

The Aeronautical Journal

Modeling sector route segment capacity under convective weather based on machine learning --Manuscript Draft--

Manuscript Number:	AeroJ-2024-0027R2
Full Title:	Modeling sector route segment capacity under convective weather based on machine learning
Article Type:	Research Article
Corresponding Author:	Baotian Yang Nanjing University of Aeronautics and Astronautics CHINA
Corresponding Author's Institution:	Nanjing University of Aeronautics and Astronautics
Order of Authors:	Shijin Wang Yinglin Li Baotian Yang Rongrong Duan
Order of Authors Secondary Information:	
Abstract:	Estimating airspace capacity under convective weather conditions is crucial for ensuring the efficiency and safety of air traffic operations. Sector route segments, which are essential components of flight routes, require timely capacity predictions during operationally critical periods. In this paper, initially, an enhanced Recursive Feature Elimination algorithm is used to select meteorological data and develop predictive features. Subsequently, the CWSRC model is established using the RF supervised learning algorithm. Finally, the paper takes ENH-YIH segment as an example to predict the capacity. Compared with other machine learning algorithms, the residual percentages for KNN, MLP, and RF are 86.03%, 77.37%, and 93.40%, respectively, within the range of [-0.2,0.2]. In three separate day cases, results show that the CWSRC model's MAE, MSE, RMSE, and R ² significantly outperform traditional methods like Maxflow/Mincut and scanning line. The results confirm the CWSRC model's superior predictive capabilities.
Response to Reviewers:	Dear Associate Editor Laurent Dala Thank you for your guidance and instructions. I have completed the final preparations for my manuscript as requested. Here is the list of files I have uploaded: The final PDF version of the paper. Individual JPEG/TIFF/EPS/PNG files for the figures that appear in the paper. All necessary source files for production, including LaTeX and WORD files. Please confirm the receipt of these files and review them. If there are any issues or further information is needed, please feel free to contact me. Thank you for your patience and support. Sincerely, Baotian Yang
Corresponding Author E-Mail:	2563765531@qq.com
Additional Information:	
Question	Response
Can you confirm that you have gained all	Yes

necessary copyrights and permissions to publish this manuscript legally	
Has this manuscript been submitted or published in any other Journal?	No - This manuscript has not been published in another Journal
Is this manuscript the work of all the authors you have listed and have all contributing authors to this manuscript been listed?	Yes
Would you be interested in receiving information about subscribing to the journal? If so, click this link.	
Would you be interested in receiving information about the benefits of becoming a member of The Royal Aeronautical Society?	
Are you submitting any supplementary material such as additional images, extensive data sets or filmed footage, which is not essential to the main argument of your article?	No
Enter the approximate word count of the manuscript. Please note, regular papers ideally should not exceed 10,000 words while survey papers should not exceed 15,000 words. Exceeding these limits may reduce the chances of your paper being approved for publication.	8000 words
Are you below the age of 30 and wish to be considered for the Young Person's Written Prize Award?	No
Select the Special Issue to which you are submitting your manuscript. If you are not submitting to a Special Issue, please select "Not part of a Special Issue."	Not part of a Special Issue
Is it your intention to have your paper published under the 'Gold' option Open Access initiative, for which there is a charge?	No
Author Comments:	

Modeling sector route segment capacity under convective weather based on machine learning

WANG shijin¹, LI yinglin¹, YANG baotian¹, DUAN rongrong¹

1. Next Generation Intelligent Air Traffic Control Laboratory, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China

Abstract

Estimating airspace capacity under convective weather conditions is crucial for ensuring the efficiency and safety of air traffic operations. Sector route segments, which are essential components of flight routes, require timely capacity predictions during operationally critical periods. In this paper, initially, an enhanced Recursive Feature Elimination algorithm is used to select meteorological data and develop predictive features. Subsequently, the CWSRC model is established using the RF supervised learning algorithm. Finally, the paper takes ENH-YIH segment as an example to predict the capacity. Compared with other machine learning algorithms, the residual percentages for KNN, MLP, and RF are 86.03%, 77.37%, and 93.40%, respectively, within the range of [-0.2,0.2]. In three separate day cases, results show that the CWSRC model's MAE, MSE, RMSE, and R² significantly outperform traditional methods like Maxflow/Mincut and scanning line. The results confirm the CWSRC model's superior predictive capabilities.

Nomenclature

ATM	air traffic management
BR	basic reflectivity[dBZ]
C	maximum flight number under clear weather of the sector route
C^p	capacity of a sector route segment during a specific period of a day
CR	composite reflectivity[dBZ]
CWAM	convective weather avoidance model
CWSRC	convective weather sector route capacity
E-WITI	en-route weather impacted traffic index
ET	echo tops[ft]
FAA	federal aviation administration
KNN	k-nearest neighbor
$Line_j$	convective weather duration of the flight on the j-th parallel line
MAE	mean absolute error
mcw	minimum available width on the route segment
MLP	multilayer perceptron
MSE	mean squared error
n	the number of grid cells
N_i^{xth}	the meteorological product value at the x-th percentile of the i-th airspace filter
NWS	national weather service
RF	random forest
RFE	recursive feature elimination
RMSE	root mean squared error
TMI	traffic management initiatives
T-WITI	terminal weather impacted traffic index
VIL	vertically integrated liquid[kg/m ²]
WITI	weather impacted traffic index

w_{ij}	the convective weather weight of the i -th grid of the j -th parallel line on the segment
W_r	the available proportion of traffic flow within the sector route segment
x^{th}	x -th percentile value of this airspace weather product
SRA_r, SRA_p	actual/ predicting airspace availability of sector route segment
SRC_r, SRC_p	actual/ predicting airspace capacity of sector route segment
Z	meteorological radar reflectivity factor
Δh	the height increment between successive radar measurements in meters.

0 Introduction

In the air transportation system, weather-related irregularities accounted for 59.56% and 66.5% of the total irregular flights in China^[1] and the United States^[2], respectively. Convective weather significantly reduces available airspace capacity, posing a major constraint on the normal operation of air traffic flow. Currently, traffic management processes heavily rely on the subjective interpretation of weather forecasts. Air traffic controllers use mental calculations to estimate reductions in airspace capacity and corresponding traffic flow rates. This approach fails to provide a clear translation of weather forecasts into tangible resource constraints. The air route is a fundamental component of the airspace system and crucial for the smooth operation of air traffic. Future ATM systems will require decision support tools. These tools must be capable of translating the impact of convective weather into precise air route capacities. Traffic demand for impacted air routes is managed through TMI. These initiatives either completely remove demand from an impacted air route or reduce demand by delaying the departure of flights filed through the impacted air route. Thus in regions of congested airspace, it is very important to be able to correctly estimate the air route capacity degradation due to convective weather effects so that avoidable delays and costs are minimized.

Many researchers have investigated the ability to translate a given weather situation into an estimate of available airspace capacity in recent years. The Lincoln Laboratory has made significant contributions in this area by developing three convective weather avoidance models, namely CWAM1, CWAM2, and CWAM3^[3-4]. They have also established a permeability calculation method and implemented the computation of airspace capacity using scanning line and Maxflow/Minicut methods^[5-6]. Additionally, the FAA developed the WITI model. This model uses historical meteorological data and traffic demand to accurately quantify the relationship between air traffic and weather. It predicts weather impacts on both E-WITI and T-WITI airspace^[6-8]. These methods rely on evaluating the severity of convective weather to estimate airspace capacity. However, the capacity of the air traffic management system is not solely dependent on convective weather conditions but is also influenced by various other factors, including route structure and the complexity of traffic flow operations. Simulation systems like SIMMOD^[9] by the FAA, TAAM^[10] by Boeing, RAMS^[11] in Europe, and ASMES^[12] in China have been developed to simulate and address the aforementioned issues. However, the complexity of the air traffic system requires sophisticated traffic flow simulation platforms and high computational performance, or complex mathematical models to accurately replicate the airspace operational environment for capacity evaluation. Moreover, these simulation systems are challenging to directly serve as decision support tools.

Recent advancements in computer data storage and computational capabilities, coupled with rapid developments in machine learning, have led to widespread applications in the civil aviation industry. This includes airport delay prediction, fleet planning, runway configuration forecasting, and route flight time estimation. Currently, machine learning is primarily focused on capacity prediction under convective weather conditions in the terminal area. Various studies have been conducted to address this area of interest. For instance, Mao et al. employed the random forest algorithm to predict the capacity of the terminal airspace in the presence of convective weather conditions^[13]. Chen et al. introduced the Spatio-Temporal Graph Convolution Network via Initial residual to effectively forecast the capacity of the terminal airspace^[14]. Wang et al., on the other hand, utilized a combination of multiple linear regression and regression tree ensemble learning methods to model and estimate the EWR arrival meter fix throughput^[15]. Additionally, Brito et al. developed a supervised machine learning-based prediction model to accurately forecast the elapsed time from a flight's takeoff to the sector crossing^[16]. However, it should be noted that the evaluation of capacity in the terminal airspace, employing machine learning algorithms, may not be directly applicable to the assessment of route resources. This discrepancy arises due to distinct differences in the structural characteristics, flight operation altitudes, control methods, and unique aspects associated with convective weather conditions between the terminal airspace and the route resources.

To quantify the impact of convective weather on airspace obstruction and accurately evaluate the capacity of air routes within the sector route segment, this paper establishes the CWSRC model. CWSRC model is based on the concept of sector

route blockage. It utilizes machine learning algorithms to predict airspace availability in the sector route during convective weather conditions, providing a consistent basis for collaborative decision-making in traffic flow operations. The structure of this paper is organized as follows: Section 1 introduces the fundamental concepts of sector route segment and the features used in our model. Section 2 describes the machine learning algorithms employed by the CWSRC model, detailing the methods used for training the model and the development processes. Section 3 analyzes the performance of three machine learning algorithms and evaluates the effectiveness of the CWSRC model. Section 4 summarizes the research findings and outlines future research directions.

1. Sector route segment and weather features

1.1 Sector route segment

An air route consists of a series of waypoints connecting multiple airspaces, typically long and complex. For instance, as depicted in Fig. 1, the routes include the light blue line A461, the dark blue line B334, the red line B339, and the green line G212. When convective weather exists in one location of a route, it does not affect the capacity of other segments farther away from that location. Therefore, it is neither accurate nor necessary to predict the capacity of a whole air route. Sector route segment refers to the segment between route points in a sector with a specific width, such as the part circled in yellow in Fig. 1. Flights usually fly according to the segments in the flight plan route, and a flight plan route contains a number of Sector route segments located in different routes. The white solid line on the left side of Fig. 1 indicates the flight plan route of the flight from ZHHH (Wuhan Tianhe International Airport) to ZYHB (Harbin Taiping International Airport), which passes through 4 routes and 11 sectors. This paper is concerned with predicting the capacity of the sector route segments.

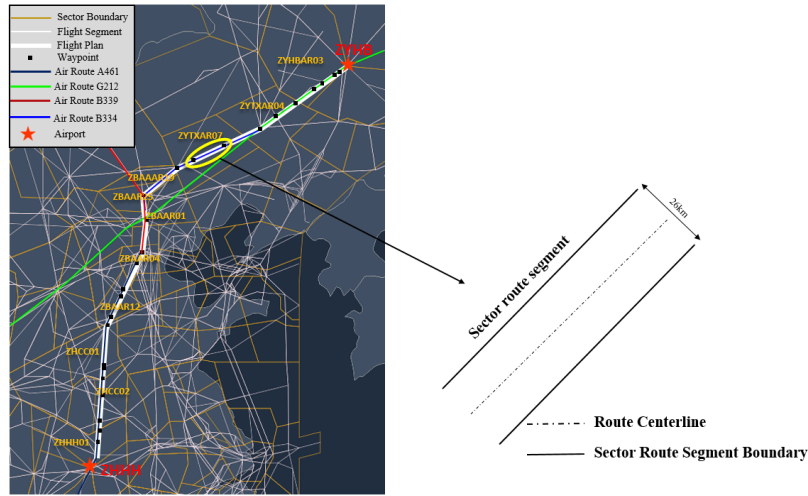


Figure 1. Schematic diagram of sector route segment

When a flight follows its planned route, deviations may occur due to factors such as weather, air traffic control decisions, and pilot operations. These deviations, while possibly altering the flight path slightly, do not constitute a rerouting. In order to accurately assess the capacity of the sector route segment, according to reference^[17], this paper takes the flight route width of 26km, as shown on the right side of Fig. 1, i.e., 13km on each side of the center-line of the flight route.

1.2 Meteorological features

Meteorological radar reflectivity Z , which reflects the strength of the radar echo signal and measures the intensity of precipitation and the physical properties of clouds, is updated every 6 minutes. In this paper, the information of meteorological features is extracted based on the derivatives of Z with the same refresh rate of 6 minutes. In addition, this paper uses the airspace filter method in extracting the percentile values of the meteorological products. The extraction process is shown in Fig. 2 and Equation. 1. The airspace filter is a square array, as shown on the right side of Fig. 2, which

consists of $n*n$ grids, each of which stores the weather product values in the latitude and longitude $0.01^\circ*0.01^\circ$ area. In this paper, a $26\text{km}*26\text{km}$ airspace filter is used, starting from one end of the sector route segment and divided based on the size of the airspace filter. There is no overlap between each airspace filter. If the remaining length of the route segment is not insufficient to form a complete filter, the grids adjacent to the segment are automatically filled to generate a complete airspace filter. The sector route segment area is divided into a number of airspace filters, extracting the x -th percentile value from each airspace filter. The x -th percentile value of this airspace weather product is obtained by computing the maximum value among all the filters within the airspace, as shown in Equation. 1.

$$x^{th} = \text{Max}\{N_1^{xth}, N_2^{xth}, \dots, N_i^{xth}\}, x \in [50, 75, 90, 100] \dots\dots\dots(1)$$

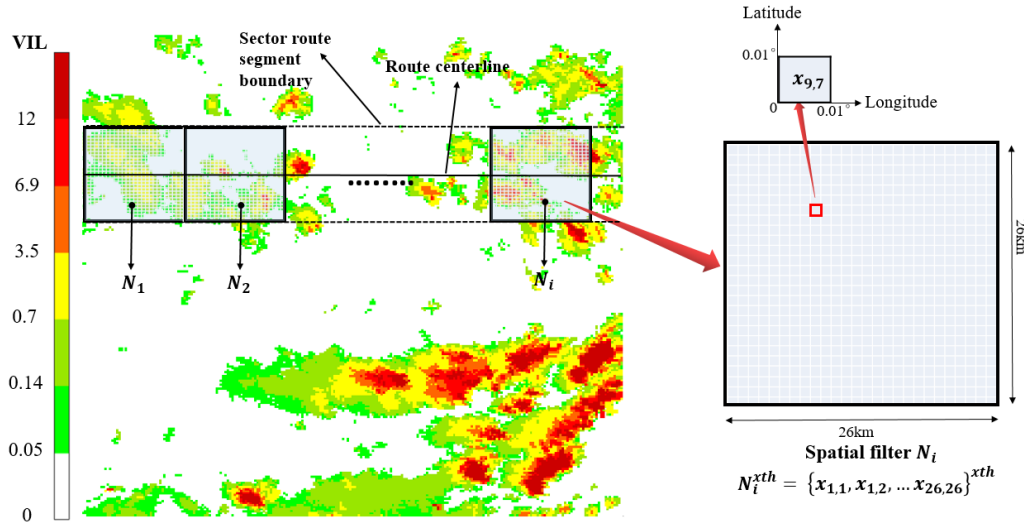


Figure 2. Airspace filter

where N_i^{xth} denotes the meteorological product value at the x -th percentile of the i -th airspace filter.

The NWS classifies convective weather as shown in Table 1^[18]. When the value of meteorological products CR, VIL and ET in the airspace exceeds Level 3 (including Level 3), it is regarded as severe weather, and flights are not allowed to fly over the airspace. Therefore, this paper selects CR, ET and VIL as weather features to predict sector segment capacity.

Table 1 Convective weather classification

Category	CR(dBZ)	VIL(kg/m ²)	ET(ft)	Level
None	<18	0-0.14		0
Light	18-30	0.14-0.7	<25,000	1
Middle	30-41	0.7-3.5		2
Heavy	41-46	3.5-6.9		3
Very heavy	46-50	6.9-12	25,000~35,000	4
Intense	50 -57	12-32		5
Extreme	≥57	≥32	>35,000	6

1.2.1 VIL

VIL is used to describe the amount of liquid water accumulated in the entire atmospheric column from the altitude of the radar to the top of the atmosphere. It is calculated as shown in Equation 2. Higher VIL values usually indicate the presence of a significant amount of liquid water within a thunderstorm, often foreshadowing the occurrence of severe convective weather such as heavy rainfall, hail, and strong winds.

$$VIL = \sum 3.44 \times 10^{-6} \left(\frac{Z_i + Z_{i+1}}{2} \right)^{4/7} \Delta h \quad (2)$$

where Z_i is the radar reflectivity and Δh is the height expressed in meters.

In this paper, the VIL data are stored in mat files, and each element in the matrix represents the VIL value of $0.01^\circ \times 0.01^\circ$ airspace cell at the specified location. In order to express the severity of convective weather in the airspace, this paper has counted the VIL percentile value of the airspace for the specified segment. Furthermore, in order to measure the coverage of convective weather in the airspace, this paper has counted the proportion of the number of $VIL \geq 3.5 \text{ kg/m}^2$ rasters in the airspace of the specified segment. Therefore the VIL features used in this paper are:

1. 50^{th} VIL: the 50th percentile VIL value in the airspace of the segment;
2. 75^{th} VIL : the 75th percentile VIL value in the airspace of the segment;
3. 90^{th} VIL : the 90th percentile VIL value in the airspace of the segment;
4. $maxVIL$: the maximum VIL value in the airspace of the segment;
5. *Proportion of $VIL \geq 3.5 \text{ kg/m}^2$* : the number of grids with $VIL \geq 3.5 \text{ kg/m}^2$ as a percentage of all grids in the airspace of the segment.

1.2.2 ET

ET is the height corresponding to an echo of 18.3 dBZ detected by meteorological radar, as shown in Equation 3. It is used to measure the intensity and vertical development of convective activity. Higher ET values usually imply higher convective cloud tops, indicating that severe convective weather might be developing or has already developed. The ET features used in this paper include:

6. 50^{th} ET: the 50th percentile ET value in the airspace of the segment;
7. 75^{th} ET: the 75th percentile ET value in the airspace of the segment;
8. 90^{th} ET: the 90th percentile ET value in the airspace of the segment;
9. $maxET$: the maximum ET value in the airspace of the segment.

$$ET = h_{BR > threshold} \quad (3)$$

1.2.3 CR

The BR after logarithmic transformation of Z can reflect the magnitude of Z value as in Equation 4.

$$BR = 10 \log(Z) \quad (4)$$

CR refers to a product derived within a single volume scan that projects the maximum reflectivity factor found in constant elevation azimuth scans onto Cartesian grid points. It represents the maximum reflectivity factor within a unit volume, as depicted in Equation 5.

$$CR = max BR \quad (5)$$

A larger value of CR represents a higher convective intensity.

