Transportation Research Part C 149 (2023) 104065

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

# Prediction of estimated time of arrival for multi-airport systems via "Bubble" mechanism

Lechen Wang [b,c], Jianfeng Mao [a,d,*], Lishuai Li [e,f], Xuechun Li [a], Yilei Tu [a]

[a] School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, China
[b] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, China
[c] Shenzhen Research Institute of Big Data, Guangdong, China
[d] Guangdong Provincial Key Laboratory of Big Data Computing, The Chinese University of Hong Kong, Shenzhen, Guangdong, China
[e] Faculty of Aerospace Engineering, Delft University of Technology, 2600 AA Delft, The Netherlands
[f] School of Data Science, City University of Hong Kong, Hong Kong, China

ARTICLE INFO

*Keywords:*
Multi-airport systems
Flight estimated time of arrival
Medium-term prediction
Trajectory pattern clustering
Sequence-to-sequence model
Spatio-temporal features

ABSTRACT

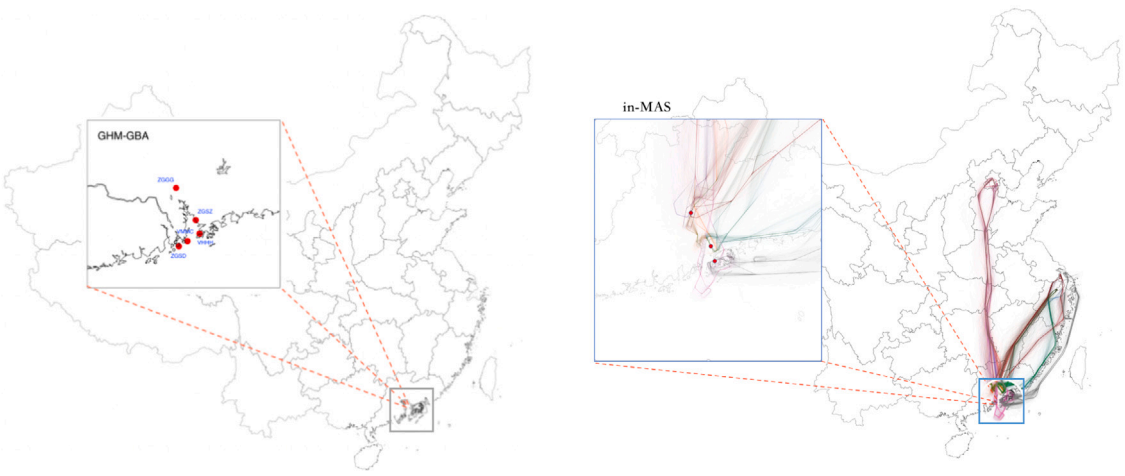
Predicting Estimated Time of Arrival (ETA) for a Multi-Airport System (MAS) is much more challenging than for a single airport system because of complex air route structure, dense air traffic volume and vagaries of traffic conditions in an MAS. In this work, we propose a novel "Bubble" mechanism to accurately predict medium-term ETA for a Multi-Airport System (MAS), in which the prediction of travel time of an origin–destination (OD) pair is decomposed into two stages, termed as out-MAS and in-MAS stages. For the out-MAS stage, Auto-Regressive Integrated Moving Average (ARIMA) is used to predict the travel time of a flight to reach the MAS boundary. For the in-MAS stage, we construct new spatio-temporal features based on clustering analysis of trajectory patterns facilitated by a novel data-driven hybrid polar sampling method. A sequence-to-sequence prediction model, Multi-variate Stacked Fully connected Bidirectional Long–Short Term Memory, is further developed to achieve multi-step-ahead predictions of in-MAS travel time for each trajectory pattern using the spatio-temporal features as input. Finally, the medium-term ETA prediction for an MAS is achieved by integrating the out-MAS and in-MAS prediction with the help of trajectory pattern prediction via random forest. A case study of predicting medium-term ETA for a typical MAS in China, Guangdong–Hong Kong–Macao Greater Bay Area, is conducted to demonstrate the usage and promising performance of the proposed method in comparison to several commonly used end-to-end learning methods.

## 1. Introduction

With developments of flight tracking technologies, like ADS-B trackers and GPS, tremendous real-time 4D trajectory data of flights (including latitude, longitude, altitude, and time stamp) prompts advanced traffic analysis methods in the new generation of Intelligent Air Transportation Systems (IATS). Among various traffic analysis research, estimated time of arrival (ETA), usually defined as when the flight touches down on the runway, is selected as one of the critical criteria in evaluating air transportation system performance by the International Civil Aviation Organization (Ma et al., 2022). The accurate ETA prediction is important to alleviate airport delays and provide insights and facts on air traffic flow management to air transportation authorities. Airlines and airports consider ETA as the essential parameter in flight recovery and runway configuration managements (Zhu et al., 2018). In

* Correspondence to: School of Data Science, The Chinese University of Hong Kong, Shenzhen, 2001 Longxiang Blvd., Shenzhen, Guangdong, 518172, China.
*E-mail address:* jfmao@cuhk.edu.cn (J. Mao).

https://doi.org/10.1016/j.trc.2023.104065
Received 21 April 2022; Received in revised form 21 December 2022; Accepted 16 February 2023
Available online 3 March 2023
0968-090X/© 2023 Elsevier Ltd. All rights reserved.




(a) The Five Airports in GHM-GBA

(b) Examples of Trajectories in The MAS

**Fig. 1.** Guangdong–Hong Kong–Macao Greater Bay Area (GHM-GBA).

addition, passengers also can increase their satisfaction and travel experience with accurate ETA information (Ayhan et al., 2018; Jie et al., 2019).

According to the time scale, prediction problems can separate into short-term and medium-term predictions (Hou and Li, 2016). Because short-term prediction is defined as a short period within an hour or even shorter, it provides evidence to tactical strategies, like trajectory sequencing improvements near airports (Ma et al., 2022; Wang et al., 2018). On the other hand, medium-term prediction that consists of a few hours usually assists in developing strategic management. For example, this kind of prediction before departure based on the flight plan is mostly used in flight recovery, fuel consumption, and traffic flow evaluation (Wang et al., 2018). Specifically, to balance the cost and benefits in solving delay problems, an accurate medium-term ETA provides more time and evidence about either deciding to cancel the flight or to substitute the delayed flight. However, compared with short-term predictions, medium-term prediction errors are relatively large (Zhang et al., 2022) since a longer time horizon results in the accumulation of prediction errors and uncertainties from the historical data.

