



## Short-term prediction of airway congestion index using machine learning methods

Zhao Yang <sup>a,b,\*</sup>, Rong Tang <sup>b</sup>, Weili Zeng <sup>b</sup>, Jiahuan Lu <sup>b</sup>, Zhijie Zhang <sup>b</sup>

<sup>a</sup> College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Jiangjun Road No. 29, Nanjing 211106, China

<sup>b</sup> National Key Laboratory of Air Traffic Flow Management, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Jiangjun Road No. 29, Nanjing 211106, China

### ARTICLE INFO

**Keywords:**

Short-term prediction  
Airway congestion  
Long short-term memory network  
Graph convolutional network  
Extreme gradient boosting

### ABSTRACT

This study presents a hybrid spatiotemporal convolutional long short-term memory network with tri-directional temporal features (SCLN-TTF) for airway congestion index prediction. The spatial features, temporal features and spatiotemporal features are taken as input variables, among which the temporal and spatiotemporal input datasets are categorized into forward time series datasets, backward time series datasets and datasets at the target time. The spatial features are extracted using the fast-approximate convolutions on graphs, and the temporal features are extracted using the long short-term memory network. The subsequent congestion index is the output variable. The model parameters are sequentially updated based on the recent collected data and the new predicting results. It is found that, in general, the proposed SCLN-TTF method achieves the most stable and satisfactory performance for all the selected airways. The models exhibit the best performance for the 60 min prediction time horizon in terms of MAPE, while the proposed SCLN-TTF shows superiority compared to the other benchmark methods particularly for the 10 min prediction. Then the graph convolutional network and extreme gradient boosting method are combined for high congestion index identification. The finding further confirms the superiority of the proposed method by considering both the spatial and temporal dependencies and integrating forward and backward features.

### 1. Introduction

With the rapid growth of air traffic demand, airspace congestion has become increasingly prominent, and the workload of air traffic controllers has increased dramatically (Bilimoria and Lee, 2005; Idris et al., 2020; Vaaben and Larsen, 2015). In order to facilitate air traffic control, the airspace is usually divided into multiple control units. Each unit is called a sector, which is a network composed of multiple airways. Accurate prediction of airway congestion is important for air traffic control, which can provide technical support and reference basis for easing airway congestion, reducing air traffic control load, and thus improving air traffic safety and operational efficiency.

Airspace congestion refers to air traffic phenomena such as aircraft queuing, waiting in the air, maneuvering around, etc., due to the inability of the airspace operating capacity to meet the continuous flight demand. Previously, a few studies focused on the problem of alleviating the airspace congestion. For example, Cai et al. formulated a multi-objective air traffic network flow optimization

\* Corresponding author.

E-mail address: [seu\\_yolanda@126.com](mailto:seu_yolanda@126.com) (Z. Yang).

(MATNFO) problem considering comprehensive air traffic flow management actions such as ground-holding, airborne-holding, rerouting and speed control. A systematic approach, namely route and time-slot assignment algorithm, is developed to solve the MATNFO problem (Cai et al., 2017). Crucilo et al. modeled the air holding problem with reinforcement learning and proposed a solution to reduce airspace congestion. The air holding problem module was applied in two case studies of the Brazilian airspace (Cruciol et al., 2015). Dai et al. presented a propagation dynamics model for congestion propagation in complex networks of airspace. An epidemiology model was applied to complex networks by comparing the similarities and differences between congestion propagation and epidemic transmission (Dai et al., 2016). Dong et al. proposed a method of designing corridor, identifying congestion state and analyzing the influence of air routes' traffic flow. The method was used to identify congestion periods and congestion segments of China's corridors, to reveal the congestion regularities on temporal dimension, spatial dimension and spatiotemporal interaction dimension, and to evaluate the congestion influence of air routes' traffic flow (Dong et al., 2019).

Although continuous efforts are being conducted towards the objective of airspace congestion alleviation, the recognition and prediction of airway congestion remains to be a tough issue. On one hand, there has no universally accepted index illustrating the airway congestion status. Currently, the most widely used method is to simply judge the congestion state of an airway by comparing the actual flow with the operating capacity threshold. It cannot demonstrate the temporal and spatial characteristics of the airway congestion, which affects to a greater extent, its pertinence, scientificity and practicality. On the other hand, the short-term prediction of airway congestion tends to be complicated, due to the stochasticity and dynamic nature of air traffic flow considering the various influencing factors such as weather conditions (Goswami and Sarkar, 2017). Further development is still needed to advance the predictive aspects of the linkage between airway congestion status and the input variables including meteorological variables, and then to predict future status.

With the advance of machine learning techniques, some efficient tools can be applied in the field of airway congestion analysis. As a matter of fact, the airway congestion evaluation indicators can be regarded as time series data. The future congestion status can be forecasted based on historical data, together with influencing factors. In recent years, machine learning models have been successfully applied to broad areas of time series forecasting problems, such as traffic flow prediction (Han et al., 2019), delay prediction (Chen and Li, 2019; Takeichi et al., 2017), etc. Theoretically, these commonly used techniques also have the potential to be applied in airway congestion analysis, as they can provide fast and real-time prediction results through high-efficient learning and training process and achieve good performance by integrating various influencing features.

To this end, the primary objective of this study is to predict the short-term airspace congestion status and identify the potential congested time periods or airspace segments incorporating multi-source data. The workflow is illustrated in Fig. 1. First, the topological structure of the neural network is constructed based on the airway network. Specifically, the waypoint of an airway network is represented as an edge and the airway is represented as a node in the neural network. Then two types of models are constructed, including regression models for congestion index prediction and classification models for high congestion index identification. As for the regression models, the graph convolution and LSTM networks are integrated for spatiotemporal feature extraction and congestion

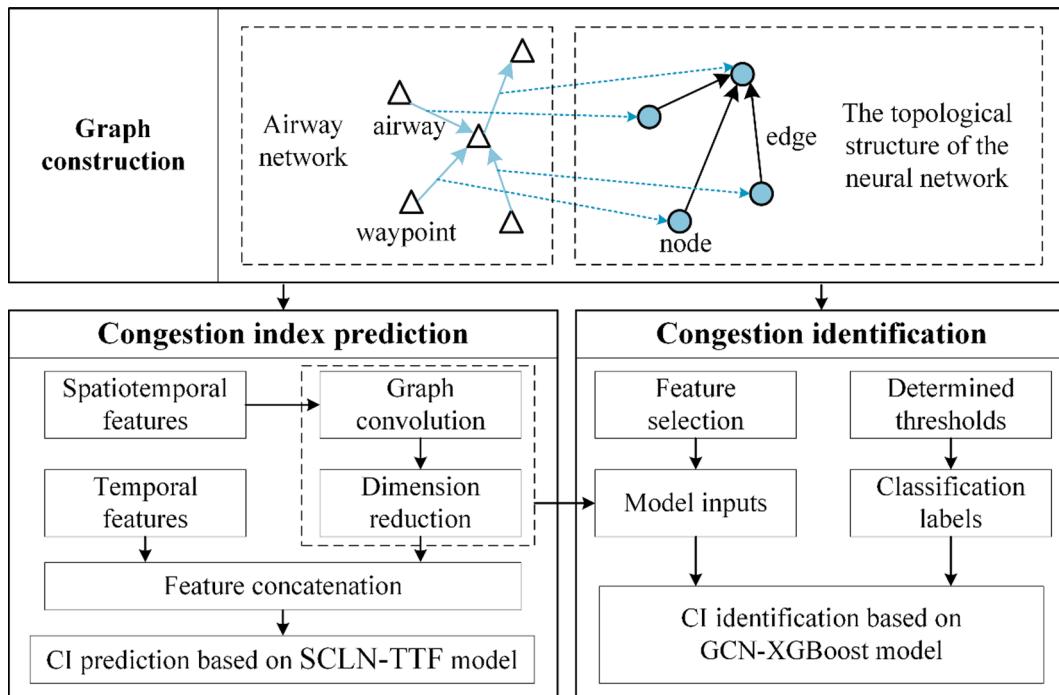


Fig. 1. Procedure of the present work.

index prediction. The proposed model is denoted as a hybrid spatiotemporal convolutional long short-term memory network with tri-directional temporal features (SCLN-TTF). As for the classification models, the XGBoost method is further incorporated for airspace congestion identification. The proposed model is denoted as combined graph convolutional network-extreme gradient boosting (GCN-XGBoost) model. The spatial and temporal patterns of a real-world traffic flow data are analyzed, and the performance gains of the proposed model relative to benchmark methods are illustrated.

The major contributions of this paper can be summarized in three parts. First, a spatiotemporal neural network model is proposed, which explicitly consider various types of short-term, long-term and spatial patterns. Both the spatial and temporal dependencies are explored, thus the impacts of congestion from adjacent time periods or congestion propagation between adjacent airways can be taken into account. Second, the proposed airway congestion index prediction method incorporates both the forward sequence and backward sequence as inputs, including the historical, current and forecasting features. Therefore, by integrating all these explanatory variables in the prediction architecture, impacts of the various influencing factors can be better reflected. Third, the proposed high congestion index identification method incorporates the strength of LSTM in time-series data analysis and XGBoost method in feature extraction. Models with different input features are compared. The results further confirm the superiority of the proposed method by taking both the time-series information as well as the various features into account.

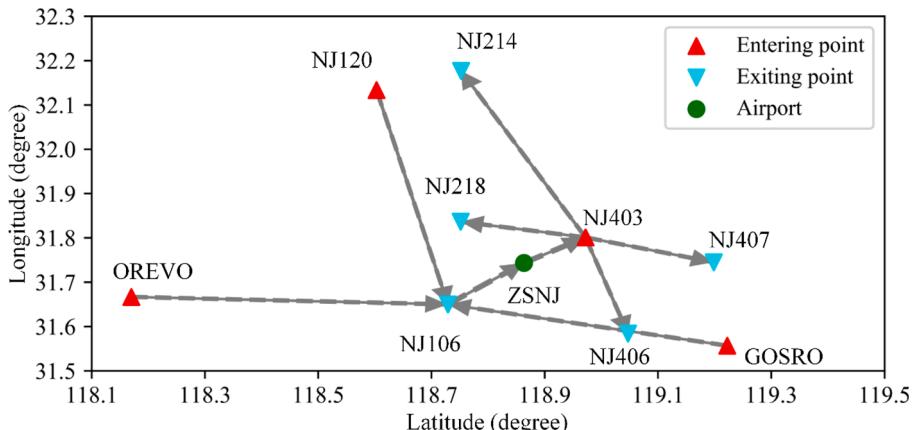
The rest of the paper is organized as follows. In [Section 2](#), the airway congestion index is defined, and the spatiotemporal forecasting problem is formulated. In [Section 3](#), the data collection and preparation process are presented. In [Section 4](#), the technical background and the proposed framework are presented. The data analysis results are provided in [Section 5](#). The conclusions and future works are summarized in [Section 6](#).