10. 50^{th} CR: the 50th percentile CR value in the airspace of the segment;
11. 75^{th} CR: the 75th percentile CR value in the airspace of the segment;
12. 90^{th} CR: the 90th percentile CR value in the airspace of the segment;
13. $maxCR$: the maximum CR value in the airspace of the segment;
14. *Proportion of $CR \geq 41 \text{ dBZ}$* : the number of grids with $CR \geq 41 \text{ dBZ}$ as a percentage of all grids in the airspace of the segment.

1.2.4 Longitudinal weather features of segment

This article analyzes longitudinal and lateral convective weather conditions in sector route segment to identify their impact on air traffic flow.

The impact of longitudinal weather in sector route segment on air traffic flow is reflected in the length of the path of a flight crossing the convective weather portion, indicating the severity of longitudinal convective weather in sector route segment. Figure. 3 presents the method for assessing longitudinal convective weather in sector route segment.

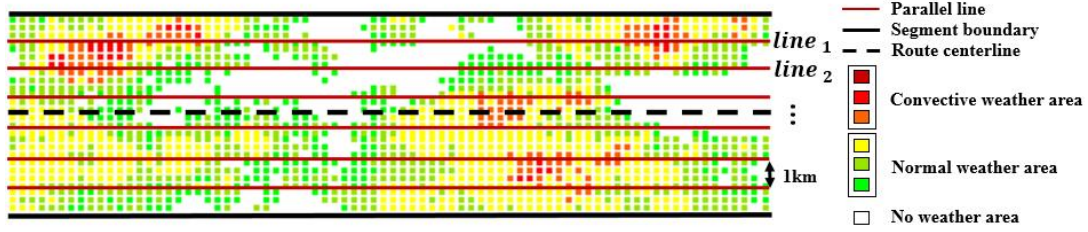


Figure 3. Schematic of longitudinal weather feature

Parallel lines are drawn along the direction of the sector route, spaced at intervals of 1 km parallel to the centerline. The VIL/CR values of all grids along each parallel line are calculated. When the VIL/CR value exceeds level 3 weather, the raster is considered to be in convective weather, and the convective weather weight value of the raster is set to be 1; Otherwise, the weight value of the raster is VIL/3.5 or CR/40. The sum of the convective weather weights of all the rasters on each parallel line indicates the longitudinal weather severity of the line as shown in Equation. 6.

$$line_j = \sum_{i=1}^n w_{ij}, w_{ij} = \begin{cases} 1, & \text{if } VIL \geq 3.5 \text{ or } CR \geq 40 \\ \frac{value}{3.5} \text{ or } \frac{value}{40}, & \text{if } VIL < 3.5 \text{ or } CR < 40 \end{cases} \quad (6)$$

Where w_{ij} denotes the convective weather weight of the i^{th} grid of the j^{th} parallel line on the segment. $Line_j$ denotes the convective weather duration of the flight on the j^{th} parallel line, and value denotes the VIL or CR value of each grid on the parallel line in this function.

The longitudinal route segment weather features are:

15/16. $VIL_parallel_mean/CR_parallel_mean$: The mean value of the impact along each parallel line represents the longitudinal average impact (based on VIL/based on CR);

17/18. $VIL_parallel_max/CR_parallel_max$: The maximum value of impact among all parallel lines (based on VIL/based on CR);

19/20. $VIL_parallel_min/CR_parallel_min$: The minimum value of impact among all parallel lines (based on VIL/based on CR);

21/22. $VIL_parallel_median/CR_parallel_median$: The median value of impact among all parallel lines (based on VIL/based on CR).

1.2.5 Lateral weather features of segment

When convective weather does not entirely cover the cross-section of a sector route segment, flights along the route can still traverse areas not affected by the weather. The lateral impact of convective weather on the sector route segment manifests in the degree of obstruction in the direction of the route's width, quantifying the proportion of the route's width covered by convective weather in the vertical direction, as indicated in Equation. 7.

$$W_r = \frac{mcw}{width} \quad (7)$$

where mcw denotes the minimum available width on the route segment, $width$ denotes the route segment width, and W_r represents the available proportion of traffic flow within the sector route segment.

The determination of the minimum available width of the segment, as illustrated in Figure. 2-4, begins by identifying the grid cells within the sector route affected by convective weather; Subsequently, the DBSCAN algorithm is employed to cluster multiple convective weather blocks within the route, depicting the progression from Figure. 4(a) to Figure. 4(b); The clustered convective weather blocks are utilized with a convex hull algorithm to generate a convex hull region, denoting the area impassable for aircraft, depicting the process from Figure. 4(b) to Figure. 4(c).

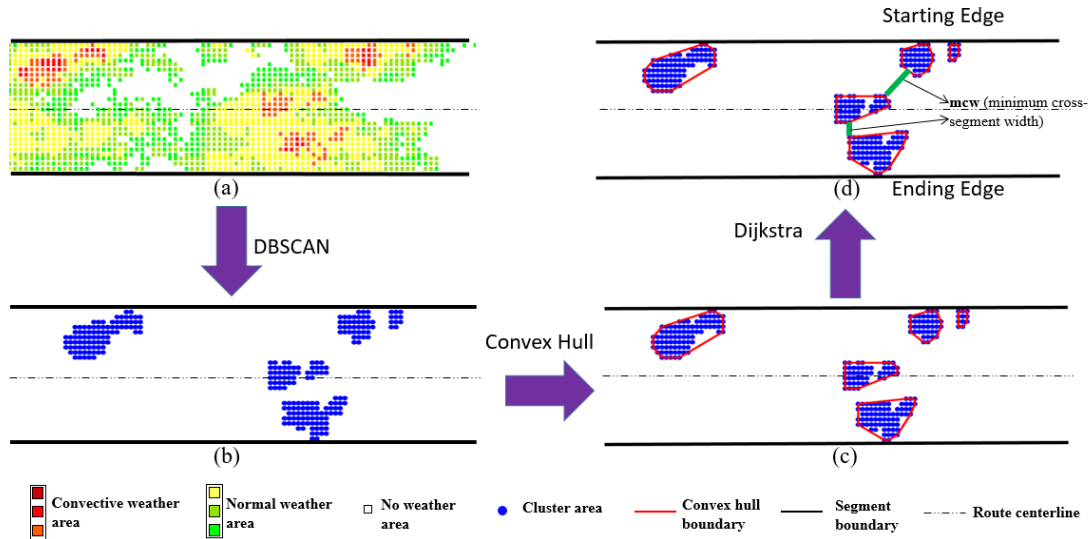


Figure 4. Flow percentage calculation process diagram for crossable sector route segments

Finally, the Dijkstra algorithm is applied to determine the *mcw* in the vertical direction within the sector route. The lateral features of the flight path in this section are included:

23. *Cross_flow_VIL*: the proportion of traffic flow that can traverse a route, calculated based on VIL;
24. *Cross_flow_CR*: the proportion of traffic flow that can traverse a route, calculated based on CR.

2 Methodology

Machine learning demonstrates excellent accuracy and generalization performance. This paper employs the CWSRC model based on the random forest algorithm for regression prediction of sector route segment capacity. The structure of the model, as depicted in Figure.5, involves meteorological product data as input and forecasts sector route segment capacity as output. This model is mainly divided into three parts: data preparation, model training and model testing.



Figure 5. Structural diagram of the CWSRC model

For the purposes of this study, airspace capacity is defined as the maximum number of aircraft that a specific unit of airspace can accommodate within a given time period. There could be a significant difference between the actual number of flights and the number of flights that can be accommodated under clear weather. As convective weather intensifies, airspace capacity decreases. To minimize flight delays and enhance operational efficiency, it is essential to match the number of flights with the capacity as closely as possible. Therefore, this paper assumes that the actual number of flights within a sector route segment during convective weather represents the segment's capacity under those conditions. Similarly, under clear weather, the maximum number of flights within an hour is considered the actual operational capacity of the sector route segment.

2.1 Data Set

The CWSRC model requires the use three meteorological product data types: VIL, ET, and CR. It extracts meteorological feature data based on sector route segment boundary information. To predict segment capacities based on varying meteorological features, the CWSRC model utilizes a dataset of historical flight radar trajectories through the sector route segment. Each flight trajectory in the dataset records the latitude and longitude coordinate position information of the flight at the corresponding time point, including the time of entering and exiting the sector route segment.

To validate the predictive performance of this model, the data is divided into two parts. The first part includes the training and validation datasets used for cross-validation to refine and validate the model. The second part comprises the test data, which is input into the trained model to analyze its performance.

2.2 Model training

To ensure robust predictive performance of the CWSRC model, we established label definitions during the model training process, performed feature selection, and fine-tuned parameters for the RF algorithm.

2.2.1 Label

This paper uses airspace availability as the label, airspace availability denoted as SRA_r , which is calculated as shown in Equation 8.

$$SRA_r = \frac{SRC_r}{C} \quad (8)$$

In Equation 8, SRC_r and C represents the actual flight number under convective weather and the maximum flight number under clear weather of the sector route in the corresponding period. The number of flights in sector route segment varies significantly across different time intervals, typically with fewer flights in the early hours and a higher volume during daylight hours. However, there exists a degree of similarity in the number of flights during equivalent time segments on different days. This study involves the calculation of sector route segment capacities for each hour within a day over a specific time span (usually a month), as derived in Equation 9.

$$C^p = 90^{\text{th}}\{C_1^p, C_2^p, \dots, C_i^p\} \quad (9)$$

where C^p represents the hourly capacity of a sector route segment during a specific period of a day, and C_i^p represents signifies the sector route segment's hourly capacity for a particular day within the statistical time span. Within a given time span, there might be instances where a day's sector route segment hourly capacity exceeds its limit. We take the 90th percentile segment capacity value to represent the maximum accommodatable flight volume for that specific hour within the sector route segment. The availability label is the current hourly flight traffic divided by the current hourly capacity.

In the later part, SRA_p and SRC_p are the predicted value of availability and capacity, respectively.

2.2.2 Model evaluation metrics

In the field of machine learning, a common metric for assessing algorithm accuracy is the generalization error. However, directly calculating the generalization error is not practical in applications due to the unknown distribution of the sample data. Therefore, we utilize a validation set constructed from sample data and compute various performance metrics to estimate the generalization performance of the RF algorithm.

This study employs four commonly used regression prediction metrics, including MSE, RMSE, MAE, and R^2 . The calculations are depicted in Equations 10,11,12,13.

$$MSE = \frac{\sum_{i=1}^n (y_i - f_i)^2}{n} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - f_i|}{n} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}} \quad (12)$$

$$R^2 = 1 - \frac{\sum (y_i - f_i)^2}{\sum (y_i - \bar{y})^2} \quad (13)$$

MSE measures the mean of the squares of the prediction errors. A smaller MSE value indicates that the model's predictions are closer to the actual values, signifying better model performance. RMSE is the square root of MSE. It provides the standard deviation of the prediction errors. RMSE offers an intuitive feel for the size of the errors; compared to MSE, RMSE has the same units as the original data, making it easier to understand. MAE provides the average size of the prediction errors and, unlike MSE and RMSE, does not assign higher weight to larger errors. R^2 is a metric that measures the model's ability to explain variability in the data. An R^2 value closer to 1 indicates that the model explains a higher variability, leading to better predictive performance.

2.2.3 Model training methods

This study utilizes K-fold cross-validation method to assess the model's generalization performance and subsequently adjust the RF parameters. Cross-validation can effectively mitigate the risk of overfitting and accurately reflects the model's generalization capabilities. The RF design in the CWSRC model is illustrated in Algorithm 1. Within the K-fold cross-validation method, the original dataset is randomly partitioned into k equally sized subsets. Each iteration selects one subset the model's validation dataset while using the remaining $k-1$ subsets as the training data. This process is repeated k times, rotating through each subset as the validation set, until all subsets have been used for validation. The metrics results obtained from each training validation are recorded, and the average of these k results yields the overall training performance for a particular set of hyperparameters. This iterative training and testing process makes the model performance assessment more reliable and reduces the sensitivity to the data partitioning method.

Algorithm1

Algorithm: RF

Input: Data set S

$$S = S_1 \cup S_2 \cup \dots \cup S_K, \quad S_i \cap S_j = \emptyset \quad (\forall i \neq j)$$

parameters: the number of decision trees T

Process:

- 1: While $k \leq K$ do
- 2: $S_{train} \leftarrow$ training set, $S - S_k$
- 3: $S_{test} \leftarrow$ test set, S_k
- 4: $N \leftarrow$ S_{train} size
- 5: $n \leftarrow$ S_{test} size
- 6: $T \leftarrow$ the number of decision trees
- 7: while $t \leq T$ do
- 8: $D_t \leftarrow$ Bootstrapping(S_{train}, N)
- 9: $f_t \leftarrow$ DTree(D_t)
- 10: end while
- 11: $f \leftarrow \frac{1}{T} (\sum_{t=1}^T f_t(S_{test}))$
- 12: $RMSE_k \leftarrow \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2}$
- 13: $MSE_k \leftarrow \frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2$
- 14: $MAE_k \leftarrow \frac{1}{n} \sum_{i=1}^n |y_i - f_i|$
- 15: $R_k^2 \leftarrow 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
- 16: end while

$$17: RMSE \leftarrow \frac{\sum_{k=1}^K RMSE_k}{K}$$

$$18: MSE \leftarrow \frac{\sum_{k=1}^K MSE_k}{K}$$

$$19: MAE \leftarrow \frac{\sum_{k=1}^K MAE_k}{K}$$

$$20: R^2 \leftarrow \frac{\sum_{k=1}^K R^2}{K}$$

2.2.4 Feature selection

This investigation incorporates 24 comprehensive features, which, while enriching the dataset, also extend the data collection phase and augment the model's complexity. Effective feature selection is pivotal, as it not only simplifies the model by eliminating features with high intercorrelations but also enhances its generalization capabilities and abbreviates the training duration. Conventionally, RFE operates by sequentially removing the least significant feature. This procedure, although systematic, may inadvertently undermine the model's performance, particularly when the excised feature plays a pivotal role in certain combinations.

To address these shortcomings, the present study proposes an advanced modification of RFE that integrates aspects of stepwise regression. Following the exclusion of each feature, a comprehensive assessment is conducted to evaluate its interaction with other features and overall impact on model performance. Should this assessment reveal that certain eliminated features substantially bolster the model's efficacy, these features are subsequently reinstated in the feature set during later iterations. This meticulous approach refines the feature selection process, culminating in an optimal feature subset that significantly enhances the model's accuracy and robustness. Algorithm 2 outlines the algorithmic process.

Algorithm2

Algorithm: Recursive Feature Elimination Algorithm

Input: feature x , label y , model $model$, number of target features $n_features_to_select$

Process:

```

1: Indicator = []
2: current_number_of_features = total number of features  $N$ 
3: function recursiveFeatureElimination( $x, y, model, n\_features\_to\_select$ ):

4:   while current_number_of_features >  $n\_features\_to\_select$ :
5:      $model.fit(x, y)$  // Train the model on current feature set
6:      $importance = model.feature\_importances\_$  // Calculate feature importances

7:     // Identify and remove the least important feature
8:      $least\_important\_index = \text{index of feature with } \min(importance)$ 
9:      $x = \text{remove feature at } least\_important\_index \text{ from } x$ 
10:     $current\_number\_of\_features = current\_number\_of\_features - 1$ 

11:    // Evaluate model performance
12:     $indicator = \text{evaluate\_model}(model, x, y)$  // Function to evaluate model performance
13:     $Indicator.append(indicator)$  // Record performance indicator

14:    // Record feature if performance worsens
15:     $delete\_feature = []$ 
16:    if  $\text{length}(Indicator) > 1$  and  $indicator$  worse than  $Indicator[-2]$ :
17:      record the removed feature as  $delete\_feature$ 

18:    // Recursive call for next iteration
19:    recursiveFeatureElimination( $x, y, model, n\_features\_to\_select$ )

20:   return subset of features:  $subset = [current\_feature, delete\_feature]$ 

21: End function

```

In Algorithm2, further stepwise explain are as follows:

- 1) Initialization (Lines 1-3): Initialize '*Indicator*' to store performance indicators and '*current_number_of_features*' to track feature count. Initialize the '*recursiveFeatureElimination*' function.

-
- 2) Loop and Training (Lines 4-5): Enter a loop to iteratively eliminate features until `'current_number_of_features'` reaches `'n_features_to_select'`. Train the model (`'model.fit'`) on the current feature set (`'x'` and `'y'`).
 - 3) Feature Importance Calculation (Line 6): Calculate feature importance using the model's `'feature_importances_'` attribute.
 - 4) Feature Removal (Lines 7-10): Identify and remove the least important feature from `'x'` based on calculated importance. Update `'current_number_of_features'` accordingly.
 - 5) Performance Evaluation (Lines 11-13): Evaluate model performance using the reduced feature set (`'x'` and `'y'`). Record the performance indicator (`'indicator'`) in `'Indicator'`.
 - 6) Comparison and Recording (Lines 14-17): Compare the current performance indicator with the last recorded one (`'Indicator[-2]'`). If performance worsens, record the removed feature as `'delete_feature'`.
 - 7) Recursive Call (Lines 18-19): Recursively call `'recursiveFeatureElimination'` to continue feature elimination with the updated feature set.
 - 8) Return (Line 20): Once `'current_number_of_features'` equals `'n_features_to_select'`, return the subset of features selected (`'subset = [current_feature, delete_feature]'`).

Begin by establishing the desired number of target features. The original features are then input into the RF algorithm, the results of the performance metrics obtained from this training will be recorded, and the importance of the features will be ranked according to the RF's feature importance evaluation function. Afterwards, selecting features based on the importance derived from random forest can identify the most influential features on the model's predictive outcomes. After eliminating the features with the least importance, proceed to the next round of training with the remaining features. Record any excluded features when the metrics from the subsequent training round show a decline compared to the previous results. The process is repeated until the number of features reaches the target number of features, and the recorded excluded features are added to the remaining features to obtain the selected subset of features after screening.