For the single airport system, medium-term ETA prediction contains the flight's travel time in the entire flight procedure, which mainly contains the origin airport's terminal maneuvering area (TMA) with take-off and climb, en-route with high-altitude cruise and the destination airport's TMA with decent and landing. As another typical airport system, the multi-airport system (MAS) is defined as a group of two or more adjacent airports whose arrival and departure operations are highly interdependent (Atkins et al., 2011; De Neufville, 1995; Bonnefoy, 2008). Compared with the single airport system, the spatial proximity of MAS airports results in a shared TMA, in which the TMA of each airport interacts and overlaps with each other (Sidiropoulos et al., 2018). As a result, the flight to the MAS experiences from the TMA to en-route, then through the shared TMA to the MAS destination airport; in such, flight's travel time in the shared TMA has a larger variance than any other parts. One of the reasons is that shared TMAs contain a dense traffic volume. Operational interdependency among airports and limited capacity in MAS' shared TMAs result in more complex and dynamic route structures. Meanwhile, airport operations, coupling effects, and convective weather mainly cause real-time flight trajectories in shared TMAs to be vectored off to the original standard approach and departure procedures (Jie et al., 2019). For instance, standard arrival routes (STARs) are usually used to guide the flight from its TMA entering location to the runway. However, controllers' suggestions always result in the actual trajectory with possible deviation. In recent MAS studies, researchers proposed several ways to analyze the characteristics of air traffic networks and airport management in the MAS to alleviate uncertainties from trajectories (Murça and Hansman, 2019; Murça et al., 2018; Ren and Li, 2018; Ramanujam and Balakrishnan, 2009; Wang and Zhang, 2021). Human interventions in previous research limit the generalities of their methods. In addition, prior research in trajectory analysis of ETA predictions is mainly based on single airport systems under the short-term time scale (Wang et al., 2017, 2018; Jie et al., 2019; Wang et al., 2020b; Ma et al., 2022). However, to the best of our knowledge, there are no related studies to focus on medium-term ETA predictions in MAS and analyze how to embed uncertainties of trajectories and traffic conditions in the shared TMAs to MAS ETA prediction accuracy.

We propose a novel "Bubble" mechanism to the MAS to fill the research gap of medium-term ETA predictions and alleviate the uncertainties of deviation between actual trajectories and STARs in the shared TMA. In the "Bubble" mechanism, the MAS is treated as a bubble, and the prediction task is decomposed into two sub-tasks separately carried out outside bubble (the out-MAS prediction) and inside bubble (the in-MAS prediction), as interpreted in the case study of GHM-GBA in Fig. 1(b). For the outside bubble, the out-MAS prediction predicts travel time between the flight's wheel-off time and when the aircraft crosses the margin of MAS's shared TMAs (this is also the flight's entering location of the shared TMA). The in-MAS prediction forecasts the travel time between the entering location of the flight and the wheel-on time (this is also the flight's actual arrival time when it touches

down on the runway). For out-MAS travel time prediction, which is more regular and steady than in-MAS, ARIMA is applied to efficiently predict with satisfactory accuracy using only temporal information. It is insufficient to realize accurate in-MAS prediction with only temporal information within the MAS bubble. To cope with the complexity and uncertainties of trajectories of an MAS, we construct new spatio-temporal features by considering the MAS's network effects based on trajectory pattern clustering. To improve the generalities of trajectory sampling methods and abilities of trajectory description in prior research, a novel data-driven hybrid polar sampling method is initiated without human interventions by considering various turning points in 4D trajectories. To achieve multi-step-ahead predictions of in-MAS travel time, we further develop a sequence-to-sequence prediction model, Multivariate Stacked Fully-connected Bidirectional Long–Short Term Memory (MSB-LSTM). Finally, the medium-term ETA prediction for an MAS is achieved by integrating the out-MAS and in-MAS prediction with the help of trajectory pattern prediction via random forest. In Section 3.1, we provide more details about the design rationale of the "Bubble" mechanism based on the analysis of traffic characteristics of out-MAS and in-MAS stages.

In summary, the main contributions of this work are listed below:

1. Medium-term ETA prediction in MAS as the new issue in air transportation has been first solved in this work.
2. A novel prediction framework, "Bubble" mechanism, is designed to accurately predict the medium-term ETA of flights to an MAS. By using the bubble, uncertainties of trajectories and network effects in shared TMAs of MAS have been first time studied in medium-term ETA predictions.
3. A data-driven hybrid polar sampling method (using concentric circles from three different aspects to sample points from 4D flight trajectory) is developed to capture spatio-temporal characteristics of the MAS network and standardize flight trajectory data in clustering analysis. Our work is the first to consider flight behaviors in trajectory sampling and clustering analysis. Compared to previous sampling methods, the proposed sampling method is more automatic without constraints of airports and runway configurations (Wang et al., 2017, 2018, 2020b).
4. We select a typical MAS in China, the Guangdong–Hong Kong–Macao Greater Bay Area (GHM-GBA) in Fig. 1(a), as the case study in our work. Two types of comparative experiments are conducted. Firstly, to verify the importance of trajectory analysis in the shared TMAs, comparisons with end-to-end ETA predictions are provided to demonstrate the implementation and performance of the proposed "Bubble" mechanism. Secondly, we demonstrate how network effects in the MAS affect medium-term ETA predictions through comparative experiments with models without MAS analysis.

The rest of this paper is organized as follows. Section 2 reviews related literature in the fields of estimated time of arrival predictions in air transportation and trajectory pattern study. Section 3 presents the proposed bubble mechanism and elaborates its design rationale and workflow in details. The GHM-GBA case study and comparative experiments are exhibited in Section 4. Section 5 concludes this paper.

## 2. Literature review

### 2.1. Estimated time of arrival predictions in air transportation

As important indexes in transportation systems, ETA and travel time have been studied in urban and air transportation by abundant research.

Instead of ETA in air transportation, travel time as another representation has been paid extensive attention to urban transportation. Based on different road types, previous work typically separate into expressways (Chen et al., 2022; Zheng et al., 2020; Li and Chen, 2014; Dion and Rakha, 2006), roads in cities (Mendes-Moreira et al., 2015; Zahid Reza and Pulugurtha, 2019), and mixed structures of road sections (Petersen et al., 2019; Lin et al., 2022). However, the fixed structure of road types ignores diverse spatial information among vehicles in the same road section. Later, researchers proposed travel time predictions based on journeys. Current studies on journey travel time predictions mainly focus on forecasting transit time for vehicles' routes. Based on road segments in the route, previous work predicted travel time for each part and accumulated them together to compute the entire journey travel time (He et al., 2019, 2020). There has been much work using freeway traffic data (Wu et al., 2004; Innamaa, 2005), country highway data (Jang, 2016), and roadway data (Salamanis et al., 2016; Billings and Yang, 2006; Jenelius and Koutsopoulos, 2013). However, those published travel time prediction has tremendous diversities in problems and properties in air transportation. One major difference is that road structures in urban transportation limit existing of diverse route networks of vehicles. For instance, buses must follow bus lines, and cars follow urban road structures. Therefore, deterministic spatial information in urban transportation causes fewer uncertainties of vehicle trajectories in travel time predictions. Nevertheless, especially in TMA, unlike defined roads in urban transportation, airport operations or air traffic control cause trajectories vectored off initial flight plans (Ma et al., 2022; Jie et al., 2019). Uncertainties of trajectories and dynamic flight routes cause spatial dependencies confronting more challenges in air transportation prediction problems. Therefore, travel time prediction frameworks in urban transportation cannot be properly applied to air transportation.

In air transportation, existing research in ETA predictions can be divided into model-based and data-driven methods. Based on these two main categories, we further summarized related previous work in terms of types of methods, research topics, and airport properties in Table 1.