## 2. Problem definition

Previous studies indicate that a series of indicators have been used illustrating airspace congestion. For example, the degree of airspace congestion can be estimated based on the ratio of predicted air traffic demand to airspace capacity over a certain period of time. Indicators such as instantaneous peak flow rate, ratio of demand to capacity, probability of overcapacity, overcapacity time duration can be calculated ([Beaton et al., 2001](#); [Cai et al., 2017](#); [Gilbo and Smith, 2011](#)). To further improve the prediction accuracy, the probability distributions of corresponding indicators are investigated, which may help to identify the impact of various uncertain factors based on the concept of probability transfer ([Mulgund et al., 2006](#); [Taylor and Wanke, 2009](#)). There are also some studies that use other indicators to demonstrate the operating conditions of air traffic flow, i.e. the average flight delay time, arrival rate within a period of time, controller workload, etc. ([Adacher et al., 2017](#); [Dong et al., 2019](#); [Enayatollahi and Atashgah, 2018](#)). These indicators reflect the airspace congestion status from different perspectives, while some of which are not easy to quantify. As a matter of fact, analysis of the ADS-B data indicates that with the same flow rate for a specific airway, the average travel time may be diverse for different time periods. Theoretically, it should be regarded as different congestion states. The present research attempts to incorporate the information from two widely used indicators, flight flow and flight time. Therefore, this study uses the total flight time of aircraft on the route in a specified time interval as an indicator to reflect the congestion of the route. Considering the different distances of each route, to facilitate the comparison of the degree of congestion among different routes, the total flight time of aircraft is divided by the length of each segment. Thus, the airway congestion evaluation index is defined as the cumulative flight time of the flights passing the airway divided by the length of the airway, which means the cumulative occupation time of a unit of route length. The defined congestion index for route  $r$  in period  $i$  can be calculated using the following formula:

$$C_{r,i} = \frac{1}{L_r} \sum_{m=1}^M \text{time}_{r,i,m} \quad (1)$$

In the formula,  $L_r$  represents the length of route  $r$  (km);  $\text{time}_{r,i,m}$  represents the flight time of aircraft  $m$  in air route  $r$  in period  $i$  (sec);



**Fig. 2.** Layout of directed airways and waypoints of ZSNJAP01.

and  $M$  represents the total number of flights in air route  $r$  in period  $i$ . The unit of  $C_{r,i}$  is sec/km.

In the following sections, the defined congestion index (CI) is taken as the indicator to be predicted, based on the historic data, as well as the given features.

### 3. Data preparation

#### 3.1. Data collection

To meet the research objective, the flight data of approach sector 01 of Nanjing Lukou International Airport (ZSNJAP01) are collected. The study period is from January 1 to December 31, 2018. The layout of the waypoints and the airways of ZSNJAP01 are shown in Fig. 2. The red points represent the entering points and the blues points represent the exiting points around the ZSNJ airport. The directed edges represent the airways.

The data sources for analysis can be divided into two categories: flight data and meteorological information. The flight ADS-B data of ZSNJAP01 in 2018 are collected, with records of information for a total of 214,334 flights. For each record, the specific flight information extracted from the dataset includes flight ID, aircraft type, departure airport, destination airport, sector code, enter point, historic arrival time at each waypoint, and estimated arrival time at each waypoint. The total number of flights in the airways during a specified time interval is calculated according to the timestamp that each aircraft entering or exiting the waypoint. Also, the historic airway congestion index is calculated using Eq. (1).

The airport meteorological information comes from Ogimet.com (Ogimet, 2020), a public meteorological website which provides local weather conditions. Data from the Meteorological Report of Aerodrome Conditions (METAR) of ZSNJ in 2018 are collected, including airport code, UTC time, wind direction, wind speed, wind gusts, visibility, weather phenomena such as precipitation, thunderstorm, fog, snowfall, haze, cloud height, cloud cover, temperature, dew point temperature, QNH, etc. Variables about some weather phenomena are set as dummy variables. The unfavorable weather conditions are categorized into two types, general unfavorable weather such as mist, light rain, shower and rain and severe unfavorable weather such as thunderstorm and snowfall.

Fig. 3 illustrates the daily average CI estimated using Eq. (1) for January 2018 as an example and unfavorable weather duration for each day. Each CI value is calculated over 30 min time period and the daily average CI is the mean of the 30-min CI values for each day. The figure depicts the randomness and fluctuation of the CI value. Previous research indicates that there may be some associations between traffic flow, delay or congestion status and unfavorable weather durations (Roh, 2020; Sridhar and Chen, 2009), especially continuous severe weather conditions, while the relationship needs to be further demonstrated.

Fig. 4 presents the distribution of CI for each airway over different time periods, including 10 min, 30 min and 60 min time intervals. The figure depicts the median, mean and the first to the third quartile values. As CI is defined as the cumulative occupation time of a unit route length, the CI value increases with the increase in time duration. Specifically, the CI value ranges from 0 to 61.09 for the 10-minute time interval, 0 to 152.74 for the 30-minute time interval and 0 to 277.89 for the 60-minute time interval. For a determined time interval, the higher the CI value, the more congested the airway is. When CI equals zero, it means that there is no aircraft in the target airway during this time period. Therefore, it can be found that among all the selected airways, ZSNJ-NJ403 tends to be the most congested, while NJ403-NJ407 is the clearest.

#### 3.2. Data processing

Considering that airspace congestion rarely occurs during nighttime, the flight information from 6:00:00 am to 23:59:59 pm are selected for congestion status analysis and prediction. The selected data are categorized into three types, including spatial features, temporal features and spatiotemporal features, as shown in Table 1. The spatial features are only spatially varied but temporally static during the study period, which refers to length of airways, latitude and longitude of waypoints. The temporal features are only

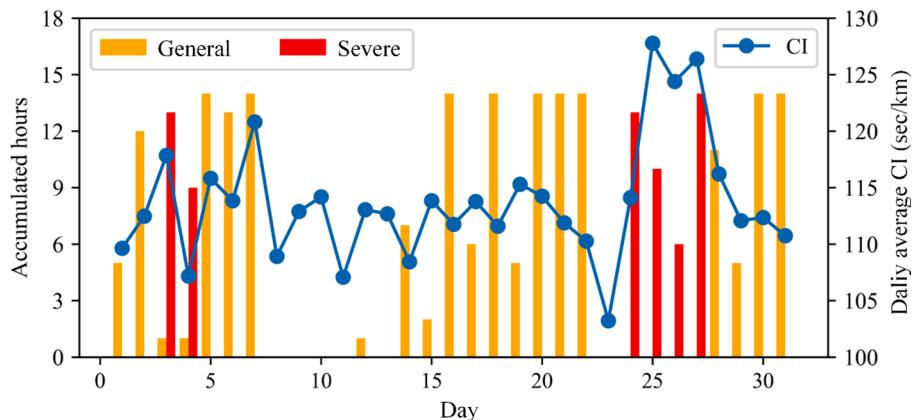
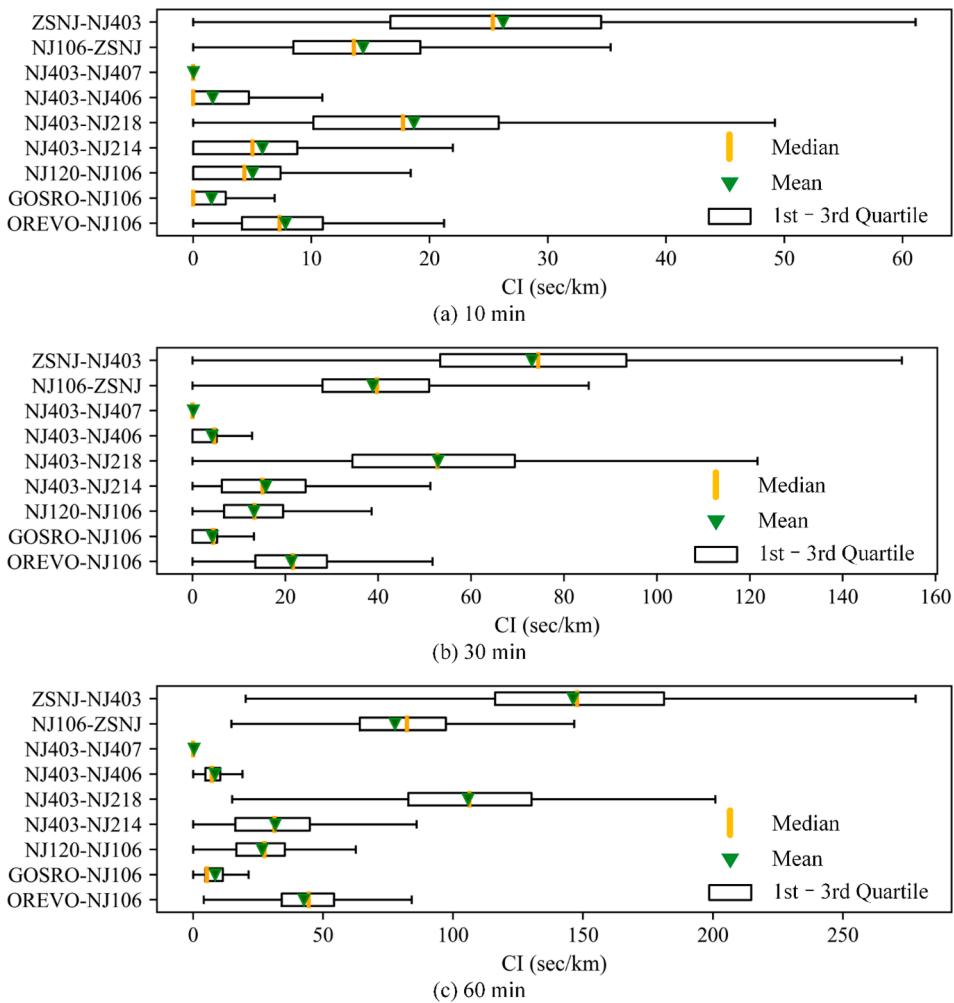


Fig. 3. Daily average CI and unfavorable weather duration.



**Fig. 4.** Distribution of CI for each airway.

temporally varied but spatially static during the study period, including month, weekday, time, general unfavorable weather and severe unfavorable weather. It is assumed that the weather conditions of each route of ZSNJAP01 are consistent during a specific time period. The spatiotemporal features are both spatially and temporally varied during the study period, for which the spatial dependencies and temporal dependencies should be considered simultaneously in the congestion prediction model. The variables include flight flow, scheduled flow, historic flow and historic congestion index.

The selected data are prepared according to the following steps.

Step 1: Transform the time-series data into a sequence  $S_t = \{t_i, i = 1, 2, 3, \dots, N\}$ , where  $N$  is the length of the sequence. According to

**Table 1**  
Description of considered variables.

Category	Name	Description
Spatial features	Airway length	Length for each airway in ZSNJAP01 sector (km)
	Latitude	Latitude of the waypoint
	Longitude	Longitude of the waypoint
Temporal features	Month	Month of the year
	Day	Weekday
	Time	Time in a day
	General unfavorable weather	Dummy variable, general unfavorable weather (mist, light rain, shower and rain) = 1; others = 0
	Severe unfavorable weather	Dummy variable, severe unfavorable weather (thunderstorm and snowfall) = 1; others = 0
Spatiotemporal features	Scheduled flow	Flow of each airway according to flight plan
	Historic flow	Historic average of flight flow for each airway
	Historic congestion index	Historic average of congestion index for each airway

the timestamp of ADS-B data and the longitude and latitude coordinates of the flights passing through the sector, the flight flow  $F_{ri}$  and congestion index  $C_{ri}$  of route  $r$  in ZSNJAP01 sector within time interval  $t_i$  are computed. The scheduled flight flow  $P_{ri}$  of route  $r$  during period  $t_i$  is obtained according to the flight plan. Thus, the basic sequence can be obtained as  $FCP = \{(F_{ri}, C_{ri}, P_{ri}), r = 1, 2, 3, \dots, R; i = 1, 2, 3, \dots, N\}$ , where  $R$  is the total number of airways in the sector. Thus,  $FCP$  can be regarded as a three-dimensional array with dimensions  $[3, R, N]$ .