2.2.5 Hyperparameter determination

The number of decision trees is a crucial hyperparameter in the RF algorithm. RF enhances the generalization and stability of the model by aggregating the predictions of multiple decision trees. Each decision tree conducts random sampling on the training data and performs splits on randomly selected subsets of features, mitigating overfitting risks. Thus, while increasing the number of decision trees in a RF typically improves performance, there's a limit to the increase as adding trees indefinitely may not guarantee improved performance. Considering the cost of computational resources and time consumption, it's essential to establish a range of decision tree numbers. In this research, we follow precedents set in prior studies by setting the number of trees in the RF algorithm between 0 and 500^[19]. By traversing this range of parameters and observing the relationship between performance metrics and the number of decision trees through visualizations, we can determine the trend of performance metrics' variations and ascertain the optimal number of decision trees.

2.3 Model test

In this paper, we test the adjusted RF algorithm using untrained data. Meteorological features are used as inputs to predict the availability of sector route segment. In addition, based on the timeframe of the samples, we determined the corresponding capacity of sector route segments for each period. Multiplying airspace availability by period capacity yielded the sector route segment capacity predicted by the CWSRC model. Analyzing the differences between the predicted results on the test set and the actual values, and comparing them with traditional capacity prediction methods, allowed us to further evaluate the effectiveness of the CWSRC model's predictions.

2.4 Model stepwise

The development steps of the CWSRC model are as follows.

- 1) Label definition: Initially, sample labels are defined, with *SRA* being utilized as the primary label for this study.

- 2) Data Synthesis: The identified labels and corresponding features are integrated to form a comprehensive sample set.
- 3) Training and Validation Set Split: This sample set is subsequently partitioned into training-validation sets and a separate test set to facilitate robust model evaluation.
- 4) Performance Metrics Definition: At this stage, performance metrics such as MAE, MSE, RMSE, and R^2 are selected to assess the effectiveness of the model.
- 5) Cross Validation: The model is trained utilizing cross-validation techniques to enhance its generalizability.
- 6) Feature Selection: An enhanced RFE algorithm is employed to identify and retain the most impactful features.
- 7) Parameter Optimization: Following feature selection, the model's hyperparameters are finely tuned to optimize performance.
- 8) Execution on Test Data: The test set, segregated in the data segmentation phase, is employed to evaluate the optimally trained model.
- 9) Results Analysis: The outcomes of the model testing are analyzed. Key performance indicators are computed and compared against those derived from traditional methodologies.

3 Result and Discussion

This section uses the ENH-YIH sector route segment in the Wuhan control area of China as an example. Based on the methods detailed in Section 2, specific feature selection and parameter optimization are conducted to enhance model performance. The optimization results of the RF-based CWSRC model are compared with those of other machine learning algorithms in a conducted comparative analysis. This study is to assess the generalization ability of the model on new, untrained data and to compare the performance of the CWSRC model against traditional methods such as Maxflow/Mincut and the scanning line approach. MAE, MSE, RMSE and R^2 are used as metrics to evaluate our objectives in the test dataset.

3.1 Data preparation

In this paper, the historical radar trajectory data, flight plan data and meteorological data from August 3 to 31, 2018 of a busy sector route in Wuhan control area is used for verification of CWSRC model. The horizontal range of Wuhan control area is shown in Figure.6, and the vertical range is from 6,000 to 7,800 meters.

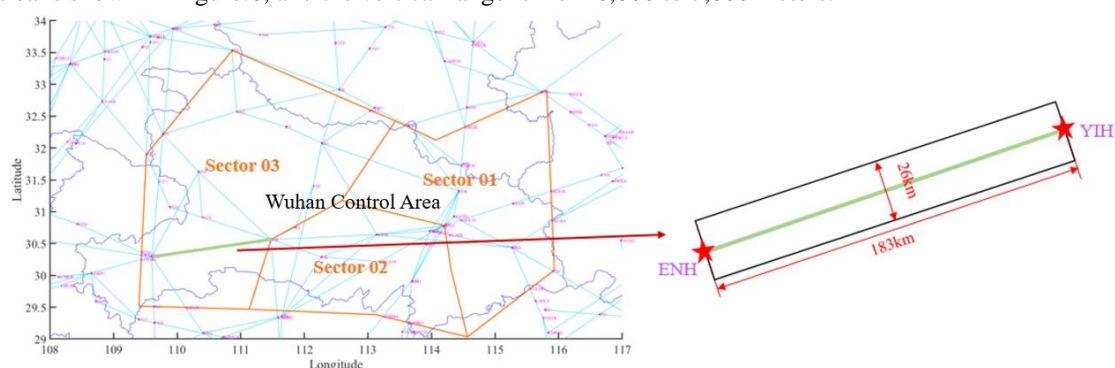


Figure 6. Position of the ENY-YIH

The sector route is ENH-YIH, located in sector 03 with a length of 183km in an east-west direction. It is the part of the sector route segment in the Shanghai-Chengdu direction and the flights accounts for about 30% of the total. Radar trajectory data is updated at a rate of 8 seconds, including the time, flight number, height, longitude, latitude and other information of historical flights. Longitude and latitude can be used to determine whether a trajectory exists within a sector area; time can be used to determine which time period the flight's traffic belongs to; and altitude is used to determine whether the flight is within the altitude range of the sector route segment. The meteorological data in the China's airspace

29°N-33°N and 109°E-116°E, VIL, CR and ET, is stored as $0.01^\circ \times 0.01^\circ$ (Longitude \times Latitude) and updated at a rate of 6 minutes.

Statistics of flights number in each period from August 3 to 31, 2018 are shown in Figure. 7. The time in the abscissa refers to one hour after this point in time. The busy periods of both two sector routes are from 8:00 to 23:00 each day, therefore historical flights data and weather data of these periods are selected in this paper. There are 4350 periods in the sector route. After screening, there are 3350 periods of ENH-YIH meeting convective weather conditions. We employed a ten-fold cross-validation approach, dividing the dataset into training and validation sets at a ratio of 9:1. Additionally, data from 46 time periods were utilized as the test set for this study.

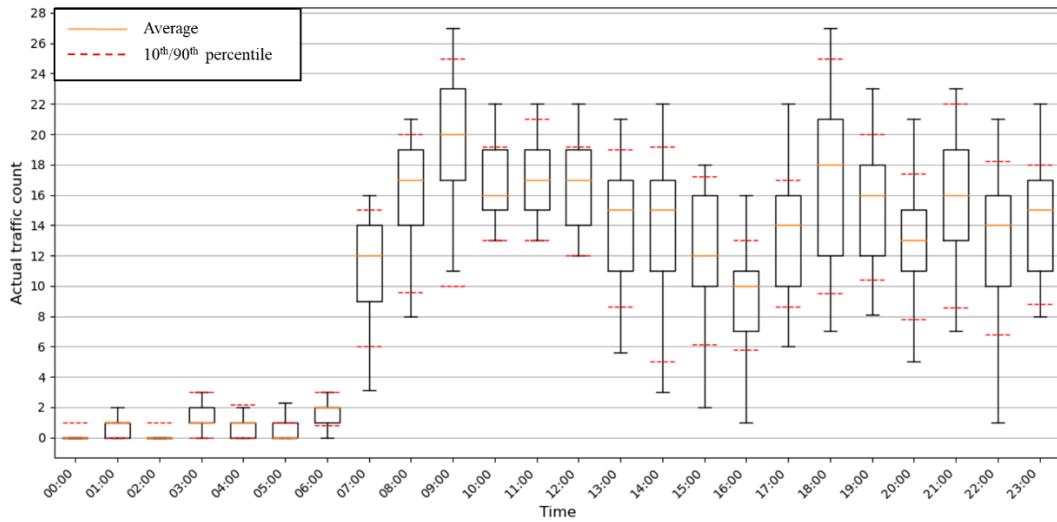


Figure 7. Distribution of operational capacity across time periods in the ENH-YIH

3.2 Model Optimization

The number of target features is set to 1, and the RF algorithm with default parameters is employed for recursive feature elimination. As the feature set iterated down to just one feature, the performance metrics curves for the entire iterative process were observed. Figure. 8 displays the results of the model's prediction for SRA using the metrics MAE, MSE, RMSE, and R^2 .

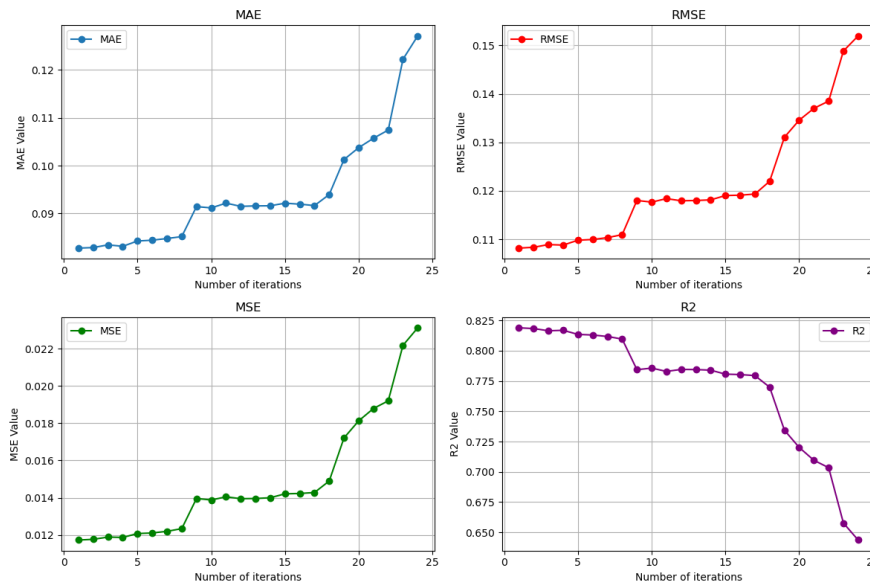


Figure 8. Changes in metrics with different number of iterations

Based on the performance metrics plotted in the aforementioned figure, Table 2 presents the impact of each iteration's feature elimination on the model's performance. An improvement in model performance upon feature elimination suggests

that the feature contributed more erroneous information, necessitating its removal. Conversely, when removal features make the model performance worse, it means that the feature is more effective. Therefore, the features causing a relative deterioration in model performance compared to the previous iteration are considered the optimal feature subset for elimination. From the table, the optimal feature subset excludes: *maxCR*, *50th VIL*, *Proportion of VIL \geq level 3*, *CR_parallel_mean*, *VIL_parallel_min*. The comparison between the model's performance using this feature subset and the original features is illustrated in Table 3. Although there is minimal change in MAE, MSE, RMSE, and R², the training prediction time decreased from 1'24" to 1'10" with the reduced feature set.

Table 2: Indicator performance after feature culling relative to the previous iteration (ρ : denotes the performance with the previous round after removing the feature)

Features	ρ	Features	ρ
<i>50th ET</i>	Worse	<i>Proportion of VIL \geq level 3</i>	Better
<i>75th ET</i>	Worse	<i>Proportion of CR \geq 41dBZ</i>	Worse
<i>90th ET</i>	Worse	<i>CR_parallel_mean</i>	Better
<i>maxET</i>	Worse	<i>CR_parallel_max</i>	Worse
<i>50th CR</i>	Worse	<i>CR_parallel_min</i>	Worse
<i>75th CR</i>	Worse	<i>CR_parallel_median</i>	Worse
<i>90th CR</i>	Worse	<i>VIL_parallel_mean</i>	Worse
<i>maxCR</i>	Better	<i>VIL_parallel_max</i>	Worse
<i>50th VIL</i>	Better	<i>VIL_parallel_min</i>	Better
<i>75th VIL</i>	Worse	<i>VIL_parallel_median</i>	Worse
<i>90th VIL</i>	Worse	<i>Cross_flow_CR</i>	Worse
<i>maxVIL</i>	Worse	<i>Cross_flow_VIL</i>	--

Table 3: The performance metrics of raw features versus filtered features

Features	MAE	MSE	RMSE	R ²	Time
Original	0.08276	0.01173	0.10817	0.81897	0:01:24.47
Filtering subsets	0.08289	0.01179	0.10846	0.81801	0:01:10.19

Table 4 presents the feature importance rankings within the RF model. Notably, the transverse and longitudinal weather characteristics of the sector rank highest in importance, validating our proposed approach that incorporating directional features of flights significantly enhances the prediction of sector route capacity.

Table 4: Importance of features

Feature	Importance	Feature	Importance
<i>Cross_flow_VIL</i>	23.89%	<i>90th ET</i>	3.37%
<i>VIL_parallel_median</i>	13.72%	<i>maxVIL</i>	3.36%
<i>VIL_parallel_mean</i>	8.03%	<i>50th ET</i>	3.15%
<i>CR_parallel_max</i>	5.87%	<i>75th CR</i>	2.69%
<i>CR_parallel_median</i>	5.31%	<i>90th CR</i>	2.54%
<i>50th CR</i>	4.74%	<i>Proportion of CR \geq 41dBZ</i>	2.43%
<i>CR_parallel_min</i>	3.88%	<i>90th VIL</i>	2.21%
<i>75th ET</i>	3.86%	<i>75th VIL</i>	1.98%
<i>maxET</i>	3.84%	<i>Cross_flow_CR</i>	1.64%
<i>VIL_parallel_max</i>	3.46%		

The selected subset of optimal features is input into the model with default parameters. The number of decision trees for RF is set from 1 to 500 to derive the model performance metrics at each training. As depicted in Figure. 9, after the number of trees is 100, there was no significant change observed in MAE, MSE, RMSE, and R², indicating convergence. However, the training time becomes longer as the number of trees increases. Considering the cost of time, this study settles on 100 trees for RF algorithm.

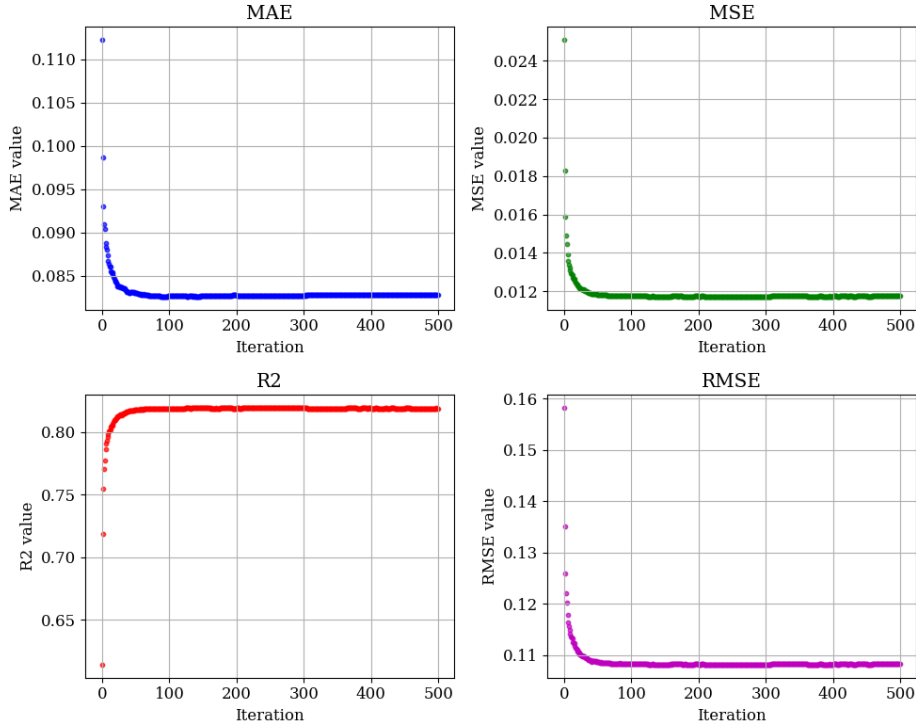


Figure 9. Different performance indicators based on tree numbers

3.3 Result analysis

To validate the model, this study excluded data from August 30 and utilized data from August 3 to August 31, 2018, for model training and testing. August 30, 2018, September 20, 2018, and September 25, 2018, were selected as case study days for model testing, as convective weather was detected in the sector route segment on these days. Due to the predominance of clear weather samples in the historical dataset, oversampling techniques were applied to the convective weather samples. The algorithm's performance was analyzed based on the model's results on the validation set. Furthermore, the model's performance on the test set was evaluated.

3.3.1 Algorithm analysis

In order to verify the superior performance of RF algorithm compared to other machine learning algorithms, this paper compares the results predicted by RF with those predicted by the KNN and MLP algorithms with default parameters. Among them, the parameters of machine learning algorithm are set as follows. The distance metric in KNN algorithm is Euclidean distance, and k is 10. The number of decision trees in RF is 100. There are two hidden layers in MLP, and the hidden layer units are 4 and 7, respectively. The activation function is *Relu*, solver is *adam*, the learning rate is 0.001, and the maximum number of iterations is 300.

Figure. 10 illustrates the residual distributions of SRA_p for KNN, MLP and RF on the cross-validated test set, where the interval of the statistics is $[-0.6, 0.6]$ with a step size of 0.05. Notably, the predominant residual distribution for RF lies within the $[-0.05, 0.05]$ range, exhibiting significantly higher sample counts compared to MLP and KNN within this interval. The number of samples with residuals between $[-0.3, 0.3]$ is 3309, 3106 and 3335 of KNN, MLP and RF respectively, accounting for 98.78%, 92.72% and 99.55% of the total data respectively. The number of samples with residuals between $[-0.2, 0.2]$ is 2882, 2592 and 3129 respectively, accounting for 86.03%, 77.37% and 93.403% of the total data respectively.

The number of samples with residuals between $[-0.1, 0.1]$ is 1637, 1435 and 2131 respectively, accounting for 48.87%, 42.84% and 63.61% of the total data respectively.

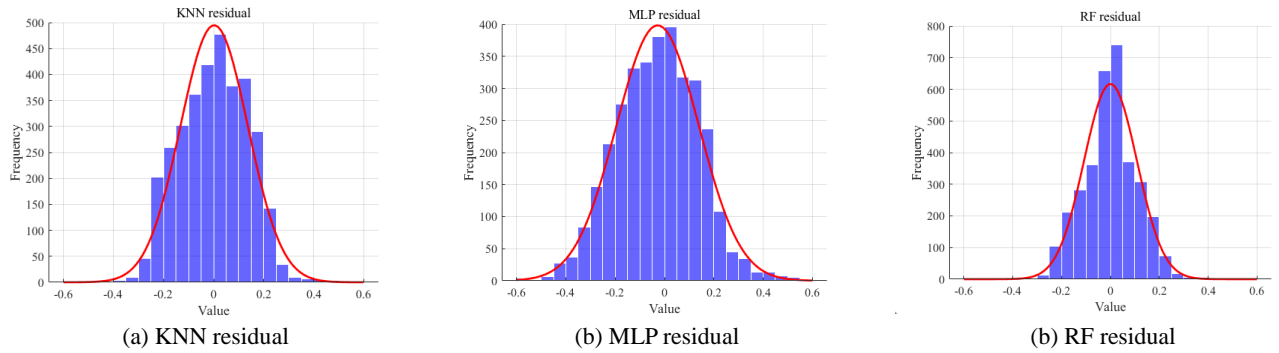


Figure 10. Residual plots of KNN, MLP, and RF algorithms

3.3.2 Analysis of SRC_p

In this section, the SRC_p of ENH-YIH obtained from the RF-based CWSRC model is analysed and compared with the results from Maxflow/Mincut and scanning lines. Considering that the weather conditions typically do not undergo drastically in a short period of time, this study represents the weather at half-hour intervals as representative of the weather for that period. Additionally, the weather conditions and flight trajectories for the prediction periods of the three case study days are depicted in Figures. 11, 12, and 13, respectively.