In Table 1, two main categories of ETA predictions, model-based and data-driven category, are further divided into three groups based on airport system, time scopes in predictions, and research targets. The first group illustrates previous work that used model-based methods only focused on short-term ETA predictions for single airport systems. They mainly studied ETA predictions of the

**Table 1**
Summary of previous research on ETA predictions in air transportation.

| Reference | Method type | | Airport system | | Time scope | | Prediction target | |
|---|---|---|---|---|---|---|---|---|
| | Model-based | Data-driven | Single airport | MAS | Medium-term | Short-term | En-route | TMA |
| Bai et al. (2016), Mueller et al. (2002), Krozel et al. (1999), Levy and Rappaport (2007), Wang et al. (2004), Huang et al. (2007), Roy et al. (2006), Wei et al. (2015) | ✓ | – | ✓ | – | – | ✓ | – | ✓ |
| Glina et al. (2012), Hong and Lee (2015), Wang et al. (2017, 2018), Jie et al. (2019), Wang et al. (2020b), Ma et al. (2022) | – | ✓ | ✓ | – | – | ✓ | – | ✓ |
| Strottmann Kern et al. (2015), Takács (2014), Zhu et al. (2018), Ayhan et al. (2018) | – | ✓ | ✓ | – | ✓ | – | ✓ | – |
| **our work: "Bubble" mechanism** | – | ✓ | – | ✓ | ✓ | – | ✓ | ✓ |

**Table 2**
Summary of previous research on data-driven methods of ETA predictions in air transportation.

| Reference | Prediction model | Endogenous features | | | External factors | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Temporal dependencies | Spatial dependencies | Dynamic Spatial dependencies | Traffic conditions | Weather conditions | Trajectory analysis | Flow Pattern uncertainties | Network effects |
| Glina et al. (2012) | Tree-based | – | ✓ | – | – | ✓ | – | – | – |
| Strottmann Kern et al. (2015), Takács (2014) | Tree-base,Ensemble | ✓ | ✓ | – | ✓ | ✓ | – | – | – |
| Zhu et al. (2018) | NN,Lasso,Ensemble | ✓ | ✓ | – | – | ✓ | – | – | – |
| Ayhan et al. (2018) | LSTM,Linear,Non-linear,Ensemble | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | – | – |
| Wang et al. (2017, 2018, 2020b) | NN | ✓ | ✓ | ✓ | – | – | ✓ | ✓ | – |
| Jie et al. (2019) | spatio-temporal model | ✓ | – | – | – | ✓ | – | – | – |
| Ma et al. (2022), Hong and Lee (2015) | Bi-LSTM,Linear | ✓ | ✓ | ✓ | – | ✓ | ✓ | ✓ | – |
| **our work** | ARIMA+seq2seq Bi-LSTM | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

flight inside the TMA. In the data-driven category, two sub-groups are described. The first partition in the data-driven category focused on short-term ETA predictions in single airports' TMAs. Most of the prior research in this partition tried to use trajectory analysis to alleviate uncertainties in TMA for ETA predictions. The second partition focused on medium-term ETA predictions in a single airport system. They usually studied the travel time of the flight before departure based on the flight plan from a strategic view.

For the model-based category in Table 1, ETA predictions usually forecast the flight ETA by considering aircraft performance, physical motions, and parametric trajectory models (Bai et al., 2016; Mueller et al., 2002; Krozel et al., 1999; Levy and Rappaport, 2007; Wang et al., 2004; Huang et al., 2007; Roy et al., 2006; Wei et al., 2015). Most of the previous work in this category first solved ETA predictions by computing possible trajectories of the flight using known parameters of aircraft performance and physical models with assumptions. Next, travel time can be computed using motion equations for computed trajectory points. Bai et al. compared two approaches to ETA predictions at downstream points of flights based on 4D trajectory generations (Bai et al., 2016). Jimmy et al. considered convective weather conditions and aircraft dynamic limits to solve ETA predictions (Krozel et al., 1999). In addition, Wang et al. considered a mathematical model with discrete event simulations to investigate miles-in-trail delays for flights (Wang et al., 2004). Some of them also consider stochastic motions of trajectories based on discrete time. For instance, Wei et al. proposed a state-dependent transition hybrid estimation algorithm by considering the descent stage of flights into discrete versions (Wei et al., 2015). Roy et al. also used discrete-time hybrid models and proposed combined multiple models for flight tracking (Roy et al., 2006). However, with ideal assumptions of flight trajectory models and airspace conditions, those computed trajectory points can be far diverse from actual trajectories (Bai et al., 2016; Mueller et al., 2002; Krozel et al., 1999; Levy and Rappaport, 2007; Wang et al., 2004). In this way, if any assumption is failed, the predicted ETA of flights will be unreliable (Ayhan et al., 2018). Moreover, in reality, besides aircraft performance and physical motions of flights, the time of arrival is always affected by other factors, like weather conditions, air traffic management, airport operations, and traffic network effects in TMA.

With the development of data-driven methods, machine learning and neural networks provide chances to researchers to initiate prediction models learning from history with fewer assumptions in reality. Compared with model-based methods in ETA predictions, researchers usually propose systematic prediction frameworks by considering diverse data types, like general flight information, weather conditions, and air traffic data. Based on prediction targets, previous work using data-driven methods can be divided into TMA ETA predictions (Glina et al., 2012; Hong and Lee, 2015; Wang et al., 2017, 2018; Jie et al., 2019; Wang et al., 2020b,a; Ma et al., 2022) and en-route ETA predictions (Strottmann Kern et al., 2015; Takács, 2014; Zhu et al., 2018; Ayhan et al., 2018), as

shown in Table 1. In addition, we also categorized prior research in data-driven ETA predictions in Table 2 from prediction models they used and different features they considered in their work to distinguish with our proposed prediction framework. As shown in Table 2, we consider features that have been used from endogenous and external aspects. In endogenous features, previous research is divided into three groups in terms of temporal dependencies, spatial dependencies, and dynamic dependencies. Specifically, spatial dependencies mainly describe features considering traffic networks and flight route structures, and dynamic spatial dependencies stand for that changes in flight trajectories have been included. For external factors, we group prior research from five different data types in Table 2. Notice that, we use traffic conditions to stand for related traffic states used in features, like traffic flow. In addition, trajectory analysis consists of studies from either point-wise analysis or route-based analysis.

In TMA ETA predictions, previous work tried to propose prediction models to forecast the time of runway arrival from the entry point of the flight. For example, Glina et al. proposed quantile regression forests to predicate ETA for individual flights based on the distances between locations in TMA to the airport and generated conditional probability distribution for the ETA (Glina et al., 2012). Their work mainly focused on correlations between flight tracking information from spatial dependencies in flight routes and ETA. In addition, weather and runway utilization also had been considered. However, spatial dependencies in the TMA and temporal dependencies behind traffic conditions are ignored.

As unique properties in TMA, various entry points and flight operations result in flight flow from different directions interacting with each other and merging based on runway configuration selections (Ma et al., 2022). In addition, high traffic volume in the limited airspace always affects trajectories vector off standard procedures (Hong and Lee, 2015). To solve this challenge in TMA ETA predictions, researchers have paid extensive attention to ETA predictions with 4D trajectory analysis. In their work, Trivedi et al. demonstrated that some machine learning clustering methods could significantly improve ETA predictions (Trivedi et al., 2015). Based on these findings, the trend of the prediction framework started to combine clustering into ETA predictions to consider dynamic spatial information and uncertainties caused by trajectories (Hong and Lee, 2015; Wang et al., 2017, 2018, 2020b,a; Ma et al., 2022).

Hong et al. proposed trajectory clustering with dynamic time wrapping in their ETA prediction framework. For each clustered major trajectory pattern, they predicted the flight's ETA from the entry point of TMA (Hong and Lee, 2015). Unfortunately, only flights from the same entry point were considered in their work. In addition, only considering trajectory points from the temporal aspect has a great chance of missing important merging points and turning arc of flight trajectories because sampling trajectory points from time scale cannot capture flight behavior information.