Step 2: Repeat step one to get the daily feature sequence  $FCP_d$ , and generate the feature sequence set  $FCP_{all} = \{FCP_d, d = 1, 2, 3, \dots, A\}$ , where  $A$  is the total number of days in the sample. The data are randomly selected  $A$  times from  $FCP_{all}$ , among which 80 percent of the selected data are denoted as the training set  $FCP_{tr}$ , and the other 20 percent of the data are denoted as the verification set  $FCP_{va}$  for parameter adjustment. The test set  $FCP_{te}$  is composed of unselected samples.

Step 3: Based on the training set  $FCP_{tr}$ , the series of historical average for  $F$  and  $C$  are calculated, denoted by  $FH = \{FH_{ri}\}$ ,  $CH = \{CH_{ri}\}$ . The formulas are written as:

$$FH_{ri} = \frac{1}{A_{tr}} \sum_{d=1}^{A_{tr}} (F_{ri})_d \quad (2)$$

$$CH_{ri} = \frac{1}{A_{tr}} \sum_{d=1}^{A_{tr}} (C_{ri})_d \quad (3)$$

where  $A_{tr}$  represents the total days of training dataset.

Step 4: The sequences  $FH$  and  $CH$  are added to each  $FCP$  of training set, verification set and test set to form a new feature sequence, written as  $FCPH = \{(F_{ri}, C_{ri}, P_{ri}, FH_{ri}, CH_{ri})\}$ , the new feature sequences are denoted as  $FCPH_{tr}$ ,  $FCPH_{te}$  and  $FCPH_{va}$ , respectively. Then the feature dataset for air route  $r$  can be expressed as  $\{F_r, C_r, P_r, FH_r, CH_r\}$ . In addition, the daily weather sequence is generated according to historical METAR message of ZSNJ airport, represented as  $W_d = \{W_{di}, i = 1, 2, 3, \dots, N; d = 1, 2, 3, \dots, A\}$ .

Step 5: According to the given prediction sequence length  $t_s$ , generate the daily input and output samples, based on processed samples from step 4. The input sample contains three types. The first type is the input incorporating spatial features, which refers to airway length. The second type are the inputs incorporating temporal features, including the forward sequence  $T_f$ , the backward sequence  $T_b$  and dataset  $T_c$  at the target time. The third type are the inputs incorporating spatiotemporal features, which are forward time series dataset  $ST_f$ , the backward time series dataset  $ST_b$  with no  $F$  or  $C$  feature, as well as dataset  $ST_c$  with no  $F$  or  $C$  feature at the target time.

### 3.3. Graph construction

As shown in Fig. 2, there are nine airways in sector ZSNJAP01, among which OREVO-NJ106, GOSRO-NJ106, and NJ120-NJ106 are the airways that pass through the sector in the approach direction, NJ403-NJ214, NJ403-NJ218, NJ403-NJ406, and NJ403-NJ407 are the airways pass through the sector in the departure direction, NJ106-ZSNJ is the approach airway to the airport, and ZSNJ-NJ403 is the departure airway from the airport. To construct the input graph for neural networks, the structure of sector ZSNJAP01 in Fig. 2 can be converted to a directed graph  $G_{se}$  shown in Fig. 5.

The airways in the airspace are connected by waypoints to form an airway network, which provides a great convenience for carrying out related research from the topological structure. In Fig. 2, the nodes in the airway network are waypoints. However, the congestion indicator (CI) defined in this article is a characteristic of the airway. Therefore, to transform the original airway network structure into the input graph for neural networks, the airways are treated as nodes, while the waypoints are treated as edges, as shown

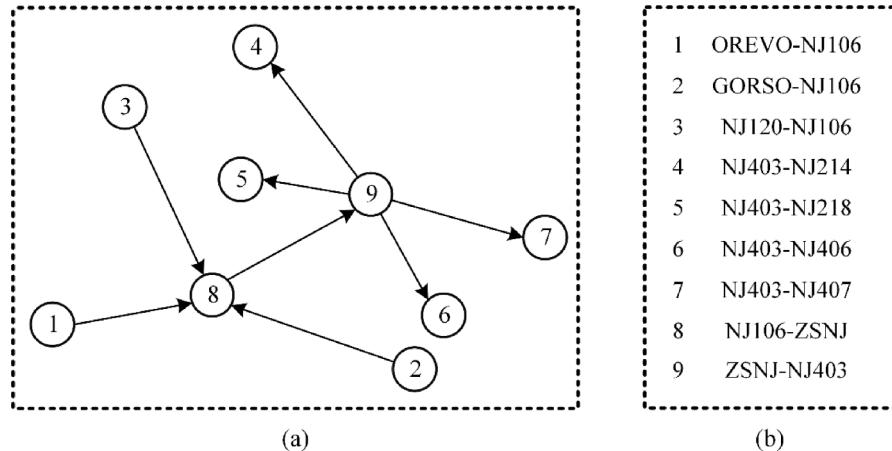


Fig. 5. (a) The constructed graph  $G_{se}$  for sector ZSNJAP01; (b) The corresponding airways in  $G_{se}$ .

in Fig. 5. For example, the node 1 represents the airway OREVO-NJ106, node 8 represents the airway NJ106-ZSNJ, and the edge between node 1 and node 8 represents the waypoint NJ106.

To this end, the original airway network is converted into an unweighted graph  $G = (V, E)$ , where  $V$  is the set of graph nodes, and  $E$  is the set of graph edges. Based on the connections of different airways in Fig. 5, the adjacency matrix ( $A$ ) of  $G$  can be obtained. If there is a connection between two airways, the corresponding element in the adjacency matrix  $A$  is 1, otherwise 0. At the same time, the airway information is reflected as the attribute features of the nodes in  $G$ . Thus, the feature matrix of  $G$  can be expressed as  $X^{n, m}$ , where  $n$  is the number of the nodes and  $m$  is the number of the features. Information about the airways is collected and treated as the attribute feature of each node in  $G_{se}$ . Then, the feature matrix  $X^{n, m}$  of  $G_{se}$  can be obtained.

#### 4. Methodology

In this section, two types of models are constructed, including regression models for congestion index prediction and classification models for congestion identification. As for the regression models, the graph convolution and LSTM networks are integrated for spatiotemporal feature extraction and congestion index prediction. As for the classification models, the XGBoost method is further incorporated for airspace congestion identification. The structures of the proposed methods are presented as follows.

##### 4.1. Congestion index prediction

Toward the objective of predicting the short-term CI of the airway, a hybrid spatiotemporal convolutional long short-term memory

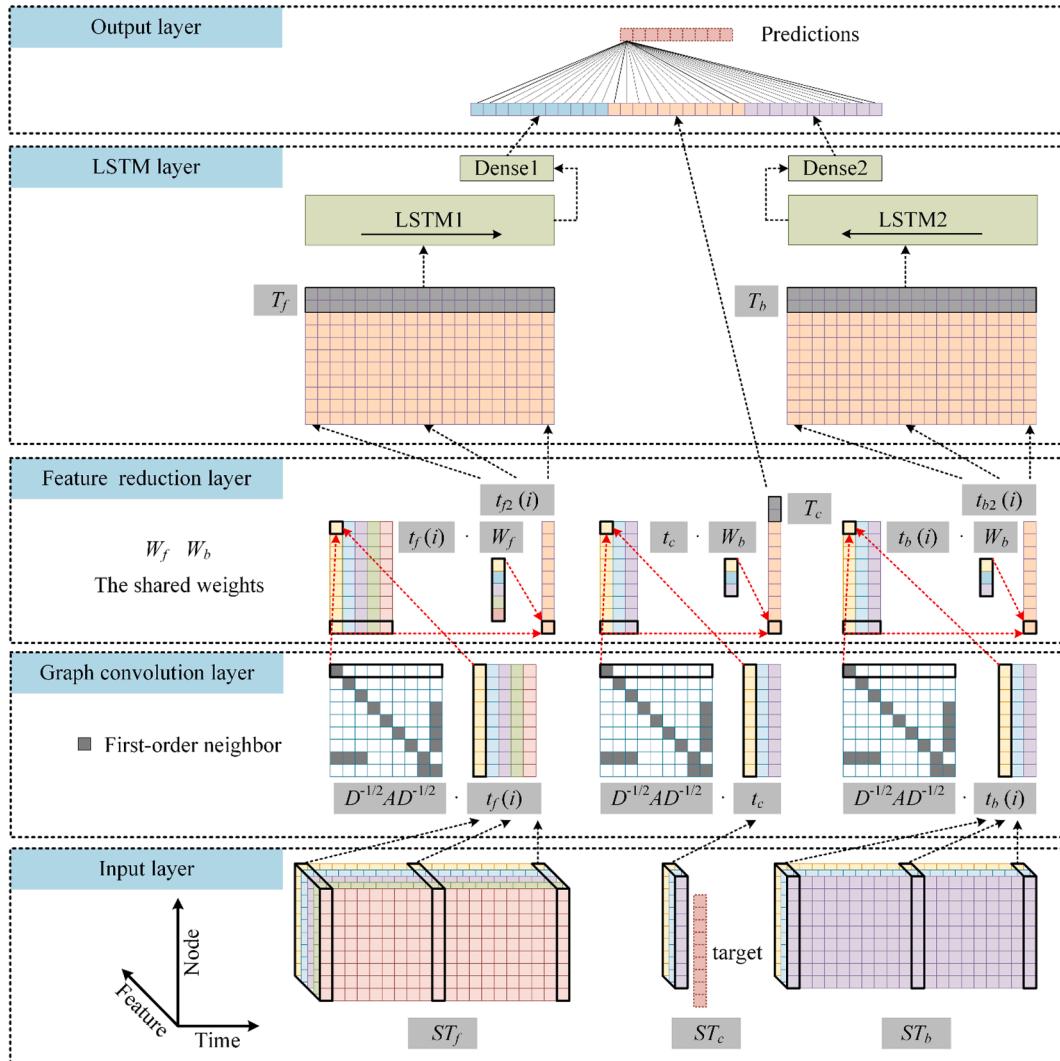


Fig. 6. The proposed SCLN-TTF network.

network with tri-directional temporal features (SCLN-TTF) is constructed. Fig. 6 presents the architecture of the LSTM-TG method. The model is composed of five layers, including the input layer, graph convolution layer, feature reduction layer, LSTM layer and output layer.

As mentioned above, the processed data are categorized into three types, including spatial features, temporal features and spatiotemporal features. The temporal and spatiotemporal input datasets are categorized into three groups. The first group are forward time series datasets including spatiotemporal sequence  $ST_f$  with dimension  $[ts, R, f_1]$ , where  $f_1$  is the number of spatiotemporal features, and temporal sequence  $T_f$  with the dimension  $[ts, f_2]$ . The second group are datasets at the target time, including spatiotemporal sequence  $ST_c$  with dimension  $[R, f_1-2]$  with no  $F$  or  $C$  feature, as well as temporal sequence  $T_c$  with dimension  $[1, f_2]$ . The third group are backward time series datasets including spatiotemporal sequence  $ST_b$  with dimension  $[ts, R, f_1-2]$  with no  $F$  or  $C$  feature, and temporal sequence  $T_b$  with dimension  $[ts, f_2]$ .