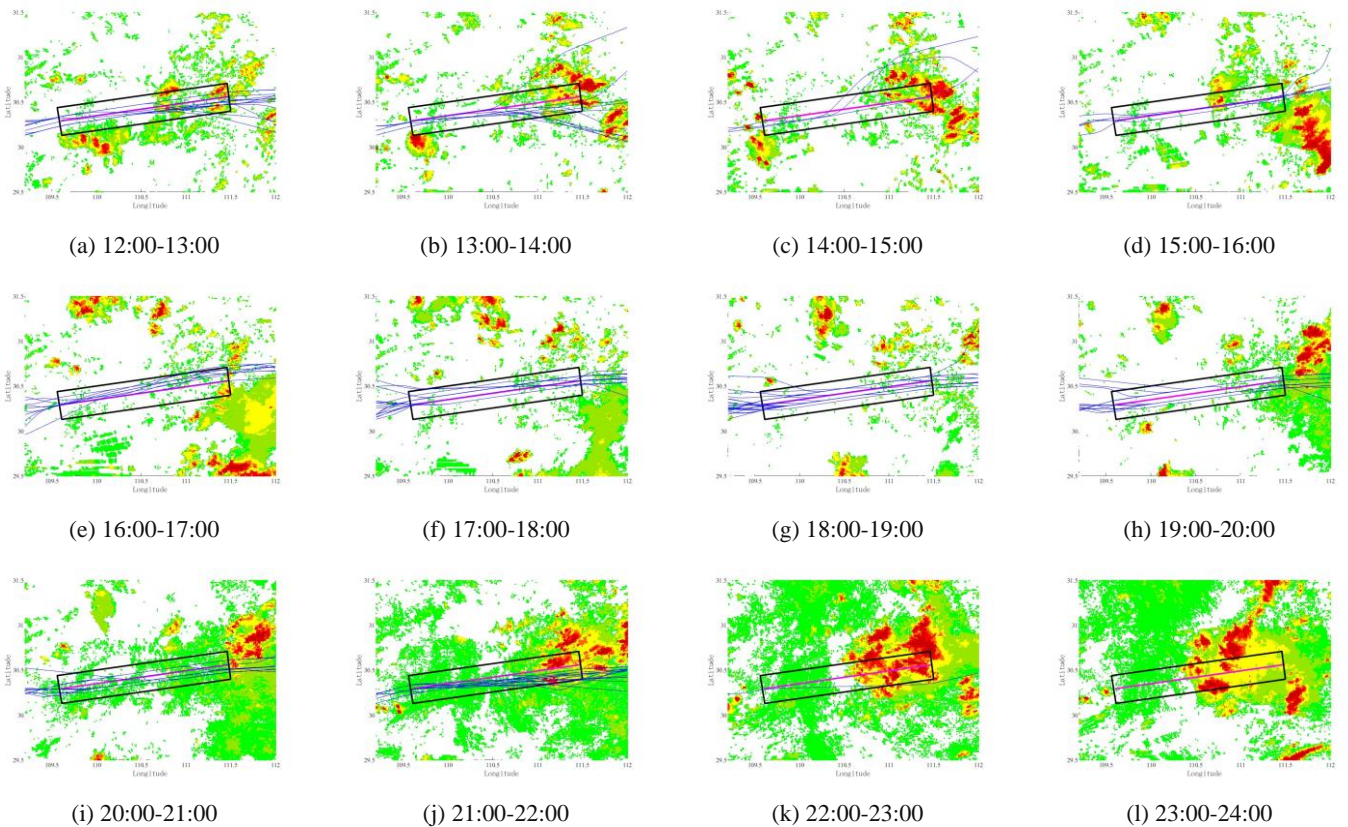


Figure 11. Flight tracks and weather on ENH-YIH from 12:00-24:00 on August 30, 2018

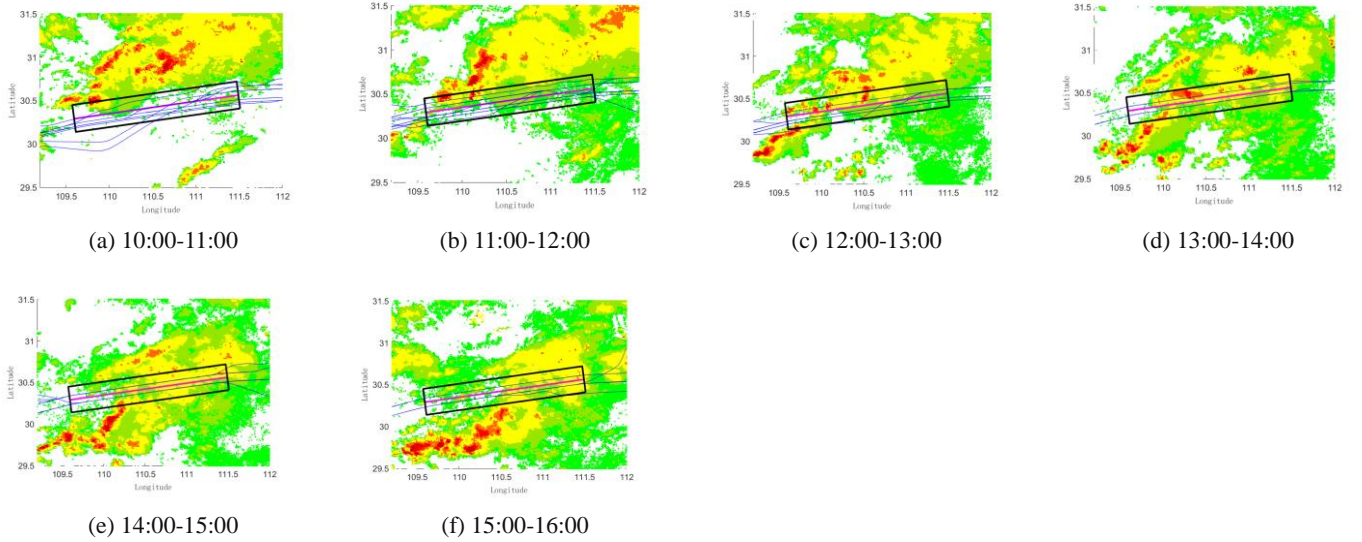


Figure 12. Flight tracks and weather on ENH-YIH from 10:00-16:00 on September 20, 2018

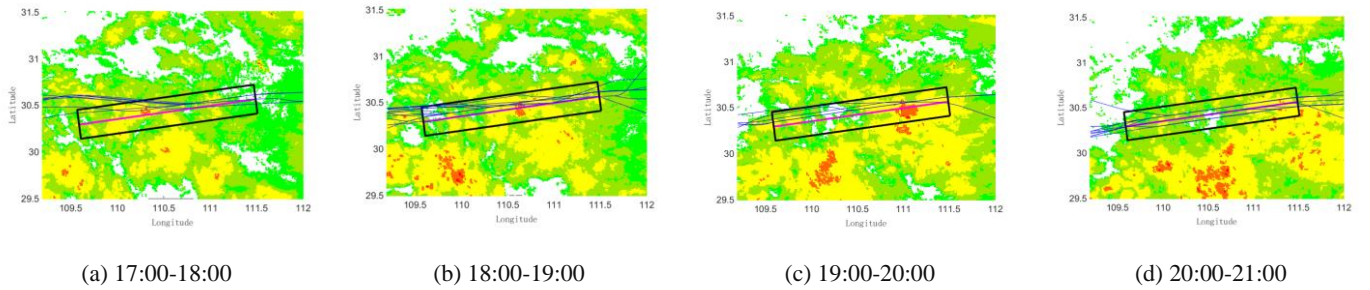


Figure 13. Flight tracks and weather on ENH-YIH from 17:00-21:00 on September 25, 2018

The prediction results for the three case study days are presented in Figure 14. Table 5 presents the overall performance of the models on the three case study days. We employed MAE, MSE, RMSE and R^2 as metrics to evaluate the predictive performance of the models. It is evident that the CWSRC model exhibits the lowest errors in metrics, indicating superior predictive performance compared to the Maxflow/Mincut and scanning line methods.

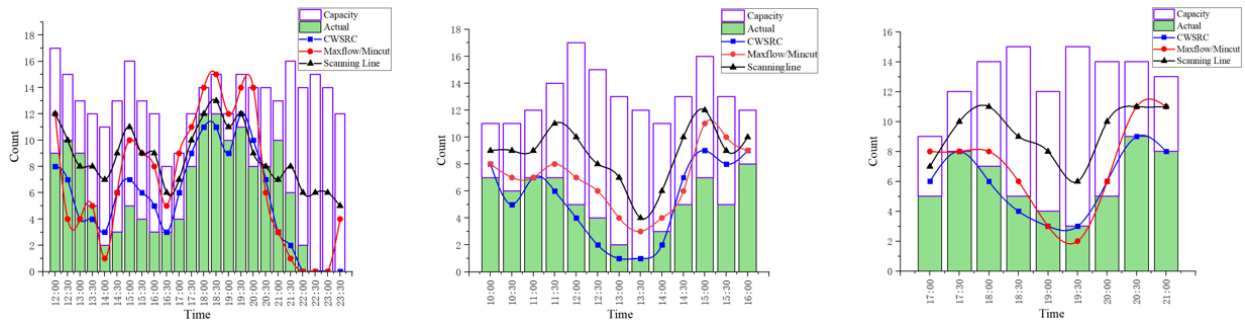


Figure 14. Comparison chart of capacity prediction results for three methods

Table 5 Deviation of predicted results from actual values for CWSRC model with Maxflow/Mincut, scanning line

Indicators	CWSRC	Maxflow/Mincut	Scanning line
MAE	1.17	2.12	3.19
MSE	2.84	7.42	12.80
RMSE	1.53	2.56	3.57
R^2	0.74	0.26	-1.22

On August 30, 2018, the Wuhan terminal area experienced an extended period of convective weather. During this day, the weather in the sector route segment transitioned from convective to clear conditions and then back to convective weather within the daytime. The sector route segment was notably affected by the fluctuations in convective weather, as illustrated in Figure. 11. Given the significant and typical nature of the convective weather on this day, we conducted a detailed analysis of the August 30 case.

Figure. 11 illustrates the hourly meteorological changes and flight trajectories on the ENH-YIH segment from 12:00 to 24:00 on August 30. The blue line represents the flight trajectories, the purple line represents the route's centerline, and the black represents the route's boundaries. From 12:00 to 15:00, convective weather progressively emerged at the segment's right end, correspondingly diminishing the flight traffic. In 15:00-17:00 the weather was gradually dissipating, but flight frequency remains low, partly due to it being a less busy time of the day with naturally lower flight volumes, and potentially because of reduced flights scheduled for that day. From 17:00-20:00 there is no weather on the segment and the traffic volume is close to the statistical capacity of the sector under clear weather. From 20:00 to 24:00, a substantial amount of convective weather appears at the right end of the sector route segment, resulting in a gradual decrease in segment capacity, eventually leading to no flight traffic.

As depicted in the left subplot of Figure. 14, the results predicted by RF are in the highest degree of conformity with the actual flow changes, and the trend remains consistent across each time interval. From 12:00 to 13:00, the scanning line method's prediction closely matched the actual traffic volume. However, RF's prediction was lower, reflecting the partial coverage of the flight route by the weather. The scanning line method reflects the effect better at this time. In 13:00-15:00, there is a wide range of convective weather in the sector, RF and Maxflow/Minicut effectively capture this change, while the scanning line method's predictions significantly diverged from the actual values. Between 15:00 and 17:00, although convective weather was gradually dissipating, the capacity under the recorded clear weather continued to decline during this period. As a result, the predictions from these methods demonstrated a decreasing trend, with RF predictions aligning more closely with the actual values. Between 17:00 and 20:00, the segment experienced clear skies, with traffic volume approaching the segment's full capacity under optimal conditions. However, from 20:00 to 24:00, significant convective weather developed at the segment's right end. This led to a steady decline in capacity and, ultimately, the cessation of flight traffic by the end of the period. RF accurately predicted a value of 0 between 22:30 and 23:30, consistent with the actual results. Yet, the prediction for 21:00-21:30 was lower than the actual traffic volume. This was because during this time, convective weather only covered the upper right tail end of the sector. There was no convective weather in other areas and flights can still fly and cross the route from the lower right side. Although flights were still navigable, the convective weather features were more pronounced during this period, resulting in the underestimated prediction. The Maxflow/Minicut prediction aligned with RF but fell short in comparison, notably deviating from the actual values at 23:30. The scanning line method gave poor prediction under convective weather. In general, the scanning line method performs better in predicting under minimal or no weather conditions but fares poorly under convective weather. Maxflow/Minicut exhibits a preference for predicting accurately during convective weather but falls short of RF's predictive performance, especially when there's no weather. RF demonstrates consistently strong predictive capabilities under both conditions

Figure. 12 and 13 illustrate the meteorological conditions observed on selected days in September 2018. The flight corridor under study is located in a region where convective weather predominantly occurs during the summer months. Notably, the intensity, duration, and coverage of convective weather events in September are generally less extensive than those observed in August. From the right panel and the center subplot of Figure. 14, the CWSRC model demonstrated superior accuracy in traffic forecasting relative to conventional methodologies. Nonetheless, the case studies revealed intriguing phenomena worthy of further investigation: On September 20, 2018, between 12:00 and 13:00, although convective weather was present, the convective cells were dispersed and individually small, permitting direct flight traversal through these areas. A similar occurrence was documented in the case study from September 25, 2018, suggesting a pattern that merits deeper exploration in future studies.

4 Conclusion

This study utilized an improved RFE algorithm to extract meteorological features and developed CWSRC model to predict sector route capacity. The model employs a RF supervised regression algorithm, using ten-fold cross-validation for training and calibrating against recorded traffic data to compute capacity. The analysis of three case studies confirms that the CWSRC model's predictions closely match the actual flight volumes during convective weather events. Key findings are outlined below:

(1) The CWSRC model achieved favorable performance metrics during training and exhibited reduced errors compared to traditional methods in the case study, underscoring its accuracy and robustness.

(2) RF algorithm showed a higher percentage (93.403%) of residuals within the range [-0.2, 0.2] on the validation set compared to KNN (86.03%) and MLP (77.37%) algorithms. This confirms that RF outperforms other machine learning algorithms in the CWSRC model.

(3) Due to the different traffic characteristics in different sector route segments, the influence of different weather factors on the prediction of sector route segment capacity is also different.

However, the model tends to underestimate capacity in scenarios with small and dispersed convective weather. It also tends to underestimate capacity when convective weather presenting at the beginning and end of sector route segment. This indicates a need for further research to enhance model accuracy: 1) incorporating adjacent airspace weather conditions as input features, 2) assessing the impact of dispersed, small convective cells that flights often traverse, and 3) considering the interaction effects of traffic volume on adjacent route segments.

References:

- [1] Civil aviation administration of china. (2022). *Civil Aviation Industry Development Statistics Bulletin for 2021* (Vol. 451). Beijing, China: Department of Development Planning. Retrieved from <https://www.mot.gov.cn/tongjishuju/minhang/202206/P020220607377281705999.pdf>.
- [2] FAA. (2022). *Air Traffic by the Numbers*. Washington: Federal Aviation Administration.
- [3] Rich DeLaura, M. An Exploratory Study of Modeling En Route Pilot Convective Storm Flight Deviation Behavior. In *12th Conference on Aviation Range and Aerospace Meteorology*. 2006.
- [4] DeLaura, R., Robinson, M., Pawlak, M., & Evans, J. Modeling convective weather avoidance in enroute airspace. In *13th Conference on Aviation, Range, and Aerospace Meteorology, AMS*, New Orleans, LA, 2008. Retrieved from <https://ci.nii.ac.jp/naid/20000753365>.
- [5] Klein, A., Cook, L., & Wood, B. Airspace availability estimation for traffic flow management using the scanning method. In *2008 IEEE/AIAA 27th Digital Avionics Systems Conference*, 2008, pp 3-B. URL: <https://doi.org/10.1109/dasc.2008.4702802>.
- [6] Song, L., Greenbaum, D., & Wanke, C. The impact of severe weather on sector capacity. In *8th USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)*, Napa, California, USA, 2009, pp 1-8. . Retrieved from http://www.atmseminar.org/seminarContent/seminar8/papers/p_075_W.pdf.
- [7] Callahan, M., DeArmon, J., Cooper, A., Goodfriend, J., Moch-Mooney, D., & Solomos, G. Assessing NAS performance: Normalizing for the effects of weather. In *4th USA/Europe Air Traffic Management R&D Symposium*. 2001, pp 3-7.
- [8] Klein, A., Jehlen, R., & Liang, D. Weather index with queuing component for national airspace system performance assessment. In *7th FAA/Eurocontrol ATM Seminar*. Spain, 2007, Retrieved from http://www.atmseminar.org/seminarContent/seminar7/papers/p_024_AAPM.pdf.
- [9] Chao, W., Xinyue, Z., & Xiaohao, X. (2008). Simulation Study on Airfield System Capacity Analysis Using SIMMOD. In *2008 International Symposium on Computational Intelligence and Design*. 2008, pp87-90, <https://doi.org/10.1109/iscid.2008.70>.
- [10] Sood, N., & Wieland, F. (2004). Total Airport and Airspace Model (TAAM) parallelization combining sequential and parallel algorithms for performance enhancement. In *Proceedings of the 2003 Winter Simulation Conference*. 2003, pp 1650-1655, <https://doi.org/10.1109/wsc.2003.1261615>.