Wang proposed a series of work considering more details in trajectory clustering (Wang et al., 2017, 2018, 2020b). In the first version of their work (Wang et al., 2017, 2018), they sampled flight trajectories based on their defined reference points, like the corner of the TMA or some reference points of the airport. For each trajectory, they sampled trajectories using concentric circles and provided features of distance and angular position of sampling points respected to those defined references. However, this kind of sampling method and feature engineering highly depends on human-defined references, which limits the generality of their methods. Moreover, sampling trajectory using concentric circles still loses important flight behavior information. For ETA predictions, they proposed fully connected neural networks for each clustered trajectory using DBSCAN without considering spatio-temporal dependencies. In their second version, Flights were clustered based on different runway-in-use. As a result, as long as flights land on the same runway, they are clustered into the same partition no matter where the flight enters the TMA and in which direction the flight merges to the runway (Wang et al., 2020b). In the end, the ensemble learning technique was used in ETA predictions. As shown in Table 2, they only provided general flight information, like the location and heading direction of the flight, as their features in ETA predictions. Related traffic conditions, weather information, and network effects were ignored.

Compared to TMA ETA predictions, there are few works on en-route ETA predictions in the data-driven category (Strottmann Kern et al., 2015; Takács, 2014; Zhu et al., 2018; Ayhan et al., 2018), as shown in Table 1. Because en-route ETA prediction is usually before the flight's actual departure, this medium-term ETA prediction confronts more uncertainties than short-term predictions. In their work, Takacs et al. spent effort on feature engineering and extracted 56 features from traffic data (Takács, 2014). Their methods consisted of six sequential stages of ridge regression and gradient boosting machine. Kern et al. tried to improve ETA prediction accuracy by iteratively adding features to the model (Strottmann Kern et al., 2015). The combination of flight information, weather data, and air traffic data was applied to the random forecast. However, information on departure routes had been ignored. Moreover, convective weather conditions also had been paid extensive attention (Zhu et al., 2018). However, dynamic spatial dependencies caused by flight trajectories had been ignored in those works, as shown in Table 2. In Ayhan's work, they collected richer features based on 3D grid trajectory points, and trajectory analysis was used to analyze changes and uncertainties behind the flight trajectory. Abundant regression models were applied for ETA prediction of commercial flights in Spain Ayhan et al. (2018). In their dataset, adaptive boosting and gradient boosting perform better than others. Even though some of the OD pair consists of MAS, they did not consider traffic network effects and specific analysis of spatial dependencies behind the MAS and ignored uncertainties of trajectories in the shared TMAs.

In summary, compared to previous work in air transportation ETA predictions, the proposed framework in this work, the "Bubble" mechanism, is distinguished from others with the following aspects:

1. "Bubble" mechanism is the first prediction framework that solves medium-term ETA prediction to the MAS by considering network effects and uncertainties of trajectories in the shared TMA, as shown in Table 1.
2. Without using flight plans and costly feature engineering of each point along the flight routes in previous work, our approach provides a more cost-effective option based on the different complexities of flight procedures. Even though traffic conditions and trajectory in the shared TMA are unknown before departure, our approach can still capture details in MAS over multiple time steps by first using seq2seq framework in MAS ETA predictions, as shown in Table 2.

**Table 3**
Summary of previous research on trajectory pattern analysis of ETA predictions in air transportation.

| Reference | Trajectory sampling | | | External features in trajectory description | | | |
|---|---|---|---|---|---|---|---|
| | Temporal aspect | Spatial aspect | Flight behaviors | Human intervention | Entry locations | Landing locations | Turning Arcs |
| Hong and Lee (2015) | ✓ | – | – | – | – | – | – |
| Wang et al. (2017, 2018), Ma et al. (2022) | – | ✓ | – | ✓ | – | – | – |
| Wang et al. (2020b) | – | ✓ | – | – | – | ✓ | – |
| **our work** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

3. Compared with previous work, trajectory analysis and traffic flow patterns have been first analyzed in MAS medium-term ETA predictions. In addition, we embed the temporal portion of air routes in shared TMAs for each clustered pattern as MAS network effects in features.
4. Compared to previous trajectory analysis in TMA ETA predictions, our proposed sampling method and trajectory description approach consider more details of flight behaviors, like turning points, entry, and merging points. The approach also does not involve human intervention and constraints of airports or specific runway configurations.

To further demonstrate the fourth aspect mentioned above, we provide a literature review of trajectory flow pattern analysis in the next section.

### *2.2. Trajectory pattern study*

Trajectory pattern clustering methods have been extensively used in air transportation research. Gariel et al. developed a framework to monitor airspace based on a density-based clustering method by learning typical patterns of flight trajectories and used principal component analysis (PCA) for dimension reduction (Gariel et al., 2011). Murca et al. also developed a framework for traffic flow patterns identification, characterization, and prediction, based on a density-based trajectory clustering scheme, to match new flight trajectories with the airspace structure identified in the previous module (Murça et al., 2018). In their work, Ayhan et al. proposed a Divide-Cluster-Merge framework to divide the flight based on three main phases, *i.e.*, climb, en-route and descent (Ayhan and Samet, 2015). From the domain knowledge, different properties and motivations of phases result in different functional trajectories. To cluster the entire trajectory, they separately clustered 3D points in each phase using K-means, then merged centroids of clusters in three phases. Compared with Stefan and Ayhan (Ayhan and Samet, 2015; Stefan et al., 2012), to introduce effects from the complex air traffic networks and alleviate workloads of analyzing each trajectory point to provide spatial information, we use the "Divide and Conquer" from the geographical aspect to provide spatial information of the region-based traffic networks. However, that mentioned literature did not aim to solve ETA predictions using trajectory pattern analysis. In addition, none of those references considered flight behaviors in detail in their trajectory clustering.

Trajectory pattern analysis can help extract traffic conditions and route network structure information to facilitate ETA predictions. Trajectory patterns in the TMA of multi-airport systems reflect recurrent utilization patterns of airspace (Murça and Hansman, 2019; Gariel et al., 2011). Identification of the trajectory patterns characterizes specific situations of air routes. Spatial information extracted from trajectory patterns can be used to construct features to represent the capacity of each airspace and play a central role in predicting ETA. We summarize related previous work on TMA ETA predictions by considering trajectory analysis in their prediction frameworks in terms of trajectory sampling methods they used and external features they proposed in trajectory description, as shown in Table 3. In the trajectory sampling category, prior research is divided into groups based on the temporal aspect, the spatial aspect, and flight behaviors. For the external features category, our work is distinguished from those works in the following aspects: (1) human intervention stands for human-defined references used in trajectory description; (2) entry and landing locations indicate starting and ending points of the flight trajectory in TMA; (3) turning arcs describe features in trajectory description that consist of turning behaviors of flights.

For the ETA prediction, the primary purpose of trajectory clustering is to recognize vectoring trajectory patterns (Hong and Lee, 2015), so the spatial shape of the flight trajectory is more important than general information along the air route. As the first group in Table 3, Hong et al. sampled trajectories from the entry fix in TMA from the temporal aspect, which means trajectory points are sampled based on defined time duration. As a result, a tiny misalignment in time may cause a far distance between flights due to different flight speeds. To tackle this issue, they used dynamic time warping (DTW) to compute distances between sampled flight trajectories. However, this sampling method from temporal aspects requires flights that have similar transit times and from fixed entry locations. Therefore, such a combination is inappropriate for TMA trajectory analysis; in this approach, the defined sample frequency based on time duration may result in missing information for the flight with longer transit time. In addition, if the point merge system is applied in airport operation, only considering temporal aspects in the sampling method cannot describe the spatial shape of the turning arcs in the flight trajectory.