The graph convolution layer is developed to capture the spatial dependencies and extract the spatial features for the three groups of data. Specifically, the data for each moment of  $ST_f$ ,  $ST_b$ ,  $ST_c$  and  $T_f$ ,  $T_b$ ,  $T_c$  in the input layer are extracted through the graph convolution layer. The shapes of  $ST_f$ ,  $ST_b$ ,  $ST_c$  are  $[R, f_1]$ ,  $[R, f_1-2]$  and  $[R, 1]$ , respectively. Then the features are reduced with shared weights. The shared weights are reflected in two aspects. On one hand, different features of all nodes are given the same weight for reducing the dimension of feature. For example, if the feature reducing weight of node1 is  $[0.2, 0.2, 0.2, 0.2, 0.2]$  in the feature reduction layer, the feature reducing weights of other nodes are also the same. On the other hand,  $ST_b$  and  $ST_c$  use the same feature reducing weight  $W_b$ , because both  $ST_b$  and  $ST_c$  represent future data input and have the same number of features. It can be considered that  $ST_b$  and  $ST_c$  are two parts of a sequence.

The forward and backward sequences obtained from the feature reduction layer are input into the LSTM layer, which is developed to capture the temporal dependencies and extract the temporal features. All the extracted high-level features from the layers above are further merged and input into multiple fully connected layers to generate the final predicted congestion index. The methods used in each component are described as follows.

#### (a) Spatial features extracted from GCN

Considering that the airways in the constructed graph are directional, the fast-approximate convolutions on graphs model based on symmetric normalized Laplacian is used to extract the spatial features (Hammond et al., 2011; Kipf and Welling, 2017). The equations are listed as follows:

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}}\bar{A}D^{-\frac{1}{2}}H^{(l)}W^{(l)}) \quad (4)$$

where  $H^{(l)}$  is the matrix of activations of  $l^{\text{th}}$  layer with dimension  $[R, M]$ , where  $M$  is equal to  $f_1$  (forward) or  $f_1-2$  (backward).  $H^{(0)} = X$ , and  $X$  is the initial feature matrix of node.  $\bar{A} = A + I_R$  is the adjacent matrix of graph  $G$  and self-connections.  $I_R$  is the identity matrix.  $D$  is the degree matrix of matrix  $\bar{A}$ ,  $D_{ii} = \sum_j \bar{A}_{ij}$ .  $W^{(l)}$  is the trainable weight matrix  $[M_l, M_{l+1}]$ , which is used to change the feature dimension of graph nodes. In this article, in order to input the spatial features extracted based on graph into the recurrent neural network layer in time series for temporal feature extraction, the output feature dimensions after graph convolution are all set to be 1.  $\sigma$  is the activation function.

#### (b) Temporal features extracted from LSTM

The LSTM algorithm is used to extract temporal features such as flight features and weather features in this paper (Hochreiter and Schmidhuber, 1997). Although there are many improved versions of LSTM, it has been proven that as long as the version containing the forget gate and the activation function, the prediction accuracy is generally comparable (Greff et al., 2017). The most commonly used LSTM architecture in literature is an improved LSTM network, which realizes the retention and forgetting of information by setting more complex structures such as input gate, forget gate, cell gate and output gate (Graves and Schmidhuber, 2005). The single

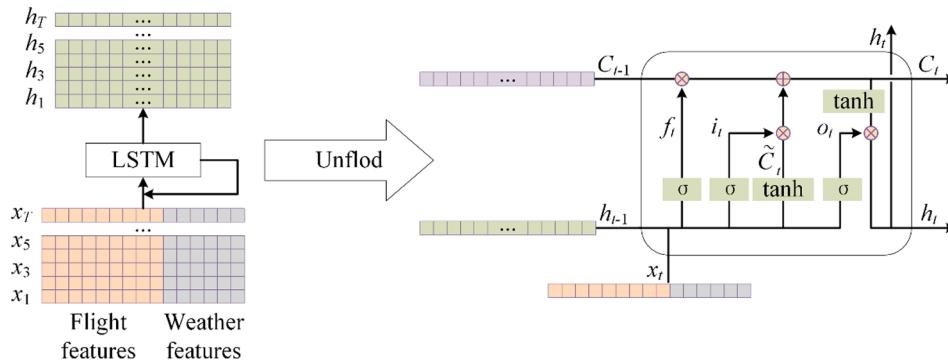


Fig. 7. A LSTM cell.

neuron of the LSTM network in this study is shown in Fig. 7.

The input of the neuron is the airway, flight and weather information  $x_t$  at time  $t$ , the memory information at time  $t-1$ , which is called the hidden state  $h_{t-1}$  and cell state  $C_{t-1}$ , and the outputs are the hidden state  $h_t$  and cell state  $C_t$  at time  $t$ . In this paper, we follow the formulation of FC-LSTM (Graves, 2014). The calculation formulas are:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

$$c_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (7)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot c_t \quad (8)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t \odot \tanh(C_t) \quad (10)$$

where  $\sigma$  is the logistic sigmoid function;  $f$ ,  $i$ ,  $o$  and  $c$  are the forget gate, input gate, output gate, and cell input activation vectors, respectively, all of which are the same size as the hidden cell size.  $W_f$ ,  $W_i$ ,  $W_c$ ,  $W_o$  are the trainable weight matrices; and  $b_f$ ,  $b_i$ ,  $b_c$ ,  $b_o$  are biases in  $f$ ,  $i$ ,  $o$ ,  $c$ , respectively.

### (c) Loss function

In order to prevent over-fitting due to insufficient samples, L2 regularization method is applied as the loss function on the mean square error for the regression model. The formula is shown as follows:

$$Loss_1 = \frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2 + \lambda \sum_k w_k^2 \quad (11)$$

where  $y_i$  and  $y'_i$  are the actual and predicted values of sample  $i$  respectively;  $\lambda$  is the proportion of regularization in the total loss;  $w$  is the trainable weights of the regression model.

### (d) Evaluation metrics

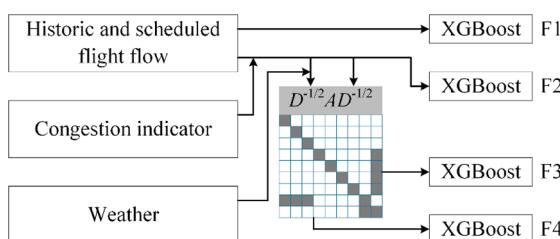
To evaluate the performance of the proposed model, mean absolute error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE) are calculated for each method, respectively. The equations are shown as follows.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y'_i - y_i| \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2} \quad (13)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y'_i - y_i}{y_i} \right| \quad (14)$$

where  $y_i$  and  $y'_i$  are the actual and predicted category of test sample  $i$  respectively, and  $n$  is the test sample size.



**Fig. 8.** Structure of the GCN- XGBoost method.

#### 4.2. Congestion identification

To further explore the congested segments and time periods of air routes, in this section, a graph convolutional network-extreme gradient boosting (GCN-XGBoost) method is proposed for congestion identification of air routes through classification. The proposed GCN-XGBoost model includes two components: the graph convolutional network (GCN) and the extreme gradient boosting (XGBoost) method. The structure of the GCN-XGBoost method is shown in Fig. 8.

As shown in Fig. 8, four different models are constructed in terms of various input features. Specifically, Model F1 takes the features associated with historic and scheduled flight flow as the inputs, including historic average of flight flow at the target time, a sequence of historic flight flow dataset and a sequence of feature flight flow dataset according to scheduled flight plan. Model F2 takes the features associated with CI as inputs in addition to the historic and scheduled flight flow features. Model F3 takes the same inputs as Model F2 but incorporates the graph convolution model. Model F4 further integrates the weather-related information based on Model F3. It should be noted that in Model F4, the spatial and spatiotemporal features are input into the graph convolution layer first. Then the XGBoost method is applied to make predictions. Description of the application of the XGBoost method in this study is illustrated as follows.

##### (a) XGBoost

XGBoost is an implementation algorithm of Granding Boosting Mechine (GBM). The difference is that GBM uses the first derivative of the loss function, while XGBoost performs the second-order Taylor expansion of the loss function, using both the first and second derivatives (Chen and Guestrin, 2016; Friedman, 2001). The objective function of  $t^{\text{th}}$  tree is:

$$Obj^{(t)} = \sum_{i=1}^n l(y_i, y_i'^{(t-1)} + f_t(x_i)) + \Omega(f_t) + c \quad (15)$$

where  $\Omega(f_t)$  is the complexity of the  $t$ -th tree;  $c$  means the information before the  $t^{\text{th}}$  tree has been determined. The regularization term can be considered constant if  $\Omega(f_t)$  is removed.

Combined with second-order Taylor expansion,  $Obj^{(t)}$  can be expressed as:

$$Obj^{(t)} \approx \sum_{i=1}^n \left( L(y_i, y_i'^{(t-1)}) + g f_t(x_i) + \frac{1}{2} h f_t^2(x_i) \right) + \Omega(f_t) + c \quad (16)$$

where  $g_i$  and  $h_i$  are the first and the second derivative of the loss function to  $y_i'^{(t-1)}$ .

As  $L(y_i, y_i'^{(t-1)})$  is known and integrated into the constant term, by combining the structure of the decision tree,  $Obj^{(t)}$  becomes:

$$\begin{aligned} Obj^{(t)} &\approx \sum_{i=1}^n \left( g f_t(x_i) + \frac{1}{2} h f_t^2(x_i) \right) + \Omega(f_t) + c \\ &= \sum_{i=1}^n \left( g f_t(x_i) + \frac{1}{2} h f_t^2(x_i) \right) + \gamma N + \frac{1}{2} \lambda \sum_{j=1}^N w_j^2 + c \end{aligned} \quad (17)$$

where  $N$  represents the number of leaf nodes of the tree,  $w_j^2$  represents the weight of the  $j^{\text{th}}$  leaf node,  $\lambda$  and  $\gamma$  are hyper parameters.

Let

$$g f_t(x_i) = (\sum_{i \in I_j} g_i) w_j = G_j w_j \quad (18)$$

$$h f_t(x_i) = (\sum_{i \in I_j} h_i) w_j = H_j w_j \quad (19)$$

The objective function is changed from the cumulative number of sample sizes to the cumulative number of leaf nodes.  $Obj^{(t)}$  becomes:

$$Obj^{(t)} = \sum_{j=1}^N \left[ G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma N + c \quad (20)$$

From Eq. (20), the derivation of objective function can be obtained as:

$$\frac{\partial [Obj^{(t)}]}{\partial w} = G_j + (H_j + \lambda) w_j \quad (21)$$

Let Eq. (21) be equal to zero, the extreme point of the objective function can be found, which is:

$$w_j = -\frac{G_j}{H_j + \lambda} \quad (22)$$

Therefore, the  $Obj^{(t)}$  can be simplified as follows:

$$Obj(t) = -\frac{1}{2} \sum_{j=1}^N \frac{G_j}{H_j + \lambda} + \gamma N \quad (23)$$

### (b) Loss function

The L2 regularization is added as the loss function on cross entropy error for classification model.

$$Loss_2 = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^T p_{ij} \log q_{ij} + \Omega_2(W_2, N) \quad (24)$$

where  $p_{ij}$  and  $q_{ij}$  are the actual and predicted probabilities of sample  $i$  with category  $j$  respectively; and  $\Omega_2$  is L2 regulation;  $W_2$  is the weight of the classification model;  $N$  is the number of leaf nodes of the tree.

### (c) Evaluation metrics

To evaluate the performance of the classification, the accuracy ( $Acc$ ), precision ( $P$ ) and recall ( $R$ ) metrics are calculated for each method, respectively. The equations are shown as follows.

$$Acc = \frac{1}{m} \sum_{i=1}^m I(y'_i = y_i) \quad (25)$$

$$P = \frac{TP}{TP + FP} \quad (26)$$

$$R = \frac{TP}{TP + FN} \quad (27)$$

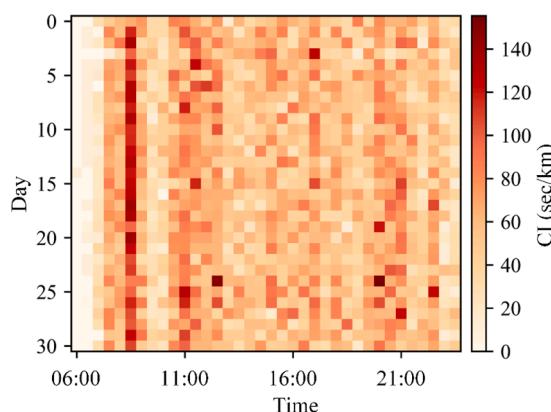
where  $y_i$  and  $y'_i$  are the actual and predicted category of test sample  $i$  respectively;  $TP$ ,  $FP$ ,  $FN$  and  $TN$  are numbers of true positive, false positive, false negative and true negative samples, and  $m$  is test sample size.

## 5. Data analysis results

This section provides the spatiotemporal pattern of the calculated CI first. Then the model training, testing and evaluation process of the proposed neural network is illustrated the [Section 5.2](#). Results from the proposed SCLN-TTF and GCN-XGBoost methods are presented for CI prediction and congestion recognition in the [Sections 5.3 and 5.4](#), respectively.

### 5.1. Spatiotemporal pattern of CI

To identify similar congested time periods or frequently congested airways, the spatiotemporal pattern of the congestion index is depicted, so that air traffic management initiatives can be scheduled in advance ([Kuhn, 2016](#)). [Fig. 9](#) presents the calculated hourly CI on airway NJ403-NJ218 for January 2018 as an example. The airway NJ403-NJ218 is a departure approach. The horizontal axis represents time of each day (from 6:00 am to 23:59 pm) and the vertical axis represents the  $i^{\text{th}}$  day of the selected month. The higher the



**Fig. 9.** Hourly CI on airway NJ403-NJ218.

CI value, the more congested the airway is. As shown in the CI distribution in Fig. 4 (b), the highest CI value is 152.74 sec/km for 30 min time interval. It is shown that there tends to be some similarities across different days. The figure presents the peak agglomeration characteristics through the whole day. The phenomena are generally consistent with that in previous study (Dong et al., 2019). As shown in Fig. 9, the most prominent peak occurs at around 9:00 am each day. There are also some small peaks at around 11:00 am and 21:00 pm.

Fig. 10 illustrates the comparison of CI for airways in the sector from 06:00 to 23:59. The figure depicts the structural characteristics of unbalanced coverage across different airways and concentrated distribution to the crossing points, e.g. ZSNJ-NJ403. Fig. 11 demonstrates the congestion index with respect to scheduled flight flows based on 80\*5 randomly selected samples. The duration for each sample is 30 min. It can be found that the trend of CI is generally consistent with planned flow rate. The larger the scheduled number of flights, the airway is more congested. However, it should also be noted that with a determined scheduled flight flow, there are some differences in CI across the randomly selected samples. It is due to the reason that a series of reasons may affect the congestion status of the airway, which deviate the actual flow rate from the scheduled one, such as real-time weather conditions, congestion propagation from neighboring airways, etc.

## 5.2. Model training and testing

First, a neural network structure including a single LSTM layer and a single dense layer is constructed. The initial batch sample size is set as 128. The mean square error (MSE) is applied as the loss function  $Loss_1$ . When using L2 regularization, the function of  $Loss_1$  is expressed as Eq. (11). The time step is set as 10. Five different optimizers are tested and compared under different learning rates (LR), including SGD, RMSProp, AdaDelta, Adam and Nadam optimizers. The results are shown in Fig. 12. It can be found that the loss of using SGD and AdaDelta optimizers converge quite slowly as compared with the other three optimizers. As for the RMSProp, Adam and Nadam optimizers, the convergence speed and training performance are quite similar. Their performances under different learning rates are tested to determine an appropriate value.

As can be seen from Fig. 13, through the case of Adam and Nadam, the two optimizers are extremely similar in performance under different learning rates. It is also found that a smaller learning rate may require more training time, while a larger learning rate may lead to the training loss unable to converge to the ideal value. But generally, the learning rate from 0.0001 to 0.001 can produce satisfactory results in this experiment. The Adam optimizer is then used in subsequent experiments. Therefore, the appropriate optimizer and initial learning rate can be determined for a single LSTM and density layer. Further, the impact of input time step, the LSTM and dense network layers on the network performance are tested. The learning rate is set as 0.001, with Adam as the optimizer. The time step takes the value from 5 to 20, with 5 as the interval. The number of LSTM layers is increased from 1 to 5. The training loss and errors of validation set are recorded. The results are shown in Fig. 14.

According to Fig. 14, it can be found that for each scenario, the  $Loss_1$  of training and validation sets increase sharply when the LSTM layer is over three, as shown in the solid lines in Fig. 14. The  $Loss_1$  for single and double LSTM layers are generally comparable. The loss of the validation set is greater than that of the training set. Considering the problem of overfitting, L2 regularization is added to the loss function, as shown in the dashed lines in Fig. 14. Due to the introduction of the regularization term, the value of  $Loss_1$  in the training set will increase compared to MSE. However, it can be found that adding L2 regularization helps to reduce the loss of the validation set. At the same time, the double-layer LSTM structure is slightly superior to the single-layer structure for this case. On the basis of determining the two-layer LSTM, the continuous time step value is further tested for performance comparison. The results are depicted as

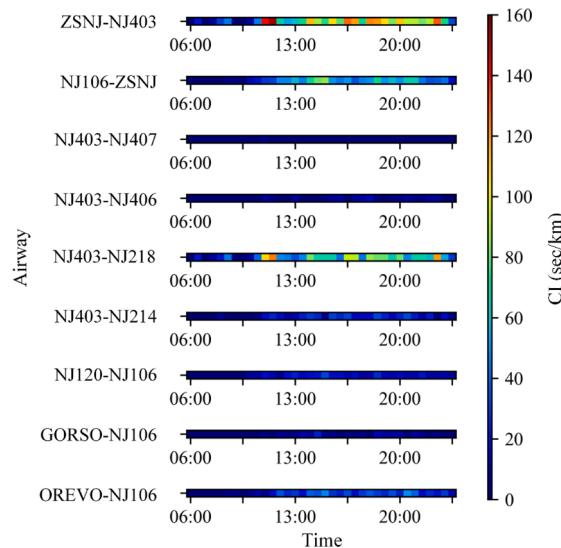


Fig. 10. Comparison of CI across different airways.

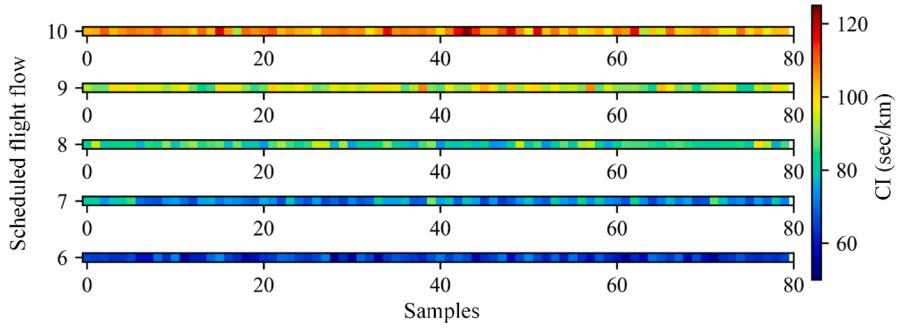


Fig. 11. Comparison of CI across different scheduled flights.

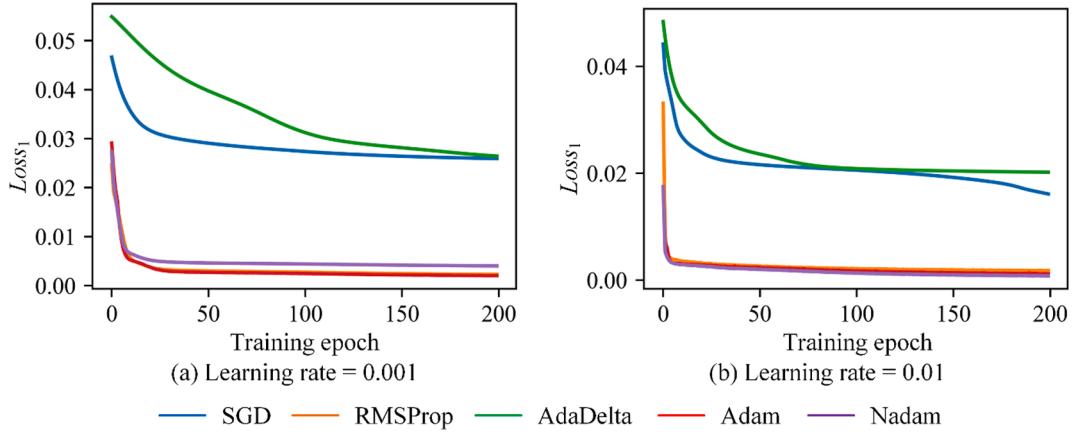


Fig. 12. Performance comparison for five different optimizers.

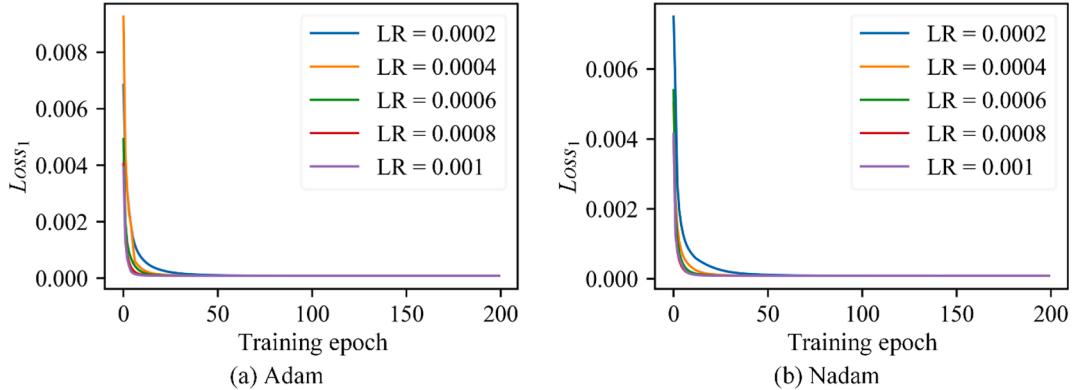


Fig. 13. Comparison of learning rate on model performance for the Adam and Nadam optimizers.

the orange “Validation” line in Fig. 15. It can be found that with the increase of time step, the loss of the validation set gradually decreases, tends to be stable, and then increases after a certain limit. The SCLN-TTF model is then constructed by adding the graph convolution layer after the input layer. The loss of the validation set based on the SCLN-TTF model is illustrated as the red “Validation\_G” line in Fig. 15. After adding the graph convolution layer, the loss of validation set can be reduced by about 0.001, while the loss decreases first and then increases with the increase of time step similar as the orange line. The final time step is set to be 5. So far, the main parameters of the SCLN-TTF network structure can be determined.

### 5.3. Results CI prediction based on the SCLN-TTF method

To testify the performance of the proposed SCLN-TTF model, several benchmark methods are also tested and compared. The

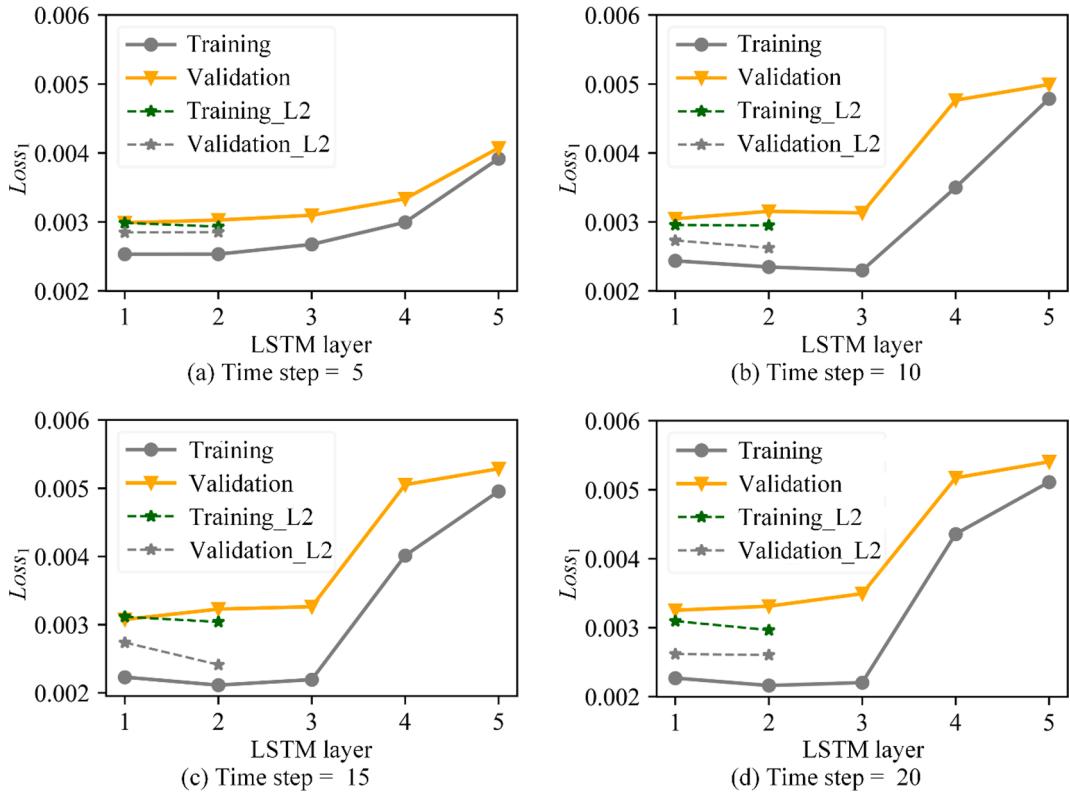
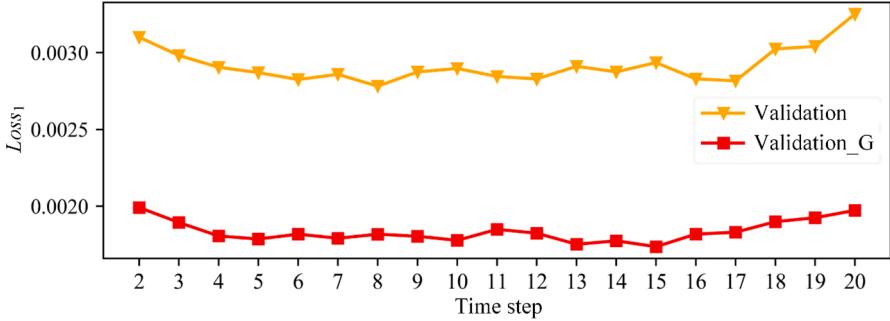
Fig. 14. Impacts of time step on  $Loss_1$ .

Fig. 15. Determination of time step for the SCLN-TTF model.

selected benchmark methods include historical average (HA), back propagation (BP) neural network, gradient boosting regression tree (GBRT), LSTM with forward temporal features (LN-TTF), spatiotemporal convolutional LSTM with forward temporal features (SCLN-TTF), and LSTM with tri-directional temporal features (LN-TTF), which are commonly used in previous studies of short-term time series data prediction. The temporal, spatial, and spatiotemporal attributes for different methods are illustrated in Fig. 16. Descriptions of the seven selected benchmark methods are shown as follows.

**Historical average (HA):** The historical average of CI at time  $t$  with scheduled flight flow of  $P$  ( $HA_{t,P}$ ) is calculated as the weighted sum of  $HA_{CI,t}$  and  $HA_{CI,P}$ , where  $HA_{CI,t}$  represents the historical average of CI at time  $t$  and  $HA_{CI,P}$  represents the historical average of CI with scheduled flight flow of  $P$ . To provide accurate weights for  $HA_{CI,t}$  and  $HA_{CI,P}$ , the predicted value is obtained through a trained single layer fully connected feedforward network.

**Back propagation neuron network (BP):** Concatenate the forecasted CI of the airway based on historical values ( $HA_{t,P}$ ) and the other features (weather condition, scheduled flow, time) that can be obtained at the target time as the inputs of the BP neural network. The single layer network is applied, with learning rate set as 0.05 and activation function set as ‘relu’.

**Gradient boosting regression tree (GBRT):** Inputs of the GBRT model are the same as that of the BP network. The learning rate is set as 0.05; the estimator number is set as 120; and the max depth is set as 5.

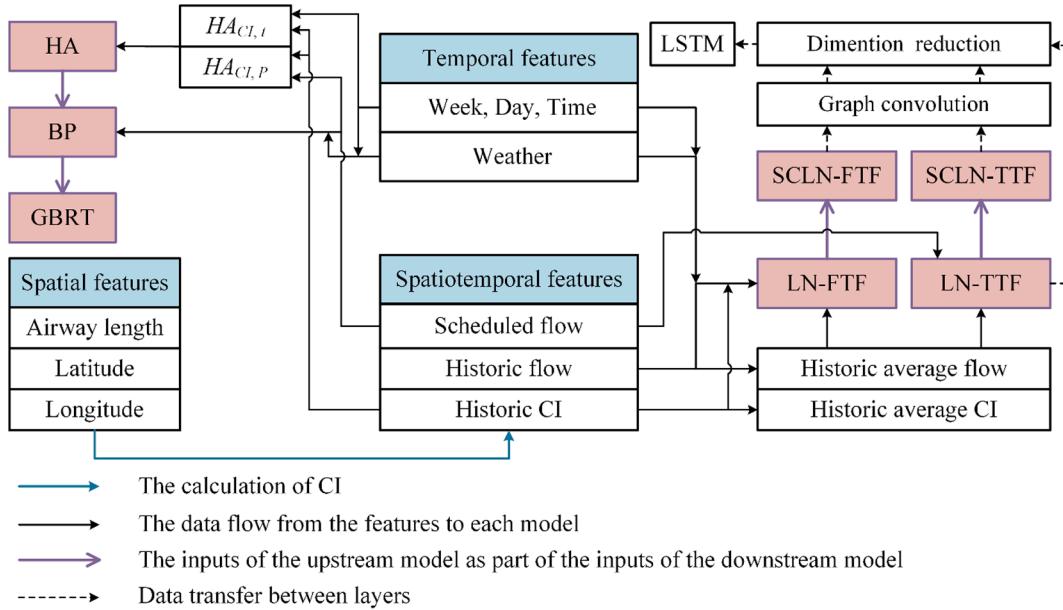


Fig. 16. The temporal, spatial, and spatiotemporal attributes for different methods.

**LSTM with forward temporal features (LN-FTF):** The input data consist of  $ST_f$  and  $T_f$  as described in Section 3.2. The shared weights  $W_f$  is applied to reduce the dimensionality. Then concatenate  $T_f$  into the LSTM layer. Apart from the input data, the graph convolution layer is not applied in the network, which is another difference from the procedure in Fig. 6.

**Spatiotemporal convolutional LSTM with forward temporal features (SCLN-FTF):** The graph convolution layer is added into the structure of LN-FTF.

**LSTM with tri-directional temporal features (LN-TTF):** The graph convolution layer is deleted from the structure in Fig. 6. Instead, the inputs are directly transferred to the feature reduction layer.

All the benchmark methods are trained and tested with the same input variables, so as to ensure that the models are comparable. The four airways with denser flights in the ZSNJAP01 sector are selected for testing, including OREVO-NJ106, NJ120-NJ106, NJ403-NJ214 and NJ403-NJ218. The prediction time interval is set as 10 min, 30 min and 60 min, respectively. Table 2 and Fig. 17 summarize the MAE, MAPE and RMSE for different methods.

Based on the experimental results, it can be found that, in general, the proposed SCLN-TTF method achieves the most stable and satisfactory performance for all four airways. Specifically, for the 10 min prediction interval, the proposed SCLN-TTF method outperforms all the other benchmark methods in terms of lower MAE, MAPE and RMSE, followed by SCLN-FTF, and then LN-TTF, LN-FTF, GBRT, BP and HA. The performance of SCLN-TTF and SCLN-FTF are much better than that of the other methods, as the two models consider both the spatial and temporal features by integrating the LSTM and GCN structures. Besides, the SCLN-TTF shows superiority compared to the SCLN-FTF method as it applies the bi-directional structure by incorporating both the forward and backward sequence. As for the 30 min and 60 min prediction, the performance of SCLN-TTF, SCLN-FTF, LN-TTF and LN-FTF are generally comparable, which are still slightly better than GBRT and outperform BP and HA methods. In addition, for most of the selected methods, the MAPE decrease sharply with the increase in prediction time horizon. The models exhibit the best performance for the 60 min prediction time horizon, while exhibiting the worst performance for the 10 min prediction time horizon, due to the randomness of the influencing factors.

#### 5.4. Results of congestion identification based on the GCN-XGBoost method

To further investigate the robustness and applicability of the proposed model, the prediction accuracy for congestion identification is examined. The predicted data are labelled as congested or non-congested according to the congestion alert threshold value. In the present study, the 80th-percentile and 90th-percentile congestion index in the sample are applied as the alarm threshold values, respectively. The comparison results in terms of accuracy, precision and recall metrics are given in Table 3, Table 4, and Fig. 18.

According to the prediction performance metrics, it can be obtained that, regardless of the prediction time interval and the alarm threshold, for each scenario, Model F4 has the best performance in terms higher accuracy rate, precision rate and recall rate metrics, followed by Model F3, F2 and F1. The finding further confirms the superiority of the proposed method by considering both the spatial and temporal dependencies and integrating forward and backward features. Similarly, the classification accuracy increases with the increase of prediction time interval, while the performances for the two alarm threshold values are generally comparable for most conditions.

**Table 2**

Performance comparison of the selected models for CI prediction in terms of MAE (km/sec), MAPE (%) and RMSE (km/sec).

Prediction Interval	Models	OREVO-NJ106			NJ120-NJ106			NJ403-NJ214			NJ403-NJ218		
		MAE	MAPE	RMSE									
10 min	HA	3.1881	39.02	4.4545	2.6151	33.85	4.0337	2.7519	29.23	4.4473	7.2539	34.02	10.3515
	BP	2.1227	25.20	2.9107	2.4631	38.46	3.4881	2.4522	27.60	3.6316	3.7511	16.82	4.9797
	GBRT	2.0544	24.05	2.7580	2.2526	35.28	3.0399	2.0640	25.75	3.0026	3.3459	15.15	4.4307
	LN-FTF	2.1951	25.74	2.8786	2.1474	34.80	2.9113	2.1229	26.00	3.0173	3.4113	14.82	4.4961
	SCLN-FTF	1.2584	21.23	1.6242	1.9487	29.37	2.6736	1.9509	23.57	2.7224	3.3851	14.01	4.4307
	LN-TTF	2.1348	24.56	2.8384	2.0617	32.41	2.7944	2.0197	25.24	2.8355	3.3982	14.62	4.4569
	SCLN-TTF	<b>1.2383</b>	<b>20.51</b>	<b>1.6081</b>	<b>1.8317</b>	<b>29.06</b>	<b>2.5333</b>	<b>1.8625</b>	<b>21.37</b>	<b>2.6291</b>	<b>3.2414</b>	<b>13.71</b>	<b>4.3523</b>
30 min	HA	6.7860	21.73	9.1736	5.7745	31.11	8.0552	6.6210	26.30	9.8938	16.1399	22.90	22.4733
	BP	4.5988	15.92	6.1611	3.8959	20.35	5.0328	4.1028	16.89	5.4232	10.0619	13.48	13.3308
	GBRT	3.0605	11.10	3.8858	2.8145	13.53	3.8404	2.7635	11.02	3.7821	4.0861	6.49	5.1842
	LN-FTF	2.8282	10.34	3.6133	2.6758	11.41	3.8196	2.6880	9.74	3.9613	4.3159	7.25	5.9248
	SCLN-FTF	<b>2.6760</b>	10.08	<b>3.4211</b>	<b>2.6273</b>	11.07	3.8127	2.6503	<b>8.71</b>	3.7538	<b>3.9839</b>	6.57	5.6949
	LN-TTF	2.7000	10.23	3.4371	2.6620	11.28	3.8612	2.6692	9.59	3.8293	4.1116	6.43	5.3885
	SCLN-TTF	2.7561	<b>9.97</b>	3.5733	2.6342	<b>10.26</b>	<b>3.6117</b>	<b>2.5466</b>	8.89	<b>3.4803</b>	<b>3.9839</b>	<b>6.30</b>	<b>5.1586</b>
60 min	HA	10.9524	45.36	15.1831	8.5178	52.04	11.8096	10.8623	47.89	15.0605	24.9223	30.58	33.8929
	BP	5.6146	11.32	6.5767	3.0713	11.88	3.8890	4.6664	10.57	6.6017	10.5094	8.57	12.3861
	GBRT	2.9391	6.73	3.8353	2.8811	10.20	3.9746	3.2462	8.05	4.4947	4.4665	4.22	6.0805
	LN-FTF	<b>2.7150</b>	6.38	<b>3.6508</b>	2.5483	9.17	3.6418	2.9497	7.54	3.9485	<b>3.9786</b>	3.80	5.1421
	SCLN-FTF	2.8732	5.65	3.7035	<b>2.4912</b>	<b>6.40</b>	<b>3.3090</b>	<b>2.3722</b>	<b>4.63</b>	3.1526	4.2038	3.45	5.3673
	LN-TTF	2.8995	6.86	3.7958	2.5103	9.37	3.5752	2.8560	6.99	3.8705	3.9786	3.70	<b>5.0670</b>
	SCLN-TTF	3.2159	<b>5.38</b>	3.7694	2.5578	6.60	3.3375	2.3878	<b>4.63</b>	<b>3.1214</b>	4.1662	<b>3.16</b>	5.5174

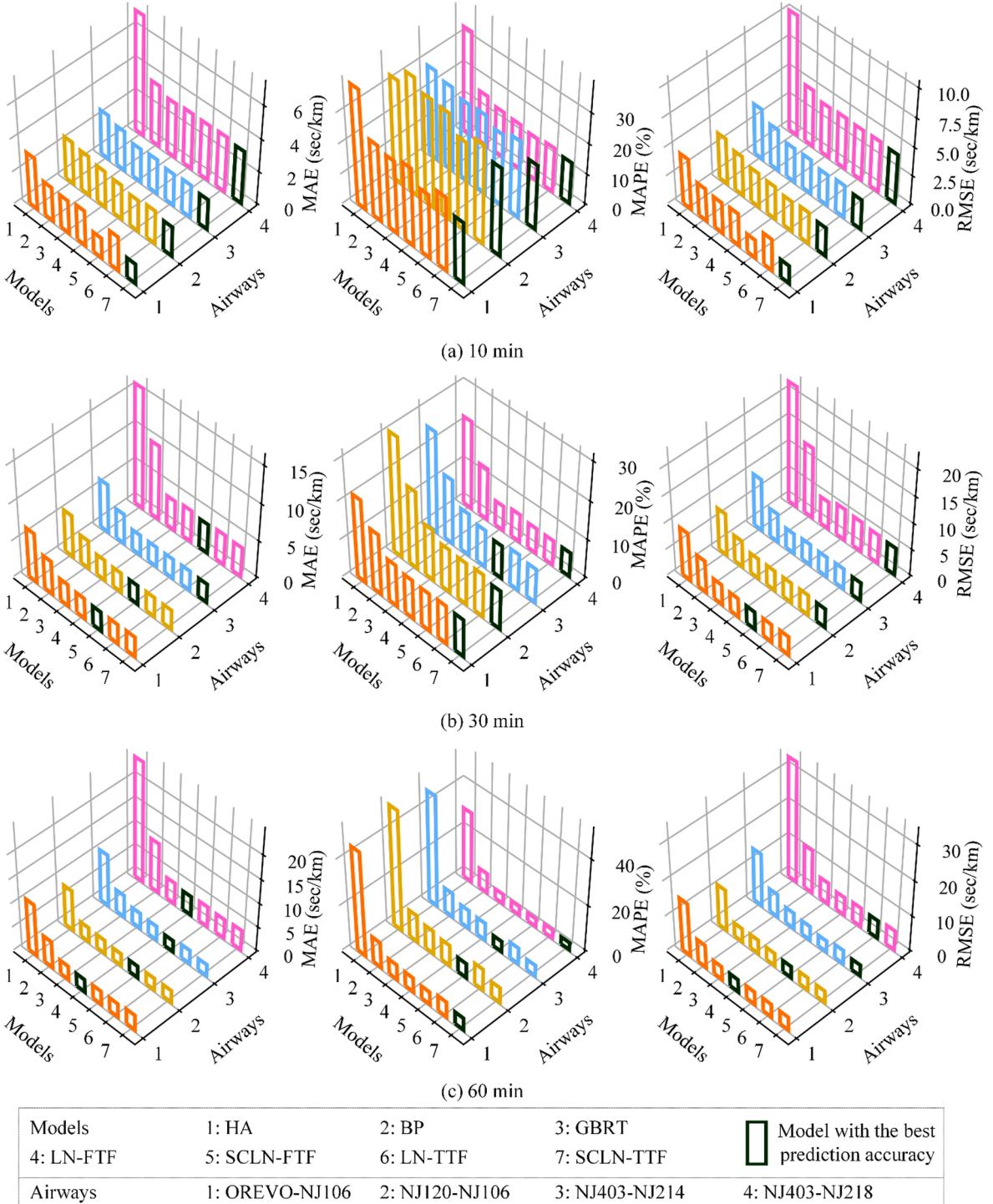


Fig. 17. Performance comparison of the selected models for CI prediction.

## 6. Conclusion and discussion

This paper proposes a new method for airway congestion prediction and identification. A CI indicator is defined for demonstrating the congestion status of the selected airway. The proposed method consists of two parts, including the SCLN-TTF method based on regression models for congestion index prediction, as well as the GCN-XGBoost method based on classification models for high congestion index identification. As for the regression models, the graph convolution and LSTM networks are integrated for

**Table 3**

Performance comparison for congestion identification (80th percentile as the alarm threshold).

Prediction Interval	Models	Accuracy	Precision	Recall
10 min	F1	86.15	78.25	87.78
	F2	88.86	82.47	88.06
	F3	89.48	82.48	90.83
	F4	90.84	85.31	90.83
30 min	F1	89.61	64.36	84.19
	F2	90.38	66.01	88.03
	F3	94.77	78.48	94.02
	F4	95.26	79.21	94.44
60 min	F1	88.29	75.30	82.27
	F2	89.39	76.77	85.00
	F3	95.86	88.54	97.73
	F4	96.34	90.11	98.18

**Table 4**

Performance comparison for congestion identification (90th percentile as the alarm threshold).

Prediction Interval	Models	Accuracy	Precision	Recall
10 min	F1	84.38	75.09	86.87
	F2	86.13	77.17	89.67
	F3	89.75	85.12	89.33
	F4	90.62	86.17	90.67
30 min	F1	86.69	72.38	86.99
	F2	89.45	77.26	90.60
	F3	94.48	87.60	94.87
	F4	95.80	90.18	95.72
60 min	F1	88.37	68.25	84.61
	F2	91.39	75.10	88.46
	F3	97.70	88.17	96.23
	F4	97.73	89.76	98.15

spatiotemporal feature extraction and congestion index prediction. The temporal and spatiotemporal input datasets are categorized into three groups, including forward time series datasets, backward time series datasets and datasets at the target time. As for the classification models, the XGBoost method is further incorporated for airspace congestion identification. Based on the model estimation results, the following conclusions can be obtained.

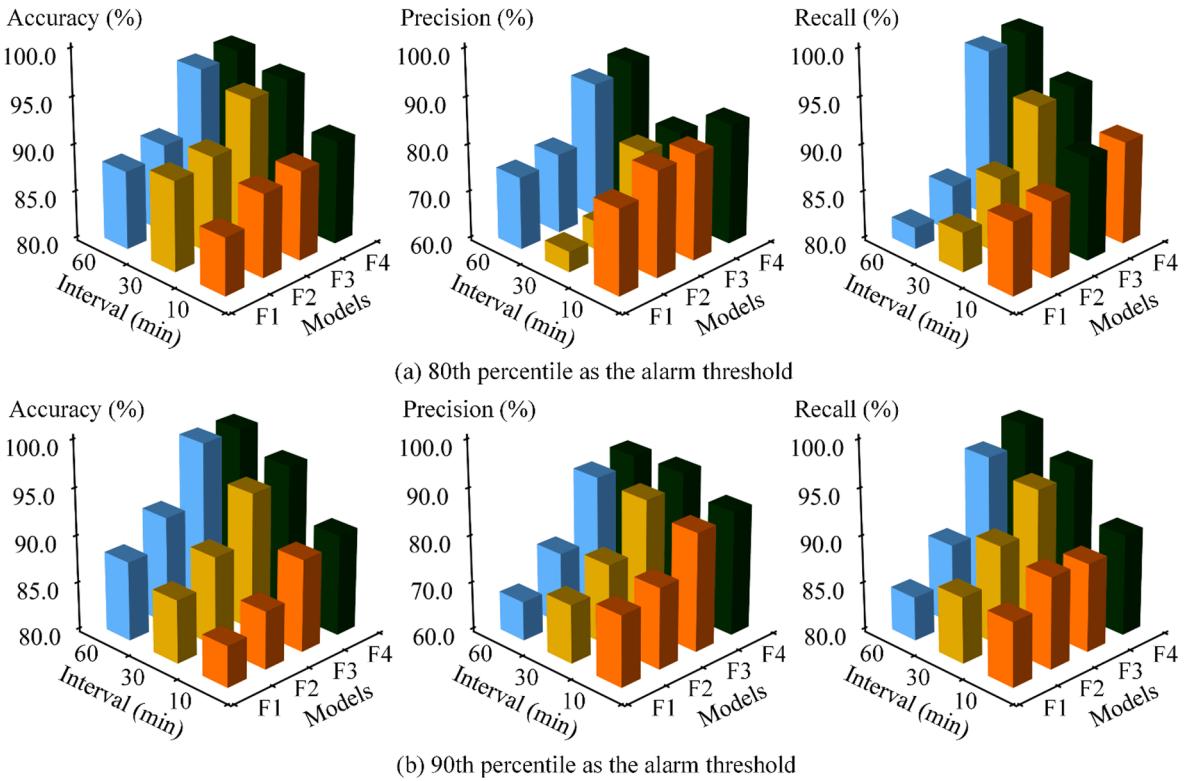
First, in general, by comparing the proposed SCLN-TTF method with benchmark methods, including HA, BP, GBRT, LN-FTF, SCLN-FTF and LN-TTF, it is found that the proposed algorithm improves the accuracy and stability for short-term congestion prediction of airways. This confirms the superiority and feasibility of the proposed model, which can successfully capture both the spatiotemporal features and influencing factors in short-term congestion prediction by considering both the spatial and temporal dependences and applying the bi-directional structure incorporating both the forward and backward sequence.

Second, most of the selected models exhibit the best performance for the 60 min prediction time horizon, while exhibiting the worst performance for the 10 min prediction time horizon. The proposed SCLN-TTF method shows its superiority especially in the 10 min prediction, indicating that the developed method can characterize well the very short-term characteristics of traffic congestion, thereby providing satisfactory prediction results.

Third, the spatiotemporal pattern of the congestion index shows that there are some common congestion periods and unbalanced flow in different airways. To identify the frequently congested airways or periods, the various influencing factors needs to be combined, together with the spatial and temporal dependencies. The performance of the high congestion index identification model can be increased by incorporating forward, current and backward features, and integrating the strength of GCN and XGBoost methods.

The present study is conducted according to the data from an approach sector (ZSNJAP01) of Nanjing Lukou International Airport as an example. The proposed method can also be applied in other cases as long as the airspace network, flight and metrological information can be provided. It should also be noted that the weather variables are considered as temporal features as data from the weather report are the same for the selected airspace. When the proposed method is applied in a larger area, the weather variables might be considered as spatiotemporal attributes. Other information associated with the airspace congestion state can also be incorporated into the neural network framework when data available.

Even though the proposed approach has exhibited great potential to short-term airway congestion prediction and identification, several limitations are still needed to be addressed in this study. First, it is acknowledged that various factors may have significant impact on airway congestion, especially meteorological conditions. This study is focused on incorporating the occurrence of general and extreme weather in airway congestion prediction. Future research is still needed to identify the impacts of other significant variables. Second, the paper used the data of ZSNJAP01 sector from Nanjing Lukou International Airport as a case study. Data from



**Fig. 18.** Performance comparison for congestion identification.

other airspace can also be applied to further investigate the robustness and applicability of the proposed model. The authors recommend that future studies could focus on these issues.

#### CRediT authorship contribution statement

**Zhao Yang:** Conceptualization, Methodology, Writing - original draft. **Rong Tang:** Data curation, Software, Visualization. **Weili Zeng:** Investigation. **Jiahuan Lu:** Writing - review & editing. **Zhijie Zhang:** Software, Validation.

#### Acknowledgement

This research was sponsored by the Fundamental Research Funds for the Central Universities of China (NS2020046) and National Natural Science Foundation of China (51608268, 62076126, 52002179).

#### References

- Adacher, L., Flamini, M., Romano, E., 2017. Rerouting algorithms solving the air traffic congestion. In: Presented at the Applied Mathematics and Computer Science: Proceedings of the 1st International Conference on Applied Mathematics and Computer Science, p. 020053. <https://doi.org/10.1063/1.4981993>.
- Beaton, E.K., Brennan, J.F., DeArmon, J.S., Formosa, J.J., Levin, K.M., Miller, S.L., Wanke, C.R., 2001. Predicting congestion in the Northeast U.S.: a search for indicators. Air Traffic Control Quarterly 9 (2), 99–131. <https://doi.org/10.2514/atcq.9.2.99>.
- Bilimoria, K., Lee, H., 2005. Analysis of aircraft clusters to measure sector-independent airspace congestion. In: AIAA 5th ATIO And16th Lighter-Than-Air Sys Tech. American Institute of Aeronautics and Astronautics, Arlington, Virginia. <https://doi.org/10.2514/6.2005-7455>.
- Cai, K.Q., Zhang, J., Xiao, M.M., Tang, K., Du, W.B., 2017. Simultaneous optimization of airspace congestion and flight delay in air traffic network flow management. IEEE Trans. Intell. Transp. Syst. 18 (11), 3072–3082. <https://doi.org/10.1109/TITS.2017.2673247>.
- Chen, J., Li, M., 2019. Chained predictions of flight delay using machine learning. In: AIAA Scitech 2019 Forum. Presented at the AIAA Scitech 2019 Forum. American Institute of Aeronautics and Astronautics, San Diego, California. <https://doi.org/10.2514/6.2019-1661>.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Presented at the KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>.
- Crucioli, L.L.B.V., Li, W., Barros, A.G., Koendjibiarie, M.W., 2015. Air holding problem solving with reinforcement learning to reduce airspace congestion. J. Adv. Transport. 49 (5), 616–633. <https://doi.org/10.1002/atr.1293>.
- Dai, X., Hu, M., Tian, W., Xie, D., Hu, B., 2016. Application of epidemiology model on complex networks in propagation dynamics of airspace congestion. PLoS ONE 11 (6), e0157945. <https://doi.org/10.1371/journal.pone.0157945>.
- Dong, Y., Lu, Z., Liu, Y., Zhang, Q., Wu, D., 2019. China's corridors-in-the-sky design and space-time congestion identification and the influence of air routes' traffic flow. J. Geog. Sci. 29 (12), 1999–2014. <https://doi.org/10.1007/s11442-019-1701-3>.

- Enayatollahi, F., Atashgah, M.A.A., 2018. Wind effect analysis on air traffic congestion in terminal area via cellular automata. *Aviation* 22 (3), 102–114. <https://doi.org/10.3846/aviation.2018.6252>.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Statist.* 29 (5), 1189–1232.
- Gilbo, E., Smith, S., 2011. Monitoring and alerting congestion at airports and sectors under uncertainty in traffic demand predictions. *Air Traffic Control Quarterly* 19 (2), 83–113. <https://doi.org/10.2514/atcq.19.2.83>.
- Goswami, P., Sarkar, S., 2017. An analogue dynamical model for forecasting fog-induced visibility: validation over Delhi: dynamical model for fog. *Meteorol. Appl.* 24 (3), 360–375. <https://doi.org/10.1002/met.1634>.
- Graves, A., 2014. Generating Sequences with Recurrent Neural Networks. arXiv:1308.0850 [cs].
- Graves, A., Schmidhuber, J., 2005. Framework phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* 18, 602–610. <https://doi.org/10.1016/j.neunet.2005.06.042>.
- Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J., 2017. LSTM: a search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* 28 (10), 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.
- Hammond, D.K., Vandergheynst, P., Gribonval, R., 2011. Wavelets on graphs via spectral graph theory. *Appl. Comput. Harmon. Anal.* 30 (2), 129–150. <https://doi.org/10.1016/j.jacha.2010.04.005>.
- Han, D., Chen, J., Sun, J., 2019. A parallel spatiotemporal deep learning network for highway traffic flow forecasting. *Int. J. Distrib. Sens. Netw.* 15 (2) <https://doi.org/10.1177/1550147719832792>.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Idris, H.R., Dao, Q., Rorie, R.C., Hashemi, K., 2020. A framework for assessment of autonomy challenges in air traffic management. In: Aiaa Aviation 2020 Forum. Presented at the Aiaa Aviation 2020 Forum, American Institute of Aeronautics and Astronautics, Virtual Event. <https://doi.org/10.2514/6.2020-3248>.
- Kipf, T.N., Welling, M., 2017. Semi-Supervised Classification with Graph Convolutional Networks. arXiv:1609.02907 [cs, stat].
- Kuhn, K.D., 2016. A methodology for identifying similar days in air traffic flow management initiative planning. *Transport. Res. Part C: Emerg. Technol.* 69, 1–15. <https://doi.org/10.1016/j.trc.2016.05.014>.
- Mulgund, S., Wanke, C., Greenbaum, D., Sood, N., 2006. A genetic algorithm approach to probabilistic airspace congestion management. In: AIAA Guidance, Navigation, and Control Conference and Exhibit. Presented at the AIAA Guidance, Navigation, and Control Conference and Exhibit, American Institute of Aeronautics and Astronautics, Keystone, Colorado. <https://doi.org/10.2514/6.2006-6226>.
- Ogimet, 2020. Metar/Speci/Taf reports selection query. Available from <http://ogimet.com/metars.shtml.en> [accessed 1 September 2019].
- Roh, H.J., 2020. Development and performance assessment of winter climate hazard models on traffic volume with four model structure types. *Nat. Hazard. Rev.* 21 (3), 04020023. [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000389](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000389).
- Sridhar, B., Chen, N.Y., 2009. Short-term national airspace system delay prediction using weather impacted traffic index. *J. Guid. Control Dyn.* 32 (2), 657–662. <https://doi.org/10.2514/1.38798>.
- Takeichi, N., Kaida, R., Shimomura, A., Yamauchi, T., 2017. Prediction of delay due to air traffic control by machine learning. In: AIAA Modeling and Simulation Technologies Conference. American Institute of Aeronautics and Astronautics, Grapevine, Texas. <https://doi.org/10.2514/6.2017-1323>.
- Taylor, C., Wanke, C., 2009. A multi-objective generalized random adaptive search procedure for resolving airspace congestion. In: AIAA Guidance, Navigation, and Control Conference. Presented at the AIAA Guidance, Navigation, and Control Conference, American Institute of Aeronautics and Astronautics, Chicago, Illinois. <https://doi.org/10.2514/6.2009-5864>.
- Vaabøen, B., Larsen, J., 2015. Mitigation of airspace congestion impact on airline networks. *J. Air Transp. Manage.* 47, 54–65. <https://doi.org/10.1016/j.jairtraman.2015.04.002>.
- Wanke, C., 2009. Continual, probabilistic airspace congestion management. In: AIAA Guidance, Navigation, and Control Conference. Presented at the AIAA Guidance, Navigation, and Control Conference. American Institute of Aeronautics and Astronautics, Chicago, Illinois. <https://doi.org/10.2514/6.2009-5863>.