framework is as follows: $X^{\text{env}}(j)=(\text{WD}, \text{WS}, \text{VIS}, \text{CC}, \text{ATC}, \text{CON}, \text{runway}, \text{aircraft})$, where CON is the convective weather, ATC the air traffic control, runway the runway configuration, and aircraft the aircraft category.

2.2 Model framework

As shown in Fig.1, the STEDL model is composed of three sub-models: Time-flow variable model and environmental model based on CNN, and spatial model based on fully connected (FC) layer. As shown in Eq.(6), the output values of the sub-model (p_t, p_e, p_d) are fused with different weights to obtain the prediction value of departure aircraft taxi time (p_i).

$$p_i = w_t p_t + w_e p_e + w_d p_d \tag{6}$$

Since the model still solves the regression problem, the mean square error (MSE) is taken as the loss function of the model.

$$\text{MSE} = \min \frac{1}{m}\sum_{i=1}^{m}(\hat{p}_i - p_i)^2 \tag{7}$$

In STEDL model, its CNN sub-model consists of two convolutional layers, a pooling layer, and a fully connected layer. The pooling layer uses maxpolling. The use of Maxpolling makes the model shield the unimportant parameters while maintaining the data characteristics, and solves the problem of excessive model data redundancy. The sub-sampling window value of the pooling layer is set to 2 to reduce the original data length to half of the origin. In terms of convolution layer settings, the research data in this paper is discrete and is not sensitive to periodic changes in time. Therefore, the horizontal sliding value and vertical sliding value of the two convolutional layers of the sub-model are set to 1, and when performing the convolution operation, padding of all 0 s of same convolution type is used. In terms of activation function, we use the Relu function (Eq.(8)) as the activation function of the sub-model, which can avoid the problem of gradient explosion and disappearance of the model.

$$f(x) = \max(0, x) \tag{8}$$

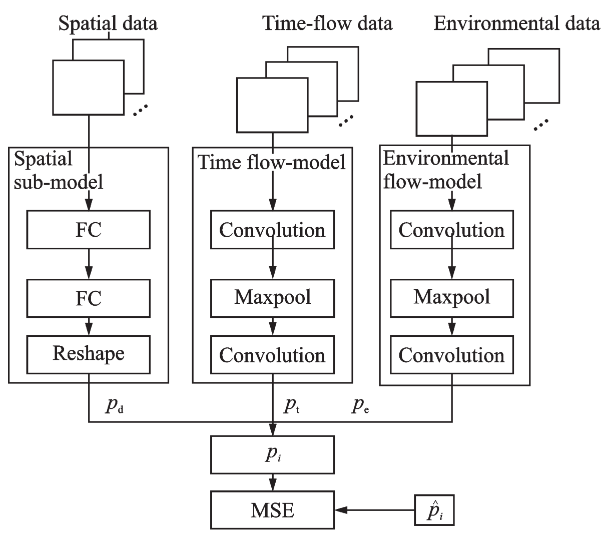

Fig.1 Model of STEDL

The three sub-models analyze spatial-temporal characteristics, external environmental characteristics, and spatial characteristics of the model, which can show the effects of spatial-temporal correlation, environmental differences, and spatial changes in the taxi time of the departure aircraft.

3 Experiment

3.1 Datasets

This data was based on observations at Hong Kong Airport from 1 July, 2019 to 10 February, 2020. The period of interest was from 0:00 to

24:00. The data information consisted of flight number, aircraft type, parking stand, block-in time, block-out time, runway configuration, take-off time, and landing time. Among them, the data affected by extreme conditions such as strong convection weather and typhoons accounted for 5% of the total data. The specific information including departure aircraft is shown in Table 3.

Table 3 Data of departure aircraft

Flight number	Aircraft type	Block-out time	Departure time	Runway	Stand
CX530	A333	8:52	9:16	25R	7
CA101	A321	9:00	9:12	25R	501

After data cleaning, a total of 77 360 valid data were obtained. Table 4 is a statistical analysis of taxi time at Hong Kong Airport. According to the data in the table, the average, the minimum, and the maximum values of the taxi- in time at Hong Kong Airport are less than the taxi-out time. Compared with the taxi-out time distribution, the difference in taxi-in time distribution at Hong Kong Airport is smaller.

Table 4 Aircraft taxi time analysis

Parameter	Min	Max	Avg	Stdev	25th	75th
Taxi-in time/min	2	25	8	3.0	6	10
Taxi-out time/min	4	67	20	6.7	16	24

Figs. 2 and 3 show the taxi-out time frequency distribution of the aircraft at Hong Kong Airport and the analysis of the residual error. As shown in the figures, the taxi-out time at Hong Kong Airport presents a right skewed distribution with skewness and kurtosis of 1.24, 4.208. The skewness and the kurtosis of the taxi-in aircraft are 1.3 and 3.57, respectively. At the same time, the 25th percentile and the 75th percentile in Table 4 indicate that the taxi-out time of Hong Kong Airport is concentrated at 16—24 min and the taxi-in time is concentrated at 6—10 min. Taxiing efficiency of arrival aircraft is higher than that of departure aircraft.

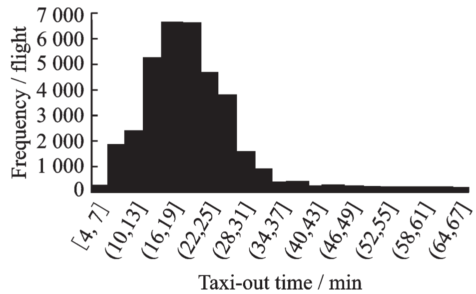

Fig.2 Frequency chart of departure taxi-out time

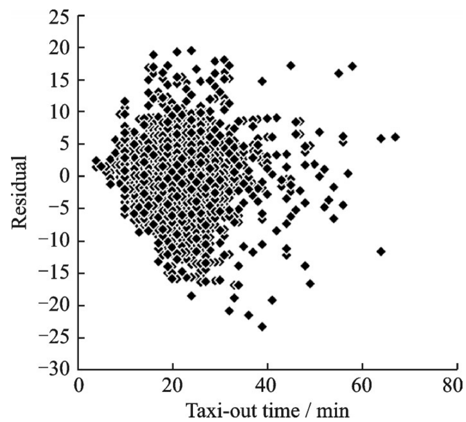

Fig.3 Residual analysis plot

Tables 5 and 6 show the statistical analysis information of numerical variables and the dummy variables in Hong Kong Airport's STEDL model, respectively.

Table 5 Numerical variable dataset

Variable	Hong Kong Airport (HKG)			
	Mean	Stdev	Min	Max
Taxiing distance/m	2 413.9	1 151.2	481	5 600
Amount of landing taxiing aircraft	8.42	4.41	0	37
Amount of take-off taxiing aircraft	8.95	4.45	0	42
Airport capacity per 15 min/flight	7.12	4.52	0	30
Wind direction	137.5	8.87	50	220
Wind speed/($m \cdot s^{-1}$)	6.88	0.34	3.1	12.3
Visibility /m	9 534.54	198.27	4 700	9 999
Ceiling/m	488.12	16.89	360	750

Table 6 Dummy variable dataset

Variable	Type	Percentage/%
Runway configuration	25L | 07R	52.7
	07L | 25R	38.08
	25R | 25R	8.5
Air traffic control	Effect	20.7
	None	79.3
Aircraft type	A	1
	B	58.15
	C	1.12
	D	38.3
	E	0.3
	F	0.9
Weather	Normal	95
	Typhoon	2.4
	Thunderstorm	2.6
Traffic flow of different time	Ⅰ	19.5
	Ⅱ	57.3
	Ⅲ	23.2

3.2 Prediction results

STEDL model was implemented in Python and run by using the TensorFlow framework. The model used Adam as the optimization parameter, the activation function was the Relu function, the number of training iteration was 1 000, and the learning rate was 0.01.

(1) Competing methods

To assess the overall model fit, we assessed three indices, including R-square (R^2), mean square error (MSE), and the mean absolute error (MAE). The predicted results are shown in Tables 7, 8.

According to Table 7, compared with the other three algorithms, the STEDL model has the smallest prediction error, and MAE and MSE are only 0.26 and 0.135. The median absolute error value of the STEDL model is 0.09 lower than that of random forest (RF) and 0.13 lower than that of support vector machine (SVM). According to the MSE value, the STEDL model is 0.165 lower than the RF model, and 0.195 lower than the SVM model. Compared with the traditional machine learning model, the STEDL model has higher prediction accuracy and smaller prediction error. At the same time, comparing the STEDL model with the traditional statistical multiple regression model, the R^2 value of the STEDL model is improved 0.202, and the model fit is better than that of the traditional statistical model. From Table 8 and Fig.4, it can be seen that the probability of the error between the predicted value and the actual value of the STEDL model within 1 min is 52%, the probability of RF is 50.2%, and the probability of SVM is 47.3%. The probability of the error between the algorithm and the actual value within 1 min is similar, but the probability of the STEDL model within 3 min and the probability within 5 min is higher than that of RF and SVM, indicating that the STEDL model has higher prediction accuracy.

Table 7 Evaluation of prediction accuracy

Method	R^2		MAE (normalization)		MSE (normalization)	
	Train	Test	Train	Test	Train	Test
SVM	0.80	0.77	0.38	0.39	0.32	0.33
RF	0.80	0.79	0.30	0.35	0.28	0.30
Multi regression	0.70	0.69	0.37	0.37	0.31	0.39
STEDL	0.902	0.87	0.23	0.26	0.10	0.135

Table 8 Prediction accuracy %

Model	±1 min	±3 min	±5 min
STEDL	52	79.5	95.4
SVM	47.3	75.4	87.4
Multi regression	40.9	69.8	85
RF	50.2	71.5	90.0

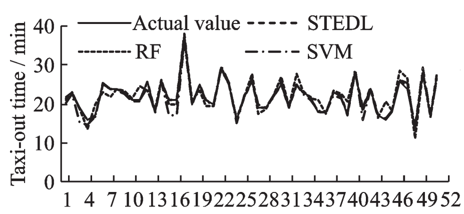

Fig.4 Plot of prediction vs. actual value

(2) Sub-model importance analysis

In order to study the impact of the three sub-models on taxi time, three sub-models are used to predict the taxi time, as shown in Table 9.

According to MSE, the time-flow sub-model error is only 0.27, the external environmental sub-model error is 0.71, and the spatial sub-model error is the largest, 0.73. The R^2 value of time-flow vari-

Table 9 Variable importance analysis

Sub-model	MSE(normalization)	R^2
Time-flow sub-model	0.27	0.765
Spatial sub-model	0.73	0.25
Environmental sub-model	0.71	0.27

able, external environmental variable, and spatial variable is in descending order. It is proved that the time-flow variable has the greatest effect on the taxi time of the aircraft. The degree of traffic congestion at the airport is positively correlated with the taxi time of the aircraft. Although the severe weather in external environmental variables also has a strong impact on the taxi time of the aircraft, due to the fewer days it takes, the number of affected aircraft accounts for less of the total number of aircraft, resulting in its importance to decline.

(3) Comparison of prediction performance in different weather conditions

In order to investigate the prediction effect of the STEDL model on the taxi-out time in different weather conditions, the test samples are classified according to weather types. Thunderstorms, heavy precipitation, fog (with visibility less than 1 km), and typhoons (including three days before and after transit) are classified as severe weather, and the rest are classified as normal weather. The three days of 19 April, 31 July, and 1 August, 2019 are as severe weather samples. The day of 19 April, 2019 is a severe thunderstorm and strong precipitation, and the days of 31 July and 1 August, 2019 are typhoon weather. The comparison of the taxi time in different weather is shown in Table 10.

Table 10 Comparison of aircraft taxi time in different weathers

Weather condition	Taxi-in time/min			Taxi-out time/min		
	Avg	Med	Mode	Avg	Med	Mode
Normal	7	7	6	20	20	18
Severe	10	7	5	34	31	28

From Table 10, the taxi time in extreme weather conditions is longer than that of normal weather. It indicates that the taxiing aircraft is affected by severe weather conditions. As shown in Table 11, the prediction error of the STEDL model increases in severe weather, and its accuracy of 1 min, 3 min, and 5 min is reduced by about 20% compared with that in normal weather. It shows that the prediction of aircraft taxi time in severe weather is complicated. At the same time, based on the analysis of the taxi time characteristics of the landing and take-off aircraft in severe weather, it was found that the taxi time in different severe weather is also quite different (Fig. 5). The average taxi time of a take-off aircraft in a thunderstorm weather is 3 min longer than in a tropical storm weather system. Judging from the impact time of severe weather, it can be concluded that two hours after the end of the severe thunderstorm weather system, the traffic congestion reaches its peak, and the taxi time also reaches the maximum accordingly. Due to the strong predictability of tropical storm systems, when the typhoon transits, the aircraft taxi time reaches the maximum, and the next day after transit, there is no scene of traffic congestion, and the aircraft taxi time is gradually decreasing.

Table 11 Accuracy of STEDL prediction in different weathers %

Accuracy	±1 min	±3 min	±5 min
Normal weather	51	79.5	95.8
Severe weather	32.1	54.5	70

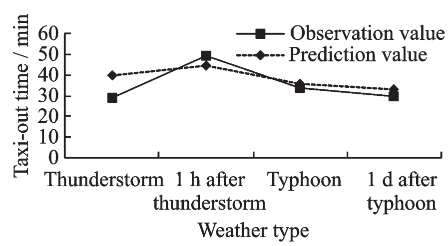

Fig.5 Different weather forecast results

4 Conclusions

To improve the accurate calculation of flight departure and delay time, a deep learning model (STEDL) based on time-space-environment data is proposed to predict the taxi time of the departure aircraft. Some conclusions can be drawn as follows.

(1) The STEDL model can effectively reflect the impact of airport surface space attributes, environmental changes and surface traffic flow changes

on aircraft taxi time, but the most important factor affecting aircraft taxi time is still the change in traffic flow.

(2) The accuracy of the STEDL model for taxi time prediction of departure aircraft is 95.4%. Its model-fitting prediction capability is higher than other machine learning algorithms such as SVM and RF, and it can be used to predict the actual taxi time of large airports.

(3) Severe weather such as strong thunderstorms and typhoons have a great impact on the taxi time of the aircraft and exhibit lagging and continuity characteristics.

(4) Due to the limitation of data acquisition, only the taxi time of Hong Kong Airport is analyzed. It is planned to add other large airports to the proposed model for prediction and comparison in the future to improve the universality of the model.

(5) Comparing the prediction results of the STEDL with the research results of other scholars. It is found that the fitting value ($R^2 = 0.90$) after using the STEDL model is higher than using the traditional machine learning model[10] ($R^2 = 0.70$), also higher than using the econometric regression method[9] ($R^2 = 0.74$). It shows that the use of STEDL algorithm is more suitable for the prediction of aircraft taxi time.

References

[1] SHUMSKY R A. Real-time forecasts of aircraft departure queues[J]. Air Traffic Control Quart, 1997, 5(4): 281-308.

[2] IDRIS H, CLARKE J P, BHUVA R, et al. Queuing model for taxi-out time estimation[J]. Air Traffic Control Quarterly, 2007, 10(1): 1-22.

[3] FLORIS H, RICHARD C, HENDRIKUS V, et al. Taxi-out time prediction model at Charles de Gaulle Airport[J]. Journal of Aerospace Information Systems, 2018, 15(3): 1-11.

[4] CLEWLOW R L, SIMAIAKIS I, BALAKRISHNAN H. Impact of arrivals on departure taxi operations at airports[C]//Proceedings of AIAA Guidance, Navigation, and Control Conference. Toronto, Ontario Canada: AIAA, 2010: 1-21.

[5] RAVIZZA S, ATKIN J A D, MAATHUIS M H, et al. A combined statistical approach and ground movement model for improving taxi time estimations at airports[J]. Journal of the Operational Research Society, 2013, 64(9): 1347-1360.

[6] CHEN J, RAVIZZA S, ATKIN J A D, et al. Aircraft taxi time prediction: Comparisons and insights[J]. Applied Soft Computing Journal, 2014, 14(1): 397-406.

[7] LEE H, MALIK W A, JUNG Y C. Taxi-out time prediction for departures at Charlotte Airport using machine learning techniques[C]//Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference. [S.l.]: AIAA, 2016.

[8] ZHANG Y. Methods for determining airport unimpeded taxi times[C]//Proceedings of Transportation Research Board 90th Annual Meeting. Washington DC: Transportation Research Board, 2011: 1-19.

[9] ZHANG Y, WANG Q. Methods for determining unimpeded aircraft taxi time and evaluating airport taxiing performance[J]. Chinese Journal of Aeronautics, 2017, 30(2): 523-537.

[10] DIANA T. Can machines learn how to forecast taxi-out time? A comparison of predictive models applied to the case of Seattle/Tacoma International Airport[J]. Transportation Research Part E-Logistics and Transportation Review, 2018,119: 149-164.

[11] LIU J, YIN M. Study on the influencing factors of departure aircraft taxi time[J]. Journal of Wuhan University of Technology, 2018, 42(2): 195-200.

[12] YAO M F. Research on key techniques of aircraft surface trajectory prediction and path planning in airport[D]. Chengdu: University of Electronic Science and Technology of China, 2018. (in Chinese)

[13] LIANG G, ZHANG Y P, XING Z W, et al. A new dynamic pushback control method for reducing fuel-burn costs: Using predicted taxi-out time[J]. Chinese Journal of Aeronautics, 2019, 32(3): 660-673.

[14] YIN J N, HU M H, MA Y, et al. Airport taxi situation awareness with a macroscopic distribution network analysis[J]. Networks and Spatial Economics, 2019,19: 669-695.

[15] XIA Z, ZHENG B, WAN J, et al. Recognition algorithm and risk assessment of airport hotspots[J]. Journal of Shanghai Jiaotong University (Science), 2019, 24(10): 769-774.

[16] LI N, LIU P, JING H H. Research on aircraft speed anomaly detection in maneuvering area[J]. Computer Simulation, 2019, 36(1): 45-50.

[17] Kim S H, FERON E, CLARKE J P, et al. Airport gate scheduling for passengers, aircraft, and operation[J]. Journal of Air Transportation, 2017, 25(4):

109-114.

[18] LEE H, COUPE J, JUNG Y C. Prediction of push-back times and ramp taxi times for departures at Charlotte Airport[C]//AIAA Aviation 2019 Forum. Dallas: AIAA, 2019(2933): 1-13.

[19] LIU T C. Hong Kong aviation safety program[R]. [S. l.]: Hong Kong Aviation Department, 2018: 3-4.

[20] LIU T C. Controlling officer's reply[R]. [S. l.]: Hong Kong Aviation Department, 2019: 2-11.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Nos.U1833103, 71801215); the China Civil Aviation Environment and Sustainable Development Research Center Open Fund(No.CESCA2019Y04).

Author Ms. **LI Nan** received the B.S. degree in computer science and application, Department of Computer Science & Technology from Civil Aviation University of China, Tianjin, China in 2000 and M.S. degree in navigation guidance and control, College of Air Traffic Management, Civil Aviation University of China in 2003. Since 2003, she has been working in College of Air Traffic Management, Civil Aviation University of China. Her research is focused on transportation planning and management, and air traffic control simulation verification.

Author contributions Ms. **LI Nan** designed the study, provided the cases and idea, conducted the analysis. Mr. **JIAO Qingyu** complied the models and wrote the manuscript. Dr. **ZHU Xinhua** contributed to the discussion and background of the study. Mr. **WANG Shaocong** contributed to the data collection and data analysis. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: XU Chengting)

基于深度学习的离场航空器滑行时间预测

李 楠[1]，焦庆宇[1]，朱新华[2,3]，王少聪[1]

（1.中国民航大学空中交通管理学院，天津300300，中国；2.中国民航大学经济与管理学院，天津300300，中国；
3.中国民航环境与可持续发展研究中心，天津300300，中国）

摘要：随着航班数量的不断增加，机场协同决策系统（Airport collaborative decision-making，A-CDM）的使用也越来越广泛。滑行时间预测的准确性对A-CDM计算离场航空器起飞排序队列和给出准确的撤轮挡时间具有重要的作用。本文提出一种基于时间-空间-环境数据的深度学习模型（Spatio-temporal-environment deep learning model，STEDL）来提高滑行时间预测的准确性。该模型由时间-流量变量（机场实际容量，场面航空器数量，时间段）、空间变量（滑行距离）、外部环境变量（天气，流控信息，跑道运行模式，机型）3部分组成。使用STEDL模型对香港机场离场航空器滑行时间进行预测验证。实验结果显示，STEDL模型预测准确率为95.4%，预测精度明显优于其他机器学习算法。

关键词：航空运输；滑行时间；深度学习；场面运行；卷积神经网络