To address mentioned issues in the temporal sampling method and DTW, Wang et al. proposed a series of works by considering spatial sampling methods and describing trajectories with human-defined referenced points (Wang et al., 2017, 2018), as shown

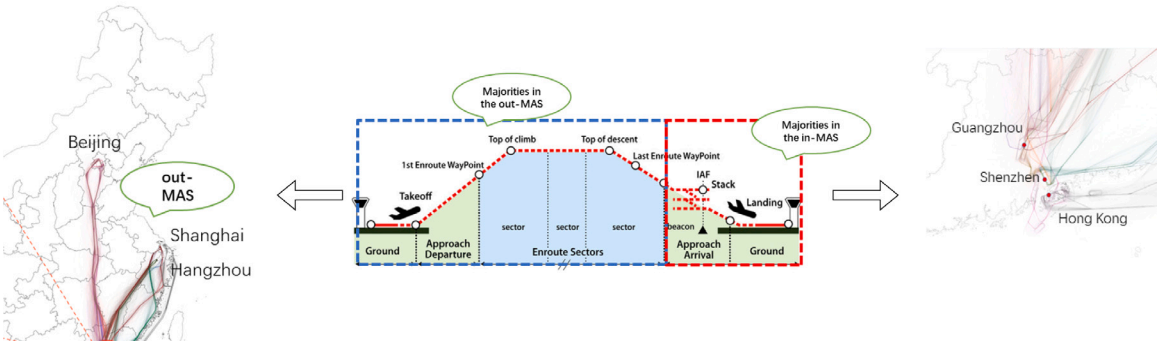



**Fig. 2.** out-MAS and in-MAS stages in the example of the MAS of GHM-GBA.

**Table 4**
The comparison of travel time in different stages.

| The OD Pair | ZBAA-ZGSZ | | | ZBAA-ZGGG | | | ZBAA-VHHH | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean(min) | STD(min) | CV | Mean(min) | STD(min) | CV | Mean(min) | STD(min) | CV |
| The Entire Journey | 170.65 | 10.53 | 6.17% | 160.00 | 10.90 | 6.81% | 185.90 | 11.12 | 5.98% |
| out-MAS | 130.30 | 5.84 | 4.48% | 129.80 | 6.10 | 4.69% | 132.22 | 5.23 | 3.95% |
| in-MAS | **40.35** | **8.71** | **21.6%** | **30.57** | **8.78** | **28.7%** | **53.70** | **10.37** | **19.3%** |

in Table 3. In their work, they used concentric circles to sample flight trajectories, then provided features in trajectory description by computing distance and angular information concerning reference points. Finally, DBSCAN was used to compute the distance between sampled trajectories. Ma et al. (2022) also used this framework in their work to analyze flight trajectories. However, human intervention limits the generality of their methods; in such an approach, human-defined reference points lack interpretations and cannot be duplicated in other airport systems. Later, Wang et al. proposed another work using the idea of runway-in-use to sample flight trajectories (Wang et al., 2020b). Even though it alleviates the issue of generality, only using landing locations in trajectory description confronts the same issue as the temporal sampling method. If the point merge system is applied, features of the shape of turning arcs and sequence legs are still missed.

As shown in Table 3, the proposed hybrid polar sampling method addresses those mentioned issues and is distinguished from previous work with the following points: 1) to the best of our knowledge, our work is the first to consider flight behaviors in flight trajectory sampling, which consists of features related to sequence legs and turning arcs used in point merge systems; 2) without any human interventions, the proposed trajectory analysis framework is more automatic than previous work.

## 3. Methodology

### *3.1. "Bubble" mechanism*

In this section, we will first illustrate the design rationale of the "Bubble" mechanism based on the analysis of the traffic characteristics of out-MAS and in-MAS stages using 4D trajectory historical data. The detailed workflow of the proposed bubble mechanism will be presented afterwards.

#### *3.1.1. Design rationale of bubble mechanism*

The Bubble mechanism can be viewed as a "Divide and Conquer" technique and is proposed to cope with the two difficulties arose in the prediction of medium-term ETA to an MAS mentioned in Section 1, *i.e.*, medium-term prediction horizon and complexities of the MAS.

The flight procedures of the OD pairs to an MAS can be partitioned into two stages: out-MAS and in-MAS stages, as shown in the example of those OD pairs to the MAS of GHM-GBA in Fig. 2. The out-MAS stage contains approach departure and en-route section outside an MAS and the in-MAS stage corresponds to approach arrival inside an MAS. In common practice, there is no need for holding in approach departure and relatively sparse traffic volume at a cruising altitude in en-route section, which results in more regular air routes and flight patterns as shown in Fig. 2. In contrast, because of dense traffic volume and complex air route structure in an MAS, approach arrival of an airport in an MAS requires dynamic and diverse flight patterns in the in-MAS stage. Therefore, the traffic situation in the out-MAS stage is much more regular and predictable than the one in the in-MAS stage.

This property can be further verified using historical trajectory data of several typical OD pairs as shown in Table 4, which lists the mean and standard deviation of the travel times in both out-MAS and in-MAS stages in minutes. To further demonstrate the property and variations of travel time in different segments, we compute Coefficient of Variance (CV) for each OD pair in Table 4. It can be observed, the mean travel time of in-MAS stage is much shorter than the one of out-MAS stage, but the variance of in-MAS

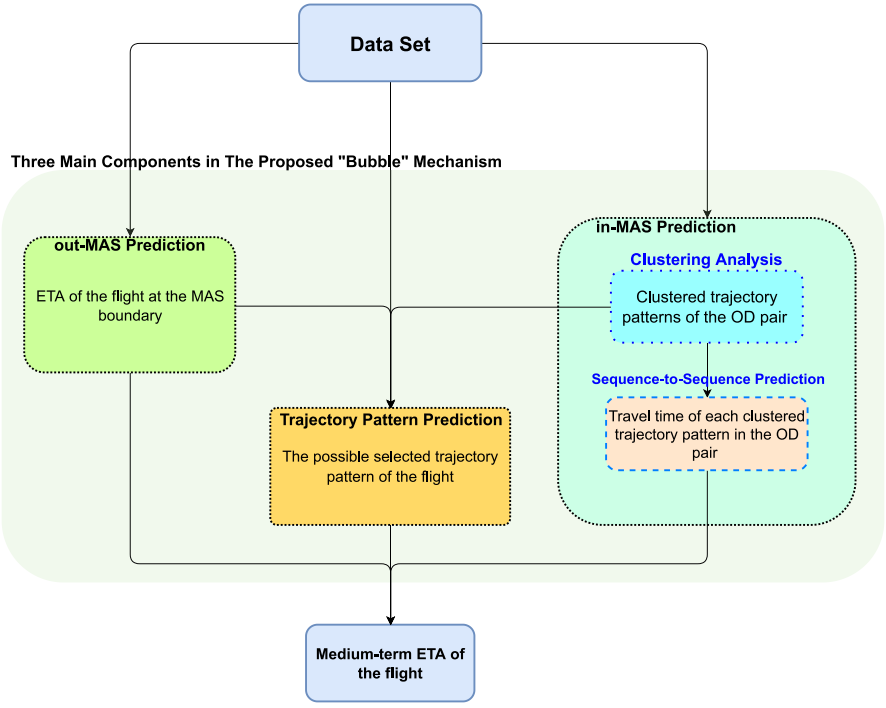



**Fig. 3.** Three main components of the "Bubble" mechanism.

stage is much higher than the one of out-MAS stage. For instance, although the travel time of in-MAS stage only accounts for less than 25% of the entire journey of ZBAA-ZGSZ, the standard deviation (STD) is almost 1.5 times the one in out-MAS stage. It becomes more obvious in ZBAA-VHHH, in which the in-MAS STD is almost 2 times the out-MAS STD. Therefore, the out-MAS stage is much more predictable than the in-MAS stage.