-
- [11] Majumdar, A., Ochieng, W. Y., McAuley, G., Lenzi, J. M., & Lepadatu, C. (2004). The factors affecting airspace capacity in Europe: A Cross-Sectional Time-Series analysis using simulated controller workload data. *Journal of Navigation*, 57(3), 385–405. <https://doi.org/10.1017/s0373463304002863>.
- [12] Shi, F., Cheng, P., Geng, R., & Yang, M. (2012). An air traffic flow analysis system using historical radar data. In *Lecture notes in electrical engineering* (pp. 541–549). https://doi.org/10.1007/978-3-642-25766-7_72.
- [13] Mao, L., Peng, Y., Li, J., Guo, C., Kang, B., & Cao, Z. (2021). Random-forest based terminal capacity prediction under convective weather. *System Engineering-Theory & Practice*, 08, 2125-2136.
- [14] Chen, J., Cai, K., Li, W., Tang, S., & Fang, J. An Airspace Capacity Estimation Model based on Spatio-Temporal Graph Convolutional Networks Considering Weather Impact. In *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, 2021, pp 1-7. <https://doi.org/10.1109/dasc52595.2021.9594417>.
- [15] Wang, Y. Weather impact on airport arrival meter fix throughput. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference*, 2017, pp 1-10. <https://doi.org/10.1109/dasc.2017.8102133>.
- [16] Brito, I. R., Murca, M. C. R., Oliveira, M. D., & Oliveira, A. V. (2021). A machine learning-based predictive model of airspace sector occupancy. *AIAA AVIATION 2021 FORUM*. <https://doi.org/10.2514/6.2021-2324>.
- [17] Shijin, W., Jingjing, L., Jiahao, L., & Jiewen, C. (2021). Determination of Flight Deviation Thresholds based on the POT Model. *Aeronautical Computing Technique*, 51(3), 5.
- [18] Steiner, M., Bateman, R., Megenhardt, D., Liu, Y., Xu, M., Pocernich, M., & Krozel, J. (2010). Translation of Ensemble Weather Forecasts into Probabilistic Air Traffic Capacity Impact. *Air Traffic Control Quarterly*, 18(3), 229–254. <https://doi.org/10.2514/atcq.18.3.229>.
- [19] Genuer, R., Poggi, J., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>.

Competing interest declaration
Competing interests: The author(s) declare none.

Modeling sector route segment capacity under convective weather based on machine learning

WANG shijin¹, LI yinglin¹, YANG baotian¹, DUAN rongrong¹

1. Next Generation Intelligent Air Traffic Control Laboratory, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China

Abstract

Estimating airspace capacity under convective weather conditions is crucial for ensuring the efficiency and safety of air traffic operations. Sector route segments, which are essential components of flight routes, require timely capacity predictions during operationally critical periods. In this paper, initially, an enhanced Recursive Feature Elimination algorithm is used to select meteorological data and develop predictive features. Subsequently, the CWSRC model is established using the RF supervised learning algorithm. Finally, the paper takes ENH-YIH segment as an example to predict the capacity. Compared with other machine learning algorithms, the residual percentages for KNN, MLP, and RF are 86.03%, 77.37%, and 93.40%, respectively, within the range of $[-0.2, 0.2]$. In three separate day cases, results show that the CWSRC model's MAE, MSE, RMSE, and R^2 significantly outperform traditional methods like Maxflow/Mincut and scanning line. The results confirm the CWSRC model's superior predictive capabilities.

Nomenclature

ATM	air traffic management
BR	basic reflectivity[dBZ]
C	maximum flight number under clear weather of the sector route
C^p	capacity of a sector route segment during a specific period of a day
CR	composite reflectivity[dBZ]
CWAM	convective weather avoidance model
CWSRC	convective weather sector route capacity
E-WITI	en-route weather impacted traffic index
ET	echo tops[ft]
FAA	federal aviation administration
KNN	k-nearest neighbor
$Line_j$	convective weather duration of the flight on the j-th parallel line
MAE	mean absolute error
mcw	minimum available width on the route segment
MLP	multilayer perceptron
MSE	mean squared error
n	the number of grid cells
N_i^{xth}	the meteorological product value at the x-th percentile of the i-th airspace filter
NWS	national weather service
RF	random forest
RFE	recursive feature elimination
RMSE	root mean squared error
TMI	traffic management initiatives
T-WITI	terminal weather impacted traffic index
VIL	vertically integrated liquid[kg/m ²]
WITI	weather impacted traffic index

w_{ij}	the convective weather weight of the i -th grid of the j -th parallel line on the segment
W_r	the available proportion of traffic flow within the sector route segment
x^{th}	x -th percentile value of this airspace weather product
SRA_r, SRA_p	actual/ predicting airspace availability of sector route segment
SRC_r, SRC_p	actual/ predicting airspace capacity of sector route segment
Z	meteorological radar reflectivity factor
Δh	the height increment between successive radar measurements in meters.

0 Introduction

In the air transportation system, weather-related irregularities accounted for 59.56% and 66.5% of the total irregular flights in China^[1] and the United States^[2], respectively. Convective weather significantly reduces available airspace capacity, posing a major constraint on the normal operation of air traffic flow. Currently, traffic management processes heavily rely on the subjective interpretation of weather forecasts. Air traffic controllers use mental calculations to estimate reductions in airspace capacity and corresponding traffic flow rates. This approach fails to provide a clear translation of weather forecasts into tangible resource constraints. The air route is a fundamental component of the airspace system and crucial for the smooth operation of air traffic. Future ATM systems will require decision support tools. These tools must be capable of translating the impact of convective weather into precise air route capacities. Traffic demand for impacted air routes is managed through TMI. These initiatives either completely remove demand from an impacted air route or reduce demand by delaying the departure of flights filed through the impacted air route. Thus in regions of congested airspace, it is very important to be able to correctly estimate the air route capacity degradation due to convective weather effects so that avoidable delays and costs are minimized.

Many researchers have investigated the ability to translate a given weather situation into an estimate of available airspace capacity in recent years. The Lincoln Laboratory has made significant contributions in this area by developing three convective weather avoidance models, namely CWAM1, CWAM2, and CWAM3^[3-4]. They have also established a permeability calculation method and implemented the computation of airspace capacity using scanning line and Maxflow/Minicut methods^[5-6]. Additionally, the FAA developed the WITI model. This model uses historical meteorological data and traffic demand to accurately quantify the relationship between air traffic and weather. It predicts weather impacts on both E-WITI and T-WITI airspace^[6-8]. These methods rely on evaluating the severity of convective weather to estimate airspace capacity. However, the capacity of the air traffic management system is not solely dependent on convective weather conditions but is also influenced by various other factors, including route structure and the complexity of traffic flow operations. Simulation systems like SIMMOD^[9] by the FAA, TAAM^[10] by Boeing, RAMS^[11] in Europe, and ASMES^[12] in China have been developed to simulate and address the aforementioned issues. However, the complexity of the air traffic system requires sophisticated traffic flow simulation platforms and high computational performance, or complex mathematical models to accurately replicate the airspace operational environment for capacity evaluation. Moreover, these simulation systems are challenging to directly serve as decision support tools.

Recent advancements in computer data storage and computational capabilities, coupled with rapid developments in machine learning, have led to widespread applications in the civil aviation industry. This includes airport delay prediction, fleet planning, runway configuration forecasting, and route flight time estimation. Currently, machine learning is primarily focused on capacity prediction under convective weather conditions in the terminal area. Various studies have been conducted to address this area of interest. For instance, Mao et al. employed the random forest algorithm to predict the capacity of the terminal airspace in the presence of convective weather conditions^[13]. Chen et al. introduced the Spatio-Temporal Graph Convolution Network via Initial residual to effectively forecast the capacity of the terminal airspace^[14]. Wang et al., on the other hand, utilized a combination of multiple linear regression and regression tree ensemble learning methods to model and estimate the EWR arrival meter fix throughput^[15]. Additionally, Brito et al. developed a supervised machine learning-based prediction model to accurately forecast the elapsed time from a flight's takeoff to the sector crossing^[16]. However, it should be noted that the evaluation of capacity in the terminal airspace, employing machine learning algorithms, may not be directly applicable to the assessment of route resources. This discrepancy arises due to distinct differences in the structural characteristics, flight operation altitudes, control methods, and unique aspects associated with convective weather conditions between the terminal airspace and the route resources.

To quantify the impact of convective weather on airspace obstruction and accurately evaluate the capacity of air routes within the sector route segment, this paper establishes the CWSRC model. CWSRC model is based on the concept of sector

route blockage. It utilizes machine learning algorithms to predict airspace availability in the sector route during convective weather conditions, providing a consistent basis for collaborative decision-making in traffic flow operations. The structure of this paper is organized as follows: Section 1 introduces the fundamental concepts of sector route segment and the features used in our model. Section 2 describes the machine learning algorithms employed by the CWSRC model, detailing the methods used for training the model and the development processes. Section 3 analyzes the performance of three machine learning algorithms and evaluates the effectiveness of the CWSRC model. Section 4 summarizes the research findings and outlines future research directions.

1. Sector route segment and weather features

1.1 Sector route segment

An air route consists of a series of waypoints connecting multiple airspaces, typically long and complex. For instance, as depicted in Fig. 1, the routes include the light blue line A461, the dark blue line B334, the red line B339, and the green line G212. When convective weather exists in one location of a route, it does not affect the capacity of other segments farther away from that location. Therefore, it is neither accurate nor necessary to predict the capacity of a whole air route. Sector route segment refers to the segment between route points in a sector with a specific width, such as the part circled in yellow in Fig. 1. Flights usually fly according to the segments in the flight plan route, and a flight plan route contains a number of Sector route segments located in different routes. The white solid line on the left side of Fig. 1 indicates the flight plan route of the flight from ZHHH (Wuhan Tianhe International Airport) to ZYHB (Harbin Taiping International Airport), which passes through 4 routes and 11 sectors. This paper is concerned with predicting the capacity of the sector route segments.

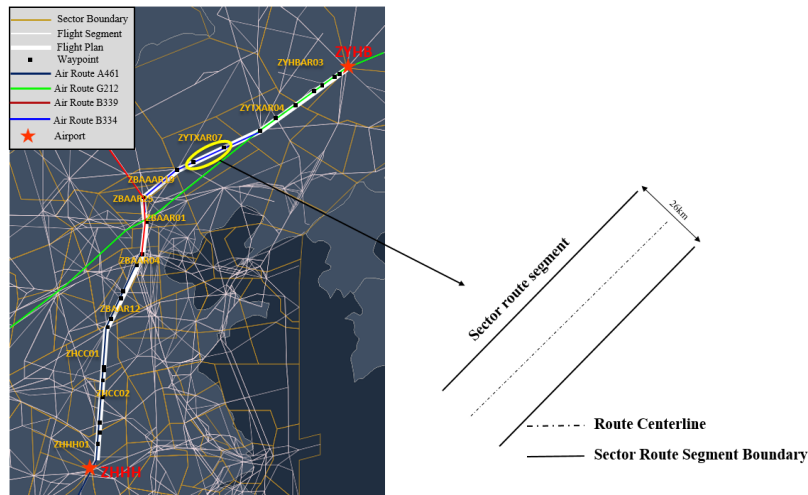


Figure 1. Schematic diagram of sector route segment

When a flight follows its planned route, deviations may occur due to factors such as weather, air traffic control decisions, and pilot operations. These deviations, while possibly altering the flight path slightly, do not constitute a rerouting. In order to accurately assess the capacity of the sector route segment, according to reference^[17], this paper takes the flight route width of 26km, as shown on the right side of Fig. 1, i.e., 13km on each side of the center-line of the flight route.

1.2 Meteorological features

Meteorological radar reflectivity Z , which reflects the strength of the radar echo signal and measures the intensity of precipitation and the physical properties of clouds, is updated every 6 minutes. In this paper, the information of meteorological features is extracted based on the derivatives of Z with the same refresh rate of 6 minutes. In addition, this paper uses the airspace filter method in extracting the percentile values of the meteorological products. The extraction process is shown in Fig. 2 and Equation. 1. The airspace filter is a square array, as shown on the right side of Fig. 2, which

consists of $n*n$ grids, each of which stores the weather product values in the latitude and longitude $0.01^\circ*0.01^\circ$ area. In this paper, a $26\text{km}*26\text{km}$ airspace filter is used, starting from one end of the sector route segment and divided based on the size of the airspace filter. There is no overlap between each airspace filter. If the remaining length of the route segment is not insufficient to form a complete filter, the grids adjacent to the segment are automatically filled to generate a complete airspace filter. The sector route segment area is divided into a number of airspace filters, extracting the x -th percentile value from each airspace filter. The x -th percentile value of this airspace weather product is obtained by computing the maximum value among all the filters within the airspace, as shown in Equation. 1.

$$x^{th} = \text{Max}\{N_1^{xth}, N_2^{xth}, \dots, N_i^{xth}\}, \quad x \in [50,75,90,100] \dots\dots\dots(1)$$

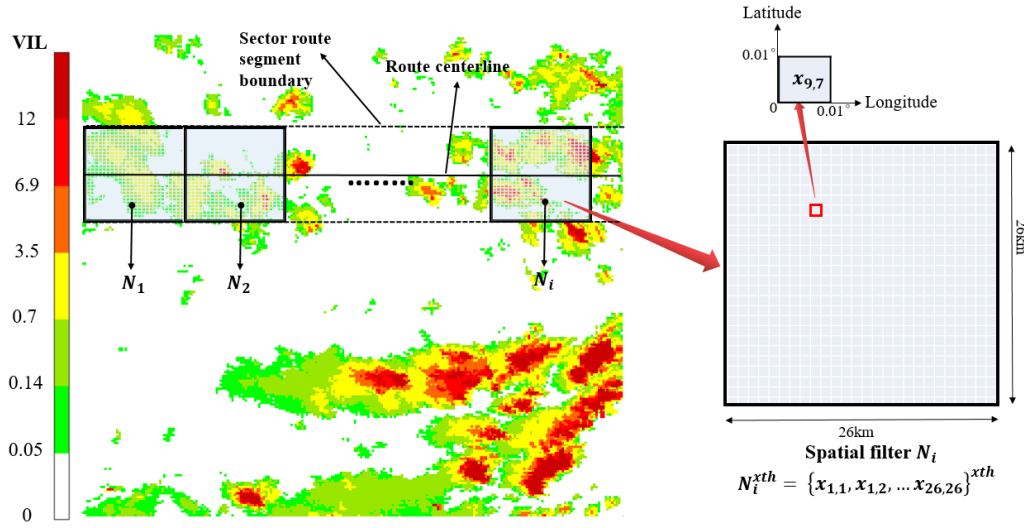


Figure 2. Airspace filter

where N_i^{xth} denotes the meteorological product value at the x -th percentile of the i -th airspace filter.

The NWS classifies convective weather as shown in Table 1^[18]. When the value of meteorological products CR, VIL and ET in the airspace exceeds Level 3 (including Level 3), it is regarded as severe weather, and flights are not allowed to fly over the airspace. Therefore, this paper selects CR, ET and VIL as weather features to predict sector segment capacity.

Table 1 Convective weather classification

Category	CR(dBZ)	VIL(kg/m ²)	ET(ft)	Level
None	<18	0-0.14		0
Light	18-30	0.14-0.7	<25,000	1
Middle	30-41	0.7-3.5		2
Heavy	41-46	3.5-6.9		3
Very heavy	46-50	6.9-12	25,000~35,000	4
Intense	50 -57	12-32		5
Extreme	≥57	≥32	>35,000	6

1.2.1 VIL

VIL is used to describe the amount of liquid water accumulated in the entire atmospheric column from the altitude of the radar to the top of the atmosphere. It is calculated as shown in Equation 2. Higher VIL values usually indicate the presence of a significant amount of liquid water within a thunderstorm, often foreshadowing the occurrence of severe convective weather such as heavy rainfall, hail, and strong winds.

$$\text{VIL} = \sum 3.44 \times 10^{-6} \left(\frac{Z_i + Z_{i+1}}{2} \right)^{4/7} \Delta h \quad (2)$$

where Z_i is the radar reflectivity and Δh is the height expressed in meters.

In this paper, the VIL data are stored in mat files, and each element in the matrix represents the VIL value of $0.01^\circ \times 0.01^\circ$ airspace cell at the specified location. In order to express the severity of convective weather in the airspace, this paper has counted the VIL percentile value of the airspace for the specified segment. Furthermore, in order to measure the coverage of convective weather in the airspace, this paper has counted the proportion of the number of $VIL \geq 3.5 \text{ kg/m}^2$ rasters in the airspace of the specified segment. Therefore the VIL features used in this paper are:

1. 50^{th} VIL: the 50th percentile VIL value in the airspace of the segment;
2. 75^{th} VIL : the 75th percentile VIL value in the airspace of the segment;
3. 90^{th} VIL : the 90th percentile VIL value in the airspace of the segment;
4. $maxVIL$: the maximum VIL value in the airspace of the segment;
5. *Proportion of $VIL \geq 3.5 \text{ kg/m}^2$* : the number of grids with $VIL \geq 3.5 \text{ kg/m}^2$ as a percentage of all grids in the airspace of the segment.

1.2.2 ET

ET is the height corresponding to an echo of 18.3 dBZ detected by meteorological radar, as shown in Equation 3. It is used to measure the intensity and vertical development of convective activity. Higher ET values usually imply higher convective cloud tops, indicating that severe convective weather might be developing or has already developed. The ET features used in this paper include:

6. 50^{th} ET: the 50th percentile ET value in the airspace of the segment;
7. 75^{th} ET: the 75th percentile ET value in the airspace of the segment;
8. 90^{th} ET: the 90th percentile ET value in the airspace of the segment;
9. $maxET$: the maximum ET value in the airspace of the segment.

$$ET = h_{BR > threshold} \quad (3)$$

1.2.3 CR

The BR after logarithmic transformation of Z can reflect the magnitude of Z value as in Equation 4.

$$BR = 10 \log(Z) \quad (4)$$

CR refers to a product derived within a single volume scan that projects the maximum reflectivity factor found in constant elevation azimuth scans onto Cartesian grid points. It represents the maximum reflectivity factor within a unit volume, as depicted in Equation 5.

$$CR = max BR \quad (5)$$

A larger value of CR represents a higher convective intensity.

10. 50^{th} CR: the 50th percentile CR value in the airspace of the segment;
11. 75^{th} CR: the 75th percentile CR value in the airspace of the segment;
12. 90^{th} CR: the 90th percentile CR value in the airspace of the segment;
13. $maxCR$: the maximum CR value in the airspace of the segment;
14. *Proportion of $CR \geq 41 \text{ dBZ}$* : the number of grids with $CR \geq 41 \text{ dBZ}$ as a percentage of all grids in the airspace of the segment.

1.2.4 Longitudinal weather features of segment

This article analyzes longitudinal and lateral convective weather conditions in sector route segment to identify their impact on air traffic flow.

The impact of longitudinal weather in sector route segment on air traffic flow is reflected in the length of the path of a flight crossing the convective weather portion, indicating the severity of longitudinal convective weather in sector route segment. Figure. 3 presents the method for assessing longitudinal convective weather in sector route segment.

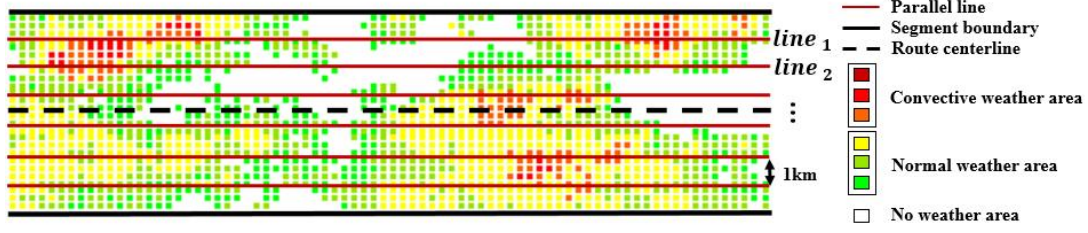


Figure 3. Schematic of longitudinal weather feature

Parallel lines are drawn along the direction of the sector route, spaced at intervals of 1 km parallel to the centerline. The VIL/CR values of all grids along each parallel line are calculated. When the VIL/CR value exceeds level 3 weather, the raster is considered to be in convective weather, and the convective weather weight value of the raster is set to be 1; Otherwise, the weight value of the raster is VIL/3.5 or CR/40. The sum of the convective weather weights of all the rasters on each parallel line indicates the longitudinal weather severity of the line as shown in Equation. 6.

$$line_j = \sum_{i=1}^n w_{ij}, w_{ij} = \begin{cases} 1, & \text{if } VIL \geq 3.5 \text{ or } CR \geq 40 \\ \frac{value}{3.5} \text{ or } \frac{value}{40}, & \text{if } VIL < 3.5 \text{ or } CR < 40 \end{cases} \quad (6)$$

Where w_{ij} denotes the convective weather weight of the i^{th} grid of the j^{th} parallel line on the segment. $Line_j$ denotes the convective weather duration of the flight on the j^{th} parallel line, and value denotes the VIL or CR value of each grid on the parallel line in this function.

The longitudinal route segment weather features are:

15/16. $VIL_parallel_mean/CR_parallel_mean$: The mean value of the impact along each parallel line represents the longitudinal average impact (based on VIL/based on CR);

17/18. $VIL_parallel_max/CR_parallel_max$: The maximum value of impact among all parallel lines (based on VIL/based on CR);

19/20. $VIL_parallel_min/CR_parallel_min$: The minimum value of impact among all parallel lines (based on VIL/based on CR);

21/22. $VIL_parallel_median/CR_parallel_median$: The median value of impact among all parallel lines (based on VIL/based on CR).

1.2.5 Lateral weather features of segment

When convective weather does not entirely cover the cross-section of a sector route segment, flights along the route can still traverse areas not affected by the weather. The lateral impact of convective weather on the sector route segment manifests in the degree of obstruction in the direction of the route's width, quantifying the proportion of the route's width covered by convective weather in the vertical direction, as indicated in Equation. 7.

$$W_r = \frac{mcw}{width} \quad (7)$$

where mcw denotes the minimum available width on the route segment, $width$ denotes the route segment width, and W_r represents the available proportion of traffic flow within the sector route segment.

The determination of the minimum available width of the segment, as illustrated in Figure. 2-4, begins by identifying the grid cells within the sector route affected by convective weather; Subsequently, the DBSCAN algorithm is employed to cluster multiple convective weather blocks within the route, depicting the progression from Figure. 4(a) to Figure. 4(b); The clustered convective weather blocks are utilized with a convex hull algorithm to generate a convex hull region, denoting the area impassable for aircraft, depicting the process from Figure. 4(b) to Figure. 4(c).

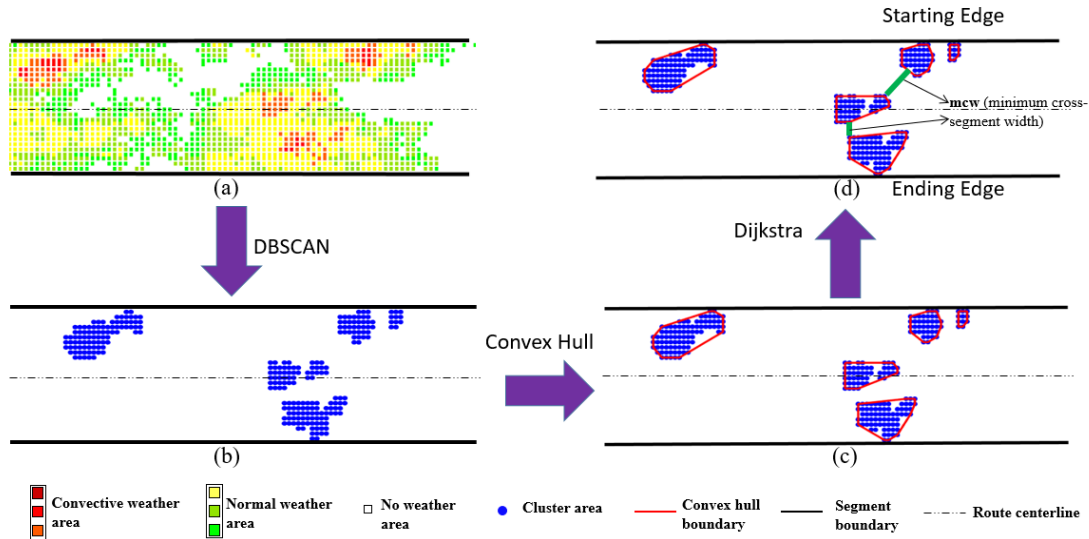


Figure 4. Flow percentage calculation process diagram for crossable sector route segments

Finally, the Dijkstra algorithm is applied to determine the *mcw* in the vertical direction within the sector route. The lateral features of the flight path in this section are included:

23. *Cross_flow_VIL*: the proportion of traffic flow that can traverse a route, calculated based on VIL;
24. *Cross_flow_CR*: the proportion of traffic flow that can traverse a route, calculated based on CR.

2 Methodology

Machine learning demonstrates excellent accuracy and generalization performance. This paper employs the CWSRC model based on the random forest algorithm for regression prediction of sector route segment capacity. The structure of the model, as depicted in Figure.5, involves meteorological product data as input and forecasts sector route segment capacity as output. This model is mainly divided into three parts: data preparation, model training and model testing.



Figure 5. Structural diagram of the CWSRC model

For the purposes of this study, airspace capacity is defined as the maximum number of aircraft that a specific unit of airspace can accommodate within a given time period. There could be a significant difference between the actual number of flights and the number of flights that can be accommodated under clear weather. As convective weather intensifies, airspace capacity decreases. To minimize flight delays and enhance operational efficiency, it is essential to match the number of flights with the capacity as closely as possible. Therefore, this paper assumes that the actual number of flights within a sector route segment during convective weather represents the segment's capacity under those conditions. Similarly, under clear weather, the maximum number of flights within an hour is considered the actual operational capacity of the sector route segment.

2.1 Data Set

The CWSRC model requires the use three meteorological product data types: VIL, ET, and CR. It extracts meteorological feature data based on sector route segment boundary information. To predict segment capacities based on varying meteorological features, the CWSRC model utilizes a dataset of historical flight radar trajectories through the sector route segment. Each flight trajectory in the dataset records the latitude and longitude coordinate position information of the flight at the corresponding time point, including the time of entering and exiting the sector route segment.

To validate the predictive performance of this model, the data is divided into two parts. The first part includes the training and validation datasets used for cross-validation to refine and validate the model. The second part comprises the test data, which is input into the trained model to analyze its performance.

2.2 Model training

To ensure robust predictive performance of the CWSRC model, we established label definitions during the model training process, performed feature selection, and fine-tuned parameters for the RF algorithm.

2.2.1 Label

This paper uses airspace availability as the label, airspace availability denoted as SRA_r , which is calculated as shown in Equation 8.

$$SRA_r = \frac{SRC_r}{C} \quad (8)$$

In Equation 8, SRC_r and C represents the actual flight number under convective weather and the maximum flight number under clear weather of the sector route in the corresponding period. The number of flights in sector route segment varies significantly across different time intervals, typically with fewer flights in the early hours and a higher volume during daylight hours. However, there exists a degree of similarity in the number of flights during equivalent time segments on different days. This study involves the calculation of sector route segment capacities for each hour within a day over a specific time span (usually a month), as derived in Equation 9.

$$C^p = 90^{\text{th}}\{C_1^p, C_2^p, \dots, C_i^p\} \quad (9)$$

where C^p represents the hourly capacity of a sector route segment during a specific period of a day, and C_i^p represents signifies the sector route segment's hourly capacity for a particular day within the statistical time span. Within a given time span, there might be instances where a day's sector route segment hourly capacity exceeds its limit. We take the 90th percentile segment capacity value to represent the maximum accommodatable flight volume for that specific hour within the sector route segment. The availability label is the current hourly flight traffic divided by the current hourly capacity.

In the later part, SRA_p and SRC_p are the predicted value of availability and capacity, respectively.

2.2.2 Model evaluation metrics

In the field of machine learning, a common metric for assessing algorithm accuracy is the generalization error. However, directly calculating the generalization error is not practical in applications due to the unknown distribution of the sample data. Therefore, we utilize a validation set constructed from sample data and compute various performance metrics to estimate the generalization performance of the RF algorithm.

This study employs four commonly used regression prediction metrics, including MSE, RMSE, MAE, and R^2 . The calculations are depicted in Equations 10,11,12,13.

$$MSE = \frac{\sum_{i=1}^n (y_i - f_i)^2}{n} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - f_i|}{n} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}} \quad (12)$$

$$R^2 = 1 - \frac{\sum (y_i - f_i)^2}{\sum (y_i - \bar{y})^2} \quad (13)$$

MSE measures the mean of the squares of the prediction errors. A smaller MSE value indicates that the model's predictions are closer to the actual values, signifying better model performance. RMSE is the square root of MSE. It provides the standard deviation of the prediction errors. RMSE offers an intuitive feel for the size of the errors; compared to MSE, RMSE has the same units as the original data, making it easier to understand. MAE provides the average size of the prediction errors and, unlike MSE and RMSE, does not assign higher weight to larger errors. R^2 is a metric that measures the model's ability to explain variability in the data. An R^2 value closer to 1 indicates that the model explains a higher variability, leading to better predictive performance.

2.2.3 Model training methods

This study utilizes K-fold cross-validation method to assess the model's generalization performance and subsequently adjust the RF parameters. Cross-validation can effectively mitigate the risk of overfitting and accurately reflects the model's generalization capabilities. The RF design in the CWSRC model is illustrated in Algorithm 1. Within the K-fold cross-validation method, the original dataset is randomly partitioned into k equally sized subsets. Each iteration selects one subset the model's validation dataset while using the remaining $k-1$ subsets as the training data. This process is repeated k times, rotating through each subset as the validation set, until all subsets have been used for validation. The metrics results obtained from each training validation are recorded, and the average of these k results yields the overall training performance for a particular set of hyperparameters. This iterative training and testing process makes the model performance assessment more reliable and reduces the sensitivity to the data partitioning method.

Algorithm1

Algorithm: RF

Input: Data set S

$$S = S_1 \cup S_2 \cup \dots \cup S_K, \quad S_i \cap S_j = \emptyset \quad (\forall i \neq j)$$

parameters: the number of decision trees T

Process:

- 1: While $k \leq K$ do
- 2: $S_{train} \leftarrow$ training set, $S - S_k$
- 3: $S_{test} \leftarrow$ test set, S_k
- 4: $N \leftarrow$ S_{train} size
- 5: $n \leftarrow$ S_{test} size
- 6: $T \leftarrow$ the number of decision trees
- 7: while $t \leq T$ do
- 8: $D_t \leftarrow$ Bootstrapping(S_{train}, N)
- 9: $f_t \leftarrow$ DTree(D_t)
- 10: end while
- 11: $f \leftarrow \frac{1}{T} (\sum_{t=1}^T f_t(S_{test}))$
- 12: $RMSE_k \leftarrow \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2}$
- 13: $MSE_k \leftarrow \frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2$
- 14: $MAE_k \leftarrow \frac{1}{n} \sum_{i=1}^n |y_i - f_i|$
- 15: $R_k^2 \leftarrow 1 - \sum_{i=1}^n (y_i - f_i)^2 / \sum_{i=1}^n (y_i - \bar{y})^2$
- 16: end while

$$17: RMSE \leftarrow \frac{\sum_{k=1}^K RMSE_k}{K}$$

$$18: MSE \leftarrow \frac{\sum_{k=1}^K MSE_k}{K}$$

$$19: MAE \leftarrow \frac{\sum_{k=1}^K MAE_k}{K}$$

$$20: R^2 \leftarrow \frac{\sum_{k=1}^K R^2}{K}$$

2.2.4 Feature selection

This investigation incorporates 24 comprehensive features, which, while enriching the dataset, also extend the data collection phase and augment the model's complexity. Effective feature selection is pivotal, as it not only simplifies the model by eliminating features with high intercorrelations but also enhances its generalization capabilities and abbreviates the training duration. Conventionally, RFE operates by sequentially removing the least significant feature. This procedure, although systematic, may inadvertently undermine the model's performance, particularly when the excised feature plays a pivotal role in certain combinations.

To address these shortcomings, the present study proposes an advanced modification of RFE that integrates aspects of stepwise regression. Following the exclusion of each feature, a comprehensive assessment is conducted to evaluate its interaction with other features and overall impact on model performance. Should this assessment reveal that certain eliminated features substantially bolster the model's efficacy, these features are subsequently reinstated in the feature set during later iterations. This meticulous approach refines the feature selection process, culminating in an optimal feature subset that significantly enhances the model's accuracy and robustness. Algorithm 2 outlines the algorithmic process.

Algorithm2

Algorithm: Recursive Feature Elimination Algorithm

Input: feature x , label y , model $model$, number of target features $n_features_to_select$

Process:

```

1: Indicator = []
2: current_number_of_features = total number of features  $N$ 
3: function recursiveFeatureElimination( $x, y, model, n\_features\_to\_select$ ):

4:   while current_number_of_features >  $n\_features\_to\_select$ :
5:      $model.fit(x, y)$  // Train the model on current feature set
6:     importance =  $model.feature\_importances\_$  // Calculate feature importances

7:     // Identify and remove the least important feature
8:     least_important_index = index of feature with  $\min(importance)$ 
9:      $x$  = remove feature at least_important_index from  $x$ 
10:    current_number_of_features = current_number_of_features - 1

11:    // Evaluate model performance
12:    indicator =  $evaluate\_model(model, x, y)$  // Function to evaluate model performance
13:    Indicator.append(indicator) // Record performance indicator

14:    // Record feature if performance worsens
15:    delete_feature = []
16:    if  $\text{length}(Indicator) > 1$  and indicator worse than Indicator[-2]:
17:      record the removed feature as delete_feature

18:    // Recursive call for next iteration
19:    recursiveFeatureElimination( $x, y, model, n\_features\_to\_select$ )

20:  return subset of features: subset = [current_feature, delete_feature]

21: End function

```

In Algorithm2, further stepwise explain are as follows:

- 1) Initialization (Lines 1-3): Initialize '*Indicator*' to store performance indicators and '*current_number_of_features*' to track feature count. Initialize the '*recursiveFeatureElimination*' function.

-
- 2) Loop and Training (Lines 4-5): Enter a loop to iteratively eliminate features until '*current_number_of_features*' reaches '*n_features_to_select*'. Train the model ('*model.fit*') on the current feature set ('*x*' and '*y*').
 - 3) Feature Importance Calculation (Line 6): Calculate feature importance using the model's '*feature_importances_*' attribute.
 - 4) Feature Removal (Lines 7-10): Identify and remove the least important feature from '*x*' based on calculated importance. Update '*current_number_of_features*' accordingly.
 - 5) Performance Evaluation (Lines 11-13): Evaluate model performance using the reduced feature set ('*x*' and '*y*'). Record the performance indicator ('*indicator*') in '*Indicator*'.
 - 6) Comparison and Recording (Lines 14-17): Compare the current performance indicator with the last recorded one ('*Indicator[-2]*'). If performance worsens, record the removed feature as '*delete_feature*'.
 - 7) Recursive Call (Lines 18-19): Recursively call '*recursiveFeatureElimination*' to continue feature elimination with the updated feature set.
 - 8) Return (Line 20): Once '*current_number_of_features*' equals '*n_features_to_select*', return the subset of features selected ('*subset = [current_feature, delete_feature]*').

Begin by establishing the desired number of target features. The original features are then input into the RF algorithm, the results of the performance metrics obtained from this training will be recorded, and the importance of the features will be ranked according to the RF's feature importance evaluation function. Afterwards, selecting features based on the importance derived from random forest can identify the most influential features on the model's predictive outcomes. After eliminating the features with the least importance, proceed to the next round of training with the remaining features. Record any excluded features when the metrics from the subsequent training round show a decline compared to the previous results. The process is repeated until the number of features reaches the target number of features, and the recorded excluded features are added to the remaining features to obtain the selected subset of features after screening.

2.2.5 Hyperparameter determination

The number of decision trees is a crucial hyperparameter in the RF algorithm. RF enhances the generalization and stability of the model by aggregating the predictions of multiple decision trees. Each decision tree conducts random sampling on the training data and performs splits on randomly selected subsets of features, mitigating overfitting risks. Thus, while increasing the number of decision trees in a RF typically improves performance, there's a limit to the increase as adding trees indefinitely may not guarantee improved performance. Considering the cost of computational resources and time consumption, it's essential to establish a range of decision tree numbers. In this research, we follow precedents set in prior studies by setting the number of trees in the RF algorithm between 0 and 500^[19]. By traversing this range of parameters and observing the relationship between performance metrics and the number of decision trees through visualizations, we can determine the trend of performance metrics' variations and ascertain the optimal number of decision trees.

2.3 Model test

In this paper, we test the adjusted RF algorithm using untrained data. Meteorological features are used as inputs to predict the availability of sector route segment. In addition, based on the timeframe of the samples, we determined the corresponding capacity of sector route segments for each period. Multiplying airspace availability by period capacity yielded the sector route segment capacity predicted by the CWSRC model. Analyzing the differences between the predicted results on the test set and the actual values, and comparing them with traditional capacity prediction methods, allowed us to further evaluate the effectiveness of the CWSRC model's predictions.

2.4 Model stepwise

The development steps of the CWSRC model are as follows.

- 1) Label definition: Initially, sample labels are defined, with *SRA* being utilized as the primary label for this study.

- 2) Data Synthesis: The identified labels and corresponding features are integrated to form a comprehensive sample set.
- 3) Training and Validation Set Split: This sample set is subsequently partitioned into training-validation sets and a separate test set to facilitate robust model evaluation.
- 4) Performance Metrics Definition: At this stage, performance metrics such as MAE, MSE, RMSE, and R^2 are selected to assess the effectiveness of the model.
- 5) Cross Validation: The model is trained utilizing cross-validation techniques to enhance its generalizability.
- 6) Feature Selection: An enhanced RFE algorithm is employed to identify and retain the most impactful features.
- 7) Parameter Optimization: Following feature selection, the model's hyperparameters are finely tuned to optimize performance.
- 8) Execution on Test Data: The test set, segregated in the data segmentation phase, is employed to evaluate the optimally trained model.
- 9) Results Analysis: The outcomes of the model testing are analyzed. Key performance indicators are computed and compared against those derived from traditional methodologies.

3 Result and Discussion

This section uses the ENH-YIH sector route segment in the Wuhan control area of China as an example. Based on the methods detailed in Section 2, specific feature selection and parameter optimization are conducted to enhance model performance. The optimization results of the RF-based CWSRC model are compared with those of other machine learning algorithms in a conducted comparative analysis. This study is to assess the generalization ability of the model on new, untrained data and to compare the performance of the CWSRC model against traditional methods such as Maxflow/Mincut and the scanning line approach. MAE, MSE, RMSE and R^2 are used as metrics to evaluate our objectives in the test dataset.

3.1 Data preparation

In this paper, the historical radar trajectory data, flight plan data and meteorological data from August 3 to 31, 2018 of a busy sector route in Wuhan control area is used for verification of CWSRC model. The horizontal range of Wuhan control area is shown in Figure.6, and the vertical range is from 6,000 to 7,800 meters.

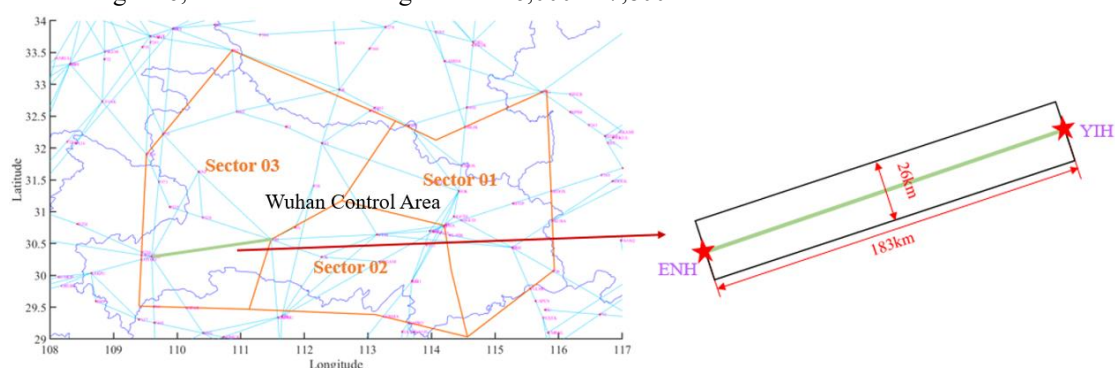


Figure 6. Position of the ENY-YIH

The sector route is ENH-YIH, located in sector 03 with a length of 183km in an east-west direction. It is the part of the sector route segment in the Shanghai-Chengdu direction and the flights accounts for about 30% of the total. Radar trajectory data is updated at a rate of 8 seconds, including the time, flight number, height, longitude, latitude and other information of historical flights. Longitude and latitude can be used to determine whether a trajectory exists within a sector area; time can be used to determine which time period the flight's traffic belongs to; and altitude is used to determine whether the flight is within the altitude range of the sector route segment. The meteorological data in the China's airspace

29°N-33°N and 109°E-116°E, VIL, CR and ET, is stored as $0.01^\circ \times 0.01^\circ$ (Longitude \times Latitude) and updated at a rate of 6 minutes.

Statistics of flights number in each period from August 3 to 31, 2018 are shown in Figure. 7. The time in the abscissa refers to one hour after this point in time. The busy periods of both two sector routes are from 8:00 to 23:00 each day, therefore historical flights data and weather data of these periods are selected in this paper. There are 4350 periods in the sector route. After screening, there are 3350 periods of ENH-YIH meeting convective weather conditions. We employed a ten-fold cross-validation approach, dividing the dataset into training and validation sets at a ratio of 9:1. Additionally, data from 46 time periods were utilized as the test set for this study.

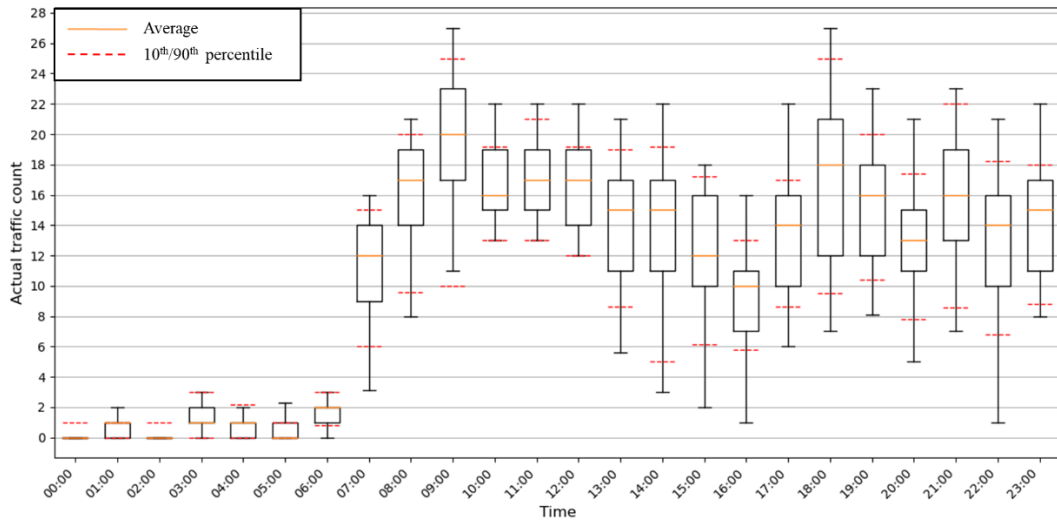


Figure 7. Distribution of operational capacity across time periods in the ENH-YIH

3.2 Model Optimization

The number of target features is set to 1, and the RF algorithm with default parameters is employed for recursive feature elimination. As the feature set iterated down to just one feature, the performance metrics curves for the entire iterative process were observed. Figure. 8 displays the results of the model's prediction for *SRA* using the metrics MAE, MSE, RMSE, and R^2 .

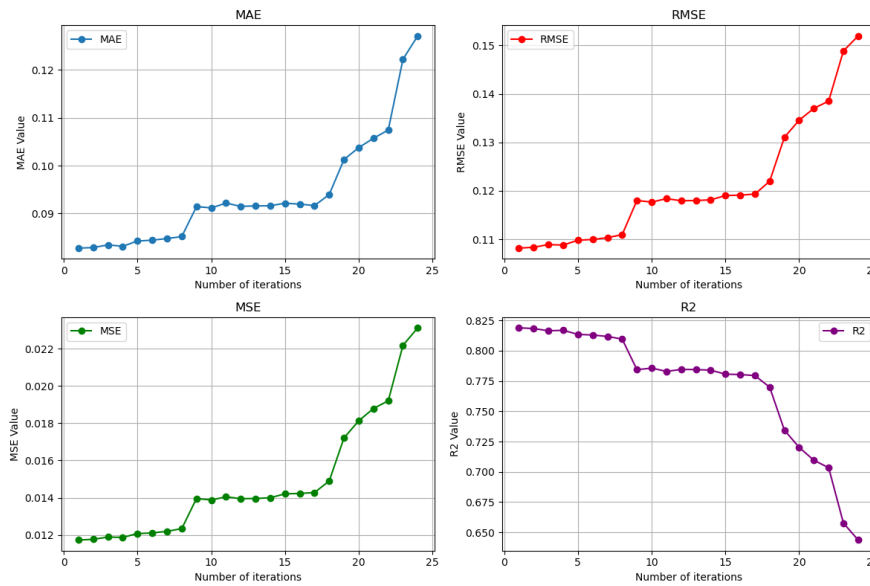


Figure 8. Changes in metrics with different number of iterations

Based on the performance metrics plotted in the aforementioned figure, Table 2 presents the impact of each iteration's feature elimination on the model's performance. An improvement in model performance upon feature elimination suggests

that the feature contributed more erroneous information, necessitating its removal. Conversely, when removal features make the model performance worse, it means that the feature is more effective. Therefore, the features causing a relative deterioration in model performance compared to the previous iteration are considered the optimal feature subset for elimination. From the table, the optimal feature subset excludes: *maxCR*, *50th VIL*, *Proportion of VIL \geq level 3*, *CR_parallel_mean*, *VIL_parallel_min*. The comparison between the model's performance using this feature subset and the original features is illustrated in Table 3. Although there is minimal change in MAE, MSE, RMSE, and R², the training prediction time decreased from 1'24" to 1'10" with the reduced feature set.

Table 2: Indicator performance after feature culling relative to the previous iteration (ρ : denotes the performance with the previous round after removing the feature)

Features	ρ	Features	ρ
<i>50th ET</i>	Worse	<i>Proportion of VIL \geq level 3</i>	Better
<i>75th ET</i>	Worse	<i>Proportion of CR \geq 41dBZ</i>	Worse
<i>90th ET</i>	Worse	<i>CR_parallel_mean</i>	Better
<i>maxET</i>	Worse	<i>CR_parallel_max</i>	Worse
<i>50th CR</i>	Worse	<i>CR_parallel_min</i>	Worse
<i>75th CR</i>	Worse	<i>CR_parallel_median</i>	Worse
<i>90th CR</i>	Worse	<i>VIL_parallel_mean</i>	Worse
<i>maxCR</i>	Better	<i>VIL_parallel_max</i>	Worse
<i>50th VIL</i>	Better	<i>VIL_parallel_min</i>	Better
<i>75th VIL</i>	Worse	<i>VIL_parallel_median</i>	Worse
<i>90th VIL</i>	Worse	<i>Cross_flow_CR</i>	Worse
<i>maxVIL</i>	Worse	<i>Cross_flow_VIL</i>	--

Table 3: The performance metrics of raw features versus filtered features

Features	MAE	MSE	RMSE	R ²	Time
Original	0.08276	0.01173	0.10817	0.81897	0:01:24.47
Filtering subsets	0.08289	0.01179	0.10846	0.81801	0:01:10.19

Table 4 presents the feature importance rankings within the RF model. Notably, the transverse and longitudinal weather characteristics of the sector rank highest in importance, validating our proposed approach that incorporating directional features of flights significantly enhances the prediction of sector route capacity.

Table 4: Importance of features

Feature	Importance	Feature	Importance
<i>Cross_flow_VIL</i>	23.89%	<i>90th ET</i>	3.37%
<i>VIL_parallel_median</i>	13.72%	<i>maxVIL</i>	3.36%
<i>VIL_parallel_mean</i>	8.03%	<i>50th ET</i>	3.15%
<i>CR_parallel_max</i>	5.87%	<i>75th CR</i>	2.69%
<i>CR_parallel_median</i>	5.31%	<i>90th CR</i>	2.54%
<i>50th CR</i>	4.74%	<i>Proportion of CR \geq 41dBZ</i>	2.43%
<i>CR_parallel_min</i>	3.88%	<i>90th VIL</i>	2.21%
<i>75th ET</i>	3.86%	<i>75^h VIL</i>	1.98%
<i>maxET</i>	3.84%	<i>Cross_flow_CR</i>	1.64%
<i>VIL_parallel_max</i>	3.46%		

The selected subset of optimal features is input into the model with default parameters. The number of decision trees for RF is set from 1 to 500 to derive the model performance metrics at each training. As depicted in Figure. 9, after the number of trees is 100, there was no significant change observed in MAE, MSE, RMSE, and R^2 , indicating convergence. However, the training time becomes longer as the number of trees increases. Considering the cost of time, this study settles on 100 trees for RF algorithm.

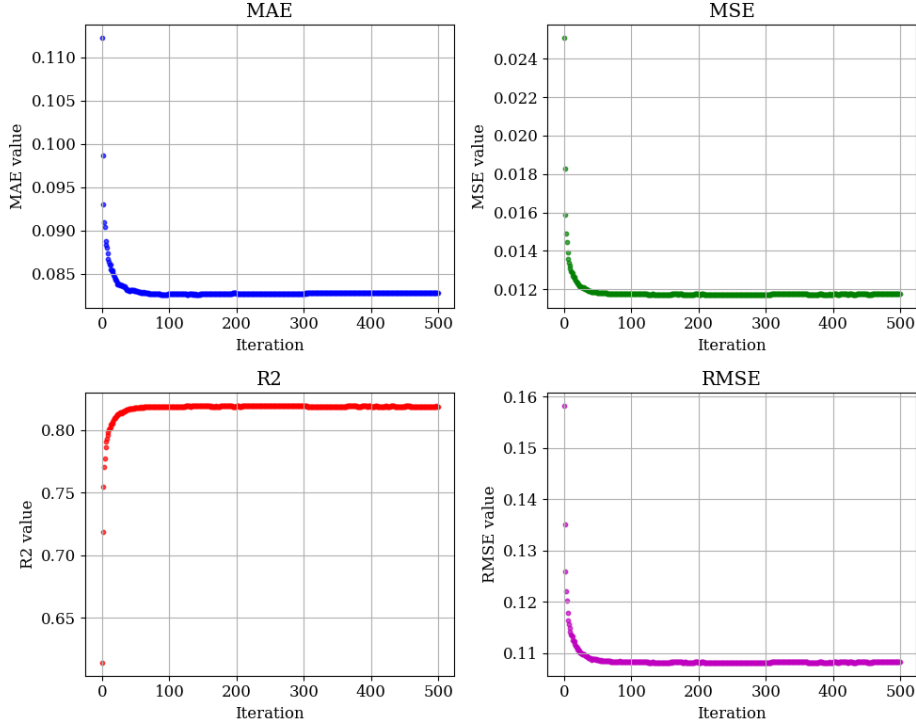


Figure 9. Different performance indicators based on tree numbers

3.3 Result analysis

To validate the model, this study excluded data from August 30 and utilized data from August 3 to August 31, 2018, for model training and testing. August 30, 2018, September 20, 2018, and September 25, 2018, were selected as case study days for model testing, as convective weather was detected in the sector route segment on these days. Due to the predominance of clear weather samples in the historical dataset, oversampling techniques were applied to the convective weather samples. The algorithm's performance was analyzed based on the model's results on the validation set. Furthermore, the model's performance on the test set was evaluated.

3.3.1 Algorithm analysis

In order to verify the superior performance of RF algorithm compared to other machine learning algorithms, this paper compares the results predicted by RF with those predicted by the KNN and MLP algorithms with default parameters. Among them, the parameters of machine learning algorithm are set as follows. The distance metric in KNN algorithm is Euclidean distance, and k is 10. The number of decision trees in RF is 100. There are two hidden layers in MLP, and the hidden layer units are 4 and 7, respectively. The activation function is *Relu*, solver is *adam*, the learning rate is 0.001, and the maximum number of iterations is 300.

Figure. 10 illustrates the residual distributions of SRA_p for KNN, MLP and RF on the cross-validated test set, where the interval of the statistics is $[-0.6, 0.6]$ with a step size of 0.05. Notably, the predominant residual distribution for RF lies within the $[-0.05, 0.05]$ range, exhibiting significantly higher sample counts compared to MLP and KNN within this interval. The number of samples with residuals between $[-0.3, 0.3]$ is 3309, 3106 and 3335 of KNN, MLP and RF respectively, accounting for 98.78%, 92.72% and 99.55% of the total data respectively. The number of samples with residuals between $[-0.2, 0.2]$ is 2882, 2592 and 3129 respectively, accounting for 86.03%, 77.37% and 93.403% of the total data respectively.

The number of samples with residuals between $[-0.1, 0.1]$ is 1637, 1435 and 2131 respectively, accounting for 48.87%, 42.84% and 63.61% of the total data respectively.

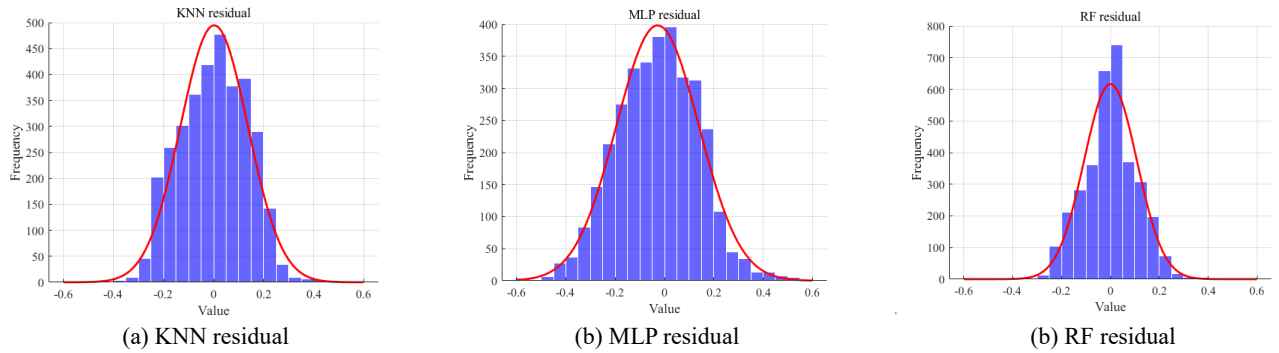


Figure 10. Residual plots of KNN, MLP, and RF algorithms

3.3.2 Analysis of SRC_p

In this section, the SRC_p of ENH-YIH obtained from the RF-based CWSRC model is analysed and compared with the results from Maxflow/Mincut and scanning lines. Considering that the weather conditions typically do not undergo drastically in a short period of time, this study represents the weather at half-hour intervals as representative of the weather for that period. Additionally, the weather conditions and flight trajectories for the prediction periods of the three case study days are depicted in Figures. 11, 12, and 13, respectively.

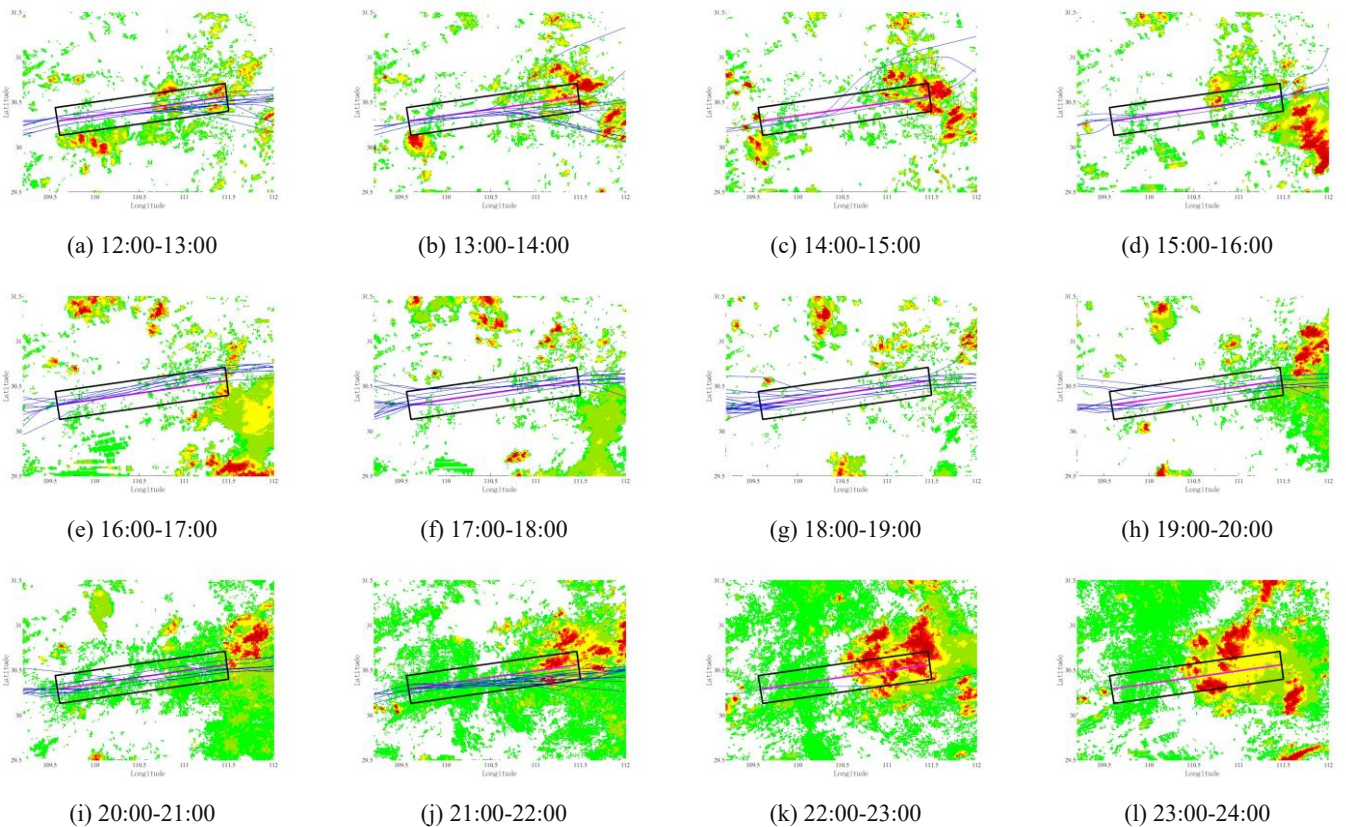


Figure 11. Flight tracks and weather on ENH-YIH from 12:00-24:00 on August 30, 2018

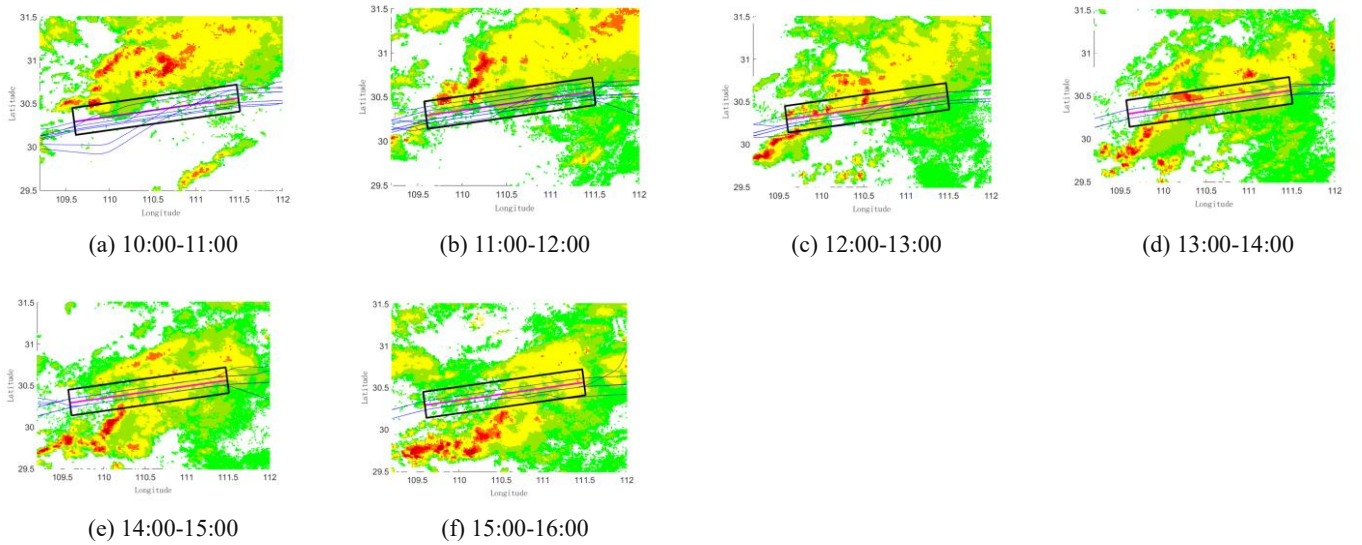


Figure 12. Flight tracks and weather on ENH-YIH from 10:00-16:00 on September 20, 2018

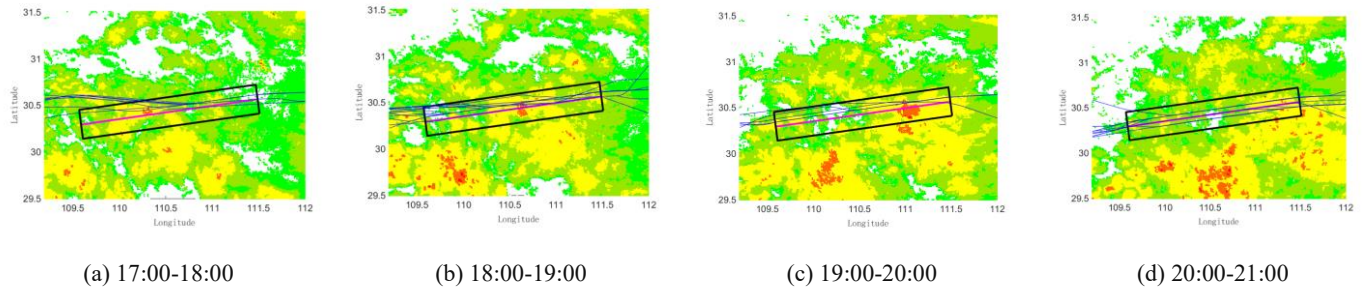


Figure 13. Flight tracks and weather on ENH-YIH from 17:00-21:00 on September 25, 2018

The prediction results for the three case study days are presented in Figure 14. Table 5 presents the overall performance of the models on the three case study days. We employed MAE, MSE, RMSE and R^2 as metrics to evaluate the predictive performance of the models. It is evident that the CWSRC model exhibits the lowest errors in metrics, indicating superior predictive performance compared to the Maxflow/Mincut and scanning line methods.

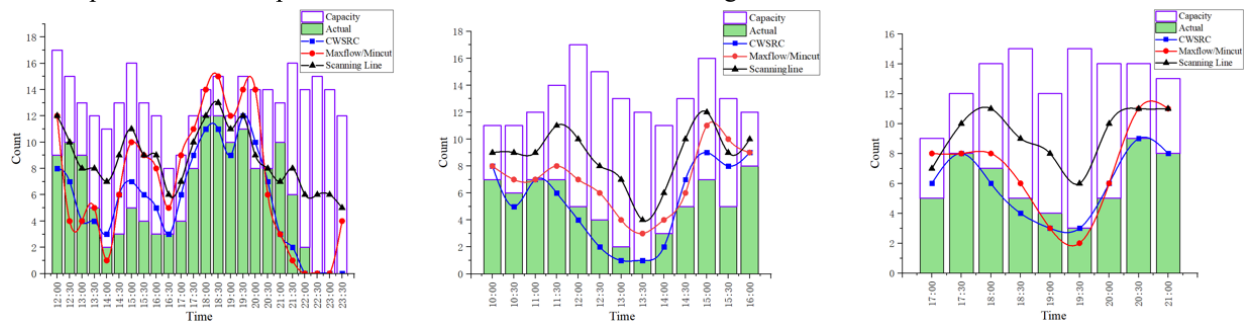


Table 5 Deviation of predicted results from actual values for CWSRC model with Maxflow/Mincut, scanning line

Indicators	CWSRC	Maxflow/Mincut	Scanning line
MAE	1.17	2.12	3.19
MSE	2.84	7.42	12.80
RMSE	1.53	2.56	3.57
R^2	0.74	0.26	-1.22

On August 30, 2018, the Wuhan terminal area experienced an extended period of convective weather. During this day, the weather in the sector route segment transitioned from convective to clear conditions and then back to convective weather within the daytime. The sector route segment was notably affected by the fluctuations in convective weather, as illustrated in Figure. 11. Given the significant and typical nature of the convective weather on this day, we conducted a detailed analysis of the August 30 case.

Figure. 11 illustrates the hourly meteorological changes and flight trajectories on the ENH-YIH segment from 12:00 to 24:00 on August 30. The blue line represents the flight trajectories, the purple line represents the route's centerline, and the black represents the route's boundaries. From 12:00 to 15:00, convective weather progressively emerged at the segment's right end, correspondingly diminishing the flight traffic. In 15:00-17:00 the weather was gradually dissipating, but flight frequency remains low, partly due to it being a less busy time of the day with naturally lower flight volumes, and potentially because of reduced flights scheduled for that day. From 17:00-20:00 there is no weather on the segment and the traffic volume is close to the statistical capacity of the sector under clear weather. From 20:00 to 24:00, a substantial amount of convective weather appears at the right end of the sector route segment, resulting in a gradual decrease in segment capacity, eventually leading to no flight traffic.

As depicted in the left subplot of Figure. 14, the results predicted by RF are in the highest degree of conformity with the actual flow changes, and the trend remains consistent across each time interval. From 12:00 to 13:00, the scanning line method's prediction closely matched the actual traffic volume. However, RF's prediction was lower, reflecting the partial coverage of the flight route by the weather. The scanning line method reflects the effect better at this time. In 13:00-15:00, there is a wide range of convective weather in the sector, RF and Maxflow/Minicut effectively capture this change, while the scanning line method's predictions significantly diverged from the actual values. Between 15:00 and 17:00, although convective weather was gradually dissipating, the capacity under the recorded clear weather continued to decline during this period. As a result, the predictions from these methods demonstrated a decreasing trend, with RF predictions aligning more closely with the actual values. Between 17:00 and 20:00, the segment experienced clear skies, with traffic volume approaching the segment's full capacity under optimal conditions. However, from 20:00 to 24:00, significant convective weather developed at the segment's right end. This led to a steady decline in capacity and, ultimately, the cessation of flight traffic by the end of the period. RF accurately predicted a value of 0 between 22:30 and 23:30, consistent with the actual results. Yet, the prediction for 21:00-21:30 was lower than the actual traffic volume. This was because during this time, convective weather only covered the upper right tail end of the sector. There was no convective weather in other areas and flights can still fly and cross the route from the lower right side. Although flights were still navigable, the convective weather features were more pronounced during this period, resulting in the underestimated prediction. The Maxflow/Minicut prediction aligned with RF but fell short in comparison, notably deviating from the actual values at 23:30. The scanning line method gave poor prediction under convective weather. In general, the scanning line method performs better in predicting under minimal or no weather conditions but fares poorly under convective weather. Maxflow/Minicut exhibits a preference for predicting accurately during convective weather but falls short of RF's predictive performance, especially when there's no weather. RF demonstrates consistently strong predictive capabilities under both conditions

Figure. 12 and 13 illustrate the meteorological conditions observed on selected days in September 2018. The flight corridor under study is located in a region where convective weather predominantly occurs during the summer months. Notably, the intensity, duration, and coverage of convective weather events in September are generally less extensive than those observed in August. From the right panel and the center subplot of Figure. 14, the CWSRC model demonstrated superior accuracy in traffic forecasting relative to conventional methodologies. Nonetheless, the case studies revealed intriguing phenomena worthy of further investigation: On September 20, 2018, between 12:00 and 13:00, although convective weather was present, the convective cells were dispersed and individually small, permitting direct flight traversal through these areas. A similar occurrence was documented in the case study from September 25, 2018, suggesting a pattern that merits deeper exploration in future studies.

4 Conclusion

This study utilized an improved RFE algorithm to extract meteorological features and developed CWSRC model to predict sector route capacity. The model employs a RF supervised regression algorithm, using ten-fold cross-validation for training and calibrating against recorded traffic data to compute capacity. The analysis of three case studies confirms that the CWSRC model's predictions closely match the actual flight volumes during convective weather events. Key findings are outlined below:

(1) The CWSRC model achieved favorable performance metrics during training and exhibited reduced errors compared to traditional methods in the case study, underscoring its accuracy and robustness.

(2) RF algorithm showed a higher percentage (93.403%) of residuals within the range $[-0.2, 0.2]$ on the validation set compared to KNN (86.03%) and MLP (77.37%) algorithms. This confirms that RF outperforms other machine learning algorithms in the CWSRC model.

(3) Due to the different traffic characteristics in different sector route segments, the influence of different weather factors on the prediction of sector route segment capacity is also different.

However, the model tends to underestimate capacity in scenarios with small and dispersed convective weather. It also tends to underestimate capacity when convective weather presenting at the beginning and end of sector route segment. This indicates a need for further research to enhance model accuracy: 1) incorporating adjacent airspace weather conditions as input features, 2) assessing the impact of dispersed, small convective cells that flights often traverse, and 3) considering the interaction effects of traffic volume on adjacent route segments.

References:

- [1] Civil aviation administration of china. (2022). *Civil Aviation Industry Development Statistics Bulletin for 2021* (Vol. 451). Beijing, China: Department of Development Planning. Retrieved from <https://www.mot.gov.cn/tongjishuju/minhang/202206/P020220607377281705999.pdf>.
- [2] FAA. (2022). *Air Traffic by the Numbers*. Washington: Federal Aviation Administration.
- [3] Rich DeLaura, M. An Exploratory Study of Modeling En Route Pilot Convective Storm Flight Deviation Behavior. In *12th Conference on Aviation Range and Aerospace Meteorology*. 2006.
- [4] DeLaura, R., Robinson, M., Pawlak, M., & Evans, J. Modeling convective weather avoidance in enroute airspace. In *13th Conference on Aviation, Range, and Aerospace Meteorology, AMS*, New Orleans, LA, 2008. Retrieved from <https://ci.nii.ac.jp/naid/20000753365>.
- [5] Klein, A., Cook, L., & Wood, B. Airspace availability estimation for traffic flow management using the scanning method. In *2008 IEEE/AIAA 27th Digital Avionics Systems Conference*, 2008, pp 3-B. URL: <https://doi.org/10.1109/dasc.2008.4702802>.
- [6] Song, L., Greenbaum, D., & Wanke, C. The impact of severe weather on sector capacity. In *8th USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)*, Napa, California, USA, 2009, pp 1-8. . Retrieved from http://www.atmseminar.org/seminarContent/seminar8/papers/p_075_W.pdf.
- [7] Callaham, M., DeArmon, J., Cooper, A., Goodfriend, J., Moch-Mooney, D., & Solomos, G. Assessing NAS performance: Normalizing for the effects of weather. In *4th USA/Europe Air Traffic Management R&D Symposium*. 2001, pp 3-7.
- [8] Klein, A., Jehlen, R., & Liang, D. Weather index with queuing component for national airspace system performance assessment. In *7th FAA/Eurocontrol ATM Seminar*. Spain, 2007, Retrieved from http://www.atmseminar.org/seminarContent/seminar7/papers/p_024_AAPM.pdf.
- [9] Chao, W., Xinyue, Z., & Xiaohao, X. (2008). Simulation Study on Airfield System Capacity Analysis Using SIMMOD. In *2008 International Symposium on Computational Intelligence and Design*. 2008, pp87-90, <https://doi.org/10.1109/iscid.2008.70>.
- [10] Sood, N., & Wieland, F. (2004). Total Airport and Airspace Model (TAAM) parallelization combining sequential and parallel algorithms for performance enhancement. In *Proceedings of the 2003 Winter Simulation Conference*. 2003, pp 1650-1655, <https://doi.org/10.1109/wsc.2003.1261615>.

-
- [11] Majumdar, A., Ochieng, W. Y., McAuley, G., Lenzi, J. M., & Lepadatu, C. (2004). The factors affecting airspace capacity in Europe: A Cross-Sectional Time-Series analysis using simulated controller workload data. *Journal of Navigation*, 57(3), 385–405. <https://doi.org/10.1017/s0373463304002863>.
- [12] Shi, F., Cheng, P., Geng, R., & Yang, M. (2012). An air traffic flow analysis system using historical radar data. In *Lecture notes in electrical engineering* (pp. 541–549). https://doi.org/10.1007/978-3-642-25766-7_72.
- [13] Mao, L., Peng, Y., Li, J., Guo, C., Kang, B., & Cao, Z. (2021). Random-forest based terminal capacity prediction under convective weather. *System Engineering-Theory & Practice*, 08, 2125-2136.
- [14] Chen, J., Cai, K., Li, W., Tang, S., & Fang, J. An Airspace Capacity Estimation Model based on Spatio-Temporal Graph Convolutional Networks Considering Weather Impact. In *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, 2021, pp 1-7. <https://doi.org/10.1109/dasc52595.2021.9594417>.
- [15] Wang, Y. Weather impact on airport arrival meter fix throughput. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference*, 2017, pp 1-10. <https://doi.org/10.1109/dasc.2017.8102133>.
- [16] Brito, I. R., Murca, M. C. R., Oliveira, M. D., & Oliveira, A. V. (2021). A machine learning-based predictive model of airspace sector occupancy. *AIAA AVIATION 2021 FORUM*. <https://doi.org/10.2514/6.2021-2324>.
- [17] Shijin, W., Jingjing, L., Jiahao, L., & Jiewen, C. (2021). Determination of Flight Deviation Thresholds based on the POT Model. *Aeronautical Computing Technique*, 51(3), 5.
- [18] Steiner, M., Bateman, R., Megenhardt, D., Liu, Y., Xu, M., Pocernich, M., & Krozel, J. (2010). Translation of Ensemble Weather Forecasts into Probabilistic Air Traffic Capacity Impact. *Air Traffic Control Quarterly*, 18(3), 229–254. <https://doi.org/10.2514/atcq.18.3.229>.
- [19] Genuer, R., Poggi, J., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>.

Competing interest declaration
Competing interests: The author(s) declare none.