The property identified above enables the Bubble mechanism proposed here to achieve both accuracy and efficiency. The so-called "Bubble" separates an MAS from its outside airspace as shown in Fig. 1(b) and decomposes the prediction of medium-term ETA into two stages, out-MAS and in-MAS stages. Since the traffic situation in the out-MAS stage is much more regular and predictable, a classical prediction model, such as ARIMA, is sufficient enough to efficiently and accurately predict the out-MAS travel time. Then, we can focus on the complexities of an MAS and develop a sophisticated model to accurately predict the in-MAS travel time. The performance of the proposed Bubble mechanism in comparison to some commonly used end-to-end learning methods can be found in the experimental results in Section 4.

#### *3.1.2. Workflows of bubble mechanism*

The proposed "Bubble" mechanism contains three main components, out-MAS prediction, in-MAS prediction and trajectory pattern prediction, as sketched in Fig. 3. The out-MAS prediction outputs the prediction of the flight travel time outside MAS and the time reaching the MAS boundary. The in-MAS prediction generates flight trajectory patterns and makes a multi-step-ahead prediction of travel time for each trajectory pattern inside MAS. The trajectory pattern prediction forecasts the selected trajectory pattern and integrates it with the outputs from the other two components to produce the medium-term ETA prediction.

The detailed workflow of the Bubble mechanism inside the three components can be found in Fig. 4. In the out-MAS prediction, due to the small variance of travel time in the out-MAS mentioned above, temporal information will be sufficient for a good prediction and ARIMA can be used to efficiently and accurately make the out-MAS travel time prediction. In the in-MAS prediction, since flight patterns in an MAS is much more dynamic and diverse than the ones outside an MAS, to enhance the prediction accuracy, we construct new spatio-temporal features based on clustering analysis of trajectory patterns facilitated by a novel data-driven hybrid polar sampling method, which standardizes and resamples the trajectory tracking data using both turning-point-based and uniform polar sampling. A sequence-to-sequence prediction model, Multi-variate Stacked Fully connected Bidirectional Long-Short Term Memory, is further developed to achieve multi-step-ahead predictions of in-MAS travel time for each trajectory pattern using the spatio-temporal features as input. Finally, the medium-term ETA prediction for an MAS is achieved by integrating the out-MAS and in-MAS prediction with the help of trajectory pattern prediction via random forest.

### *3.2. out-MAS prediction: ARIMA*

ARIMA is a widely used time series model and capable of capturing temporal characteristics of relatively steady and regular traffic in the out-MAS stage as shown in the experimental results below. A time series of the average travel time of the OD pair

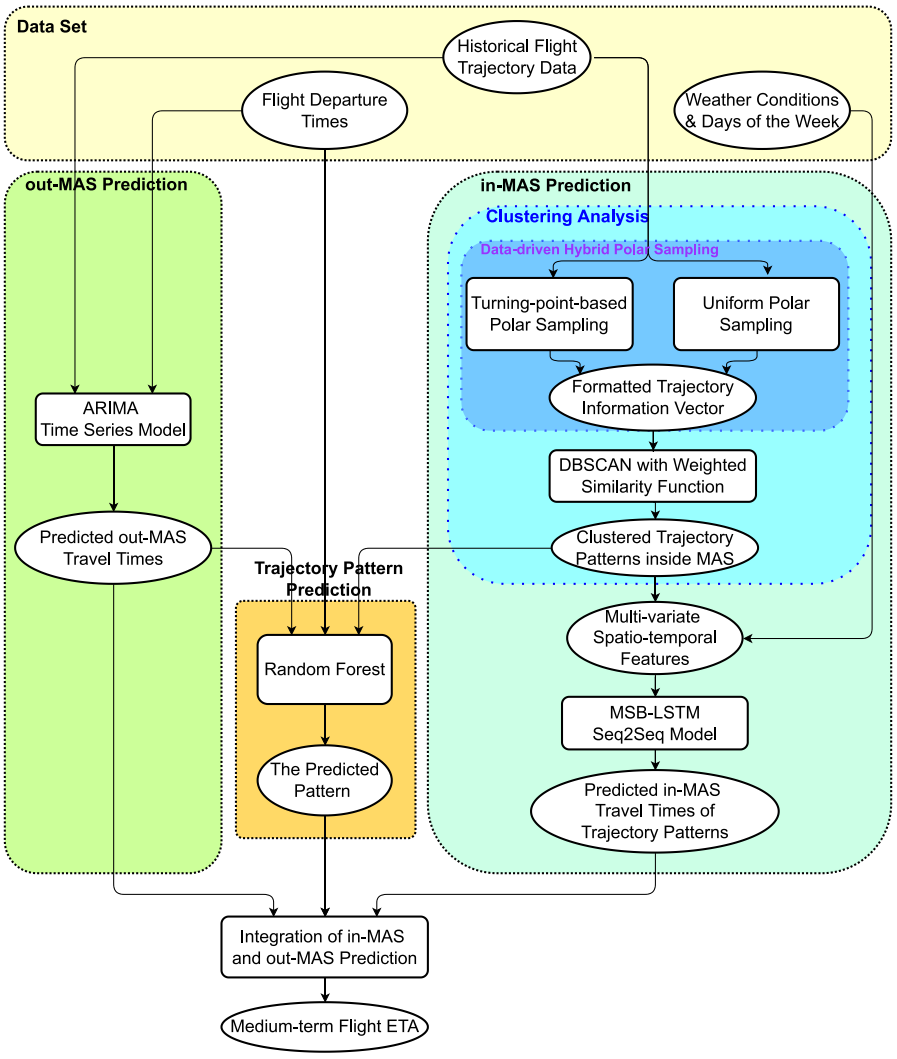



**Fig. 4.** The workflow of The "Bubble" mechanism.

versus flight departure time $\{x_t\}$ is accordingly constructed using historical 4D trajectories, where $x_t$ represents the average en-route travel time of flights $f$ the OD pair departing in the $t$th time segment. For the configuration and the training process of ARIMA in this study, we provide details in Section 4.7.

### 3.3. in-MAS prediction: Trajectory flow pattern clustering

We aim at predicting the travel time of trajectory patterns within an MAS. Different from out-MAS prediction, only temporal information of 4D Trajectory data is not sufficient for a good prediction because of the complex air route network structure, aggregated high density of traffic volume in the limited and shared airspace, and intersections among various approach and departure procedures of multiple airports. To improve the prediction accuracy, a trajectory pattern clustering method is developed to extract spatial information of 4D trajectory data and facilitate the construction of spatio-temporal features.

Each trajectory $\mathbf{F}_j$ is a sequence of waypoints:

$$\mathbf{F}_j = \{ (x_{j,t},\ y_{j,t},\ \theta_{j,t}),\quad t = 1, 2, \ldots, T \}, \tag{1}$$

where $j$ is the flight index, $t$ is the time index, $(x_{j,t}, y_{j,t}, \theta_{j,t})$ are the latitude, longitude and heading direction of flight $j$ at time $t$. These trajectories are grouped according to combinations of airports and modes (departure or arrival) in an MAS. In the case study of GHM-GBA, there are 10 groups corresponding to the combinations of 5 airports, {ZGSZ, ZGGG, VHHH, VMMC, ZGSD} and 2 operation modes {arrival, departure}. Since different combinations of airport and operation mode implies different trajectory pattern, we can separately conduct trajectory pattern clustering for each combination.

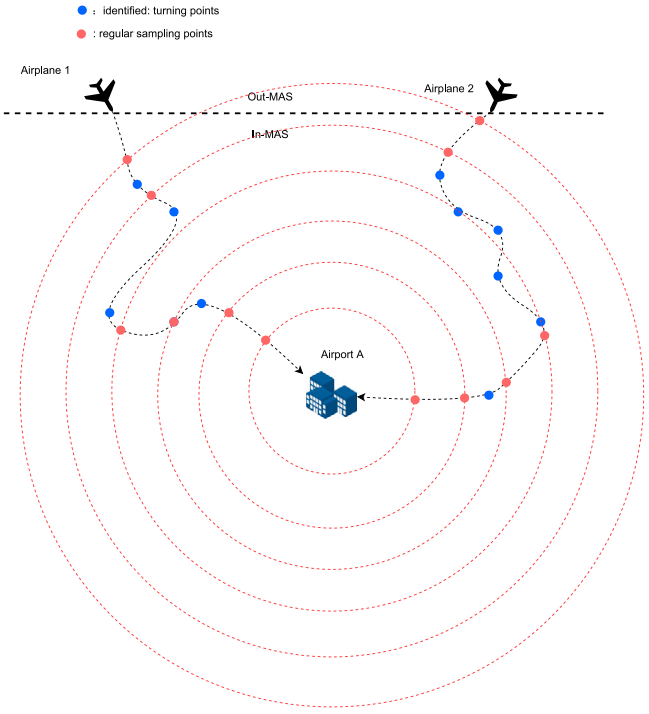



**Fig. 5.** Illustration of identified turning points.

Before clustering trajectories, a data-driven spatial sampling method, *Hybrid Polar Sampling*, is developed to standardize trajectories into formatted trajectory information vectors without relying on user-defined reference points, as shown in Fig. 4. The hybrid polar sampling mainly includes two parts: uniform polar sampling and turning-point-based polar sampling.

The uniform polar sampling is to sample flight trajectories $\mathbf{F}_j$ defined in (1) with equally spaced concentric rings whose center is the corresponding flight arrival or departure airport in an MAS. In the case study, the radius of these equally spaced concentric rings varies from 15 km to 300 km in increments of 15 km and the uniformly sampled points for each flight $j$ is the set $\mathbf{R}_j$ below

$$\mathbf{R}_j = \left\{ (x_{j,t}, y_{j,t}) : \sqrt{(x_{j,t} - x^a)^2 + (y_{j,t} - y^a)^2} \approx d,\ d = 15, 30, \dots, 300 \right\}, \tag{2}$$

where $(x^a, y^a)$ is the location of the corresponding airport $a$. The representativeness using uniform polar sampling is largely dependent on the sampling frequency. The larger the sampling frequency is, the better the representativeness is, but the higher the computational complexity is. Uniform polar sampling may miss certain parts of a flight trajectory containing important features that can differentiate it from others if the sampling frequency is not sufficiently high.

To improve sampling efficiency and effectiveness, the turning-point-based polar sampling is designed to adaptively sample more frequently in feature-rich areas according to historical trajectory data.

The feature-rich areas generally correspond to the ones where trajectories or curves vary a lot. The area of the air space where flights fly straight is feature-less, while the one where flights turn is feature-rich. The feature-rich areas, *i.e.*, flight turning areas, can be determined by clustering turning-point locations. The turning-point locations of flight $j$, $\mathbf{T}_j$, are identified as the ones whose change of heading direction is greater than $\Delta\theta$, which is set as 15° in the case study, *i.e.*,

$$\mathbf{T}_j = \{(x_{j,t}, y_{j,t}) : \ \theta_{j,t} - \theta_{j,t-1} > \Delta\theta,\ t = 1, 2, \dots, T\}.$$

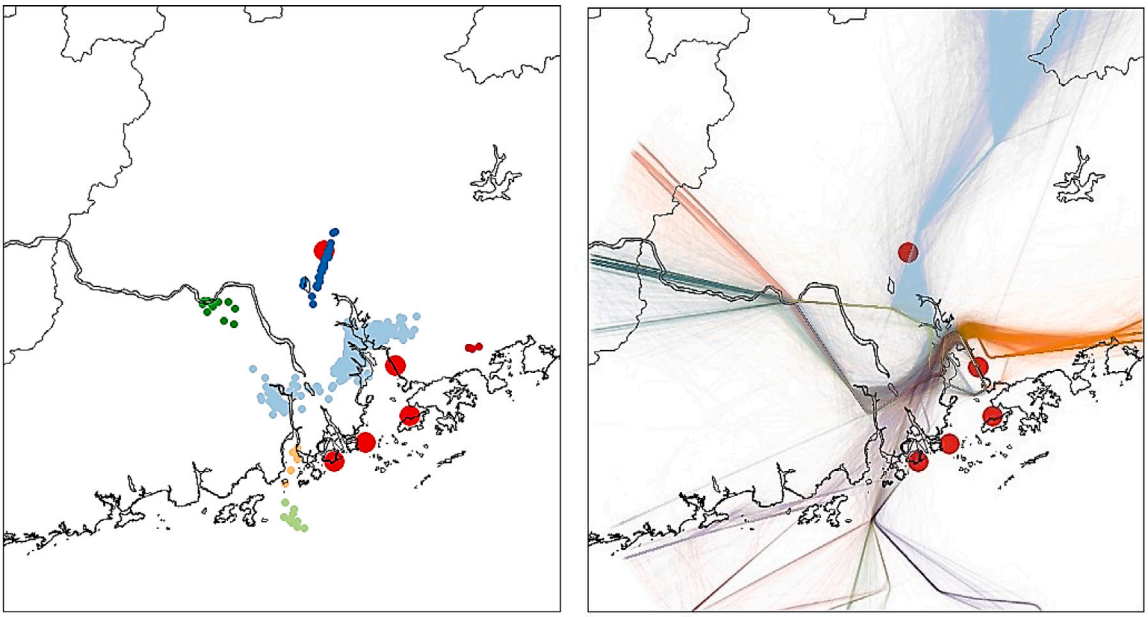
An example of the collection of turning points in flight trajectories is depicted in Fig. 5. Then, turning point clusters can be obtained using DBSCAN. Fig. 6(a) is an example of the turning-point clustering result of ZGSZ under arrival mode.

After clustering turning-points, the distance $r_k$ between the centroid of each cluster $k$ and the corresponding airport is calculated. Flight trajectories $\mathbf{F}_j$ is further sampled by the concentric rings with radius of $r_k$ for all $k$ and the turning-point-based sampled points is the set $\mathbf{D}_j$ below

$$\mathbf{D}_j = \left\{ (x_{j,t}, y_{j,t}) : \sqrt{(x_{j,t} - x^a)^2 + (y_{j,t} - y^a)^2} \approx r_k, \forall k \right\}. \tag{3}$$

Based on $\mathbf{R}_j$ in (2) and $\mathbf{D}_j$ in (3), we can construct the formatted trajectory information vector $\mathbf{V}_j$ for each flight $j$,

$$\mathbf{V}_j = \left[ \mathbf{R}_j,\ \mathbf{D}_j,\ \mathbf{E}_j \right], \tag{4}$$



(a) Turning-point Clustering of ZGSZ (b) Trajectory Pattern Clustering of ZGSZ

**Fig. 6.** Clustering analysis: Turning-point clustering, trajectory pattern clustering.

where $\mathbf{E}_j = \{(x_{j,e}, y_{j,e})\}$ is the location of flight $j$ entering or exiting an MAS.

We apply DBSCAN to cluster flight trajectories based on their corresponding formatted trajectory information vectors $\mathbf{V}_j$ in (4). DBSCAN is a very popular clustering method suitable for data sets with noise (Ester et al., 1996), which enables the identification of significant patterns in the presence of abnormal trajectory profiles. Parameters *Minpts* and $\epsilon$ can be tuned by the $k$-distance algorithm (Ester et al., 1996). Moreover, the similarity (distance) between two flight trajectories $i$ and $j$ is defined as the weighted L2-norm of the difference between $\mathbf{V}_i$ and $\mathbf{V}_j$:

$$\left\|\mathbf{V}_i - \mathbf{V}_j\right\|_2 = w_R\left\|\mathbf{R}_i - \mathbf{R}_j\right\|_2 + w_D\left\|\mathbf{D}_i - \mathbf{D}_j\right\|_2 + w_E\left\|\mathbf{E}_i - \mathbf{E}_j\right\|_2$$

where $w_R, w_D, w_E$ are the parameters to prioritize the importance of three sets of sampled points in reflecting spatial characteristics of flight trajectories. The point in $\mathbf{E}_j$ marks the gate information of an MAS, which has the highest priority, *i.e.* $w_E$ is the largest. The points in $\mathbf{D}_j$ corresponding to the points in feature-rich areas are more representative than the ones in $\mathbf{R}_j$ for differentiating with other flight trajectories, *i.e.*, $w_D > w_R$. Fig. 6(b) exhibits an example of identified trajectory patterns for ZGSZ under arrival mode.

### 3.4. in-MAS: Spatio-temporal features and MSB-LSTM

In the limited and shared TMA of an MAS, traffic and weather situation along air routes are critical factors impacting travel time. After clustering trajectory patterns, we can specify more information about those situations for different trajectory patterns in an MAS. Only those trajectory patterns affect each other if they share same parts of the airspace in an MAS. For the example in Fig. 7, in which we exhibit ZGSZ arrival trajectory patterns as an example, the trajectory pattern in blue traverses the parts of airspace highlighted at the top right corner, which are shared with the trajectory pattern in green, in that way the in-MAS travel time of the two trajectory patterns are correlated. Moreover, if we plot trajectory patterns for other airports in the same figure, we can conclude that the highlighted area in Fig. 7 will contain more diverse patterns at the same time. This inspires us that effects to each trajectory pattern from the MAS network-scale can be introduced by traffic flows of flights in the network-level and traffic flows in the trajectory pattern-level, which are features $c_{t,l}$ and $r_{t,l}$.

To capture these traffic flow situations and environmental conditions that may affect the flights using trajectory pattern $l$, we construct a multi-variate spatial–temporal feature $I_{t,l}$ below for each pattern $l$ at each time $t$,

$$I_{t,l} = \left[c_{t,l},\ r_{t,l},\ p_{t,l},\ w_{t,l},\ d_t\right],$$

where

- $c_{t,l}$: The total number of flights sharing the airspace inside an MAS with the flights using trajectory pattern $l$ during time slot $t$;

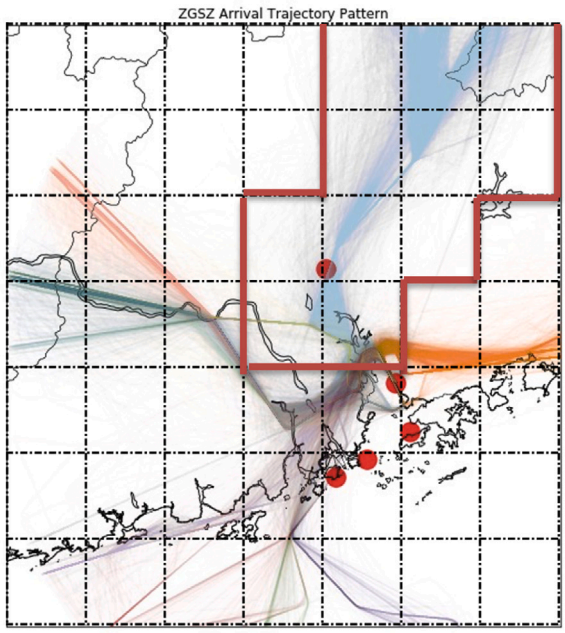



**Fig. 7.** Illustration of spatio-temporal features in GHM-GBA.

- $e_{t,l}$: The total number of flights using the pattern $l$ during time slot $t$;
- $r_{t,l} = \frac{e_{t,l}}{c_{t,l}}$: The ratio between $e_{t,l}$ and $c_{t,l}$, which can be regarded as the traffic situation (congestion level) along the air routes flown by the flights using pattern $l$;
- $p_{t,l}$: The average travel time of flights using pattern $l$ inside an MAS during time slot $t$;
- $w_{t,l}$: The weather condition at the airport serviced by flights using pattern $l$ during time slot $t$, measured by its visibility;
- $d_t$: The day of the week for time slot $t$.

To incorporate the historical information, the multi-variate spatio-temporal features in the past $q$ time windows,

$$[I_{t-q,l},\ \ldots,\ I_{t-1,l},\ I_{t,l}],$$

are all included in the formatted inputs as shown in Fig. 8 and the this time duration $q$ can be appropriately selected after experiments.

Although the multi-variate spatio-temporal features can reflect the unique complexity and vagaries of traffic in an MAS, it is much harder to extract meaningful information from them and make multi-step-ahead predictions than single-variate inputs. A powerful and representative learning model, Multi-variate Stacked Fully connected Bidirectional-LSTM (MSB-LSTM), is developed to facilitate such purpose.

As shown in Fig. 8, a Stacked Bidirectional-LSTM (B-LSTM) is employed to initialize the data processing and facilitate extracting representative information and meaningful statistics from multi-variate spatio-temporal features $[I_{t-q}, \ldots, I_t]$ and embedding them into high dimensional hidden-state vectors $[H_{t-q}, \ldots, H_t]$.

Furthermore, since we aim at predicting medium-term ETA at a flight's departure time, a multi-step-ahead prediction is required for in-MAS prediction. Assume if the flight's out-MAS travel time is predicted as $m$ time periods, then $m$-step-ahead prediction is needed for its in-MAS travel time. To further solve the issue of the accumulation of prediction errors and make multi-step-ahead predictions, a sequence-to-sequence model using encoder–decoder architecture is proposed as shown in Fig. 8. After generating the high dimensional hidden-state vectors $[H_{t-q}, \ldots, H_{t-1}, H_t]$, another LSTM layer is used as the encoder to map them to a single hidden state $e_t$. Then, the decoder uses $e_t$ to generate future information for the incoming $m$ time periods. With the help of an LSTM layer in the decoder, the future hidden sequence is separately processed by using several dense layers, *i.e.*, the $m$ parallel fully connected layers $\mathrm{MLP}_1, \mathrm{MLP}_2, \ldots, \mathrm{MLP}_m$. Finally, the travel times of flights using pattern $l$ during the incoming $m$ intervals, $[p_{t+1,l}, p_{t+2,l}, \ldots, p_{t+m,l}]$, are predicted through the decoder from the hidden-state vector $e_t$. By using this proposed MSB-LSTM, the model can make multi-step-ahead predictions of in-MAS travel time for each trajectory pattern.

### 3.5. Trajectory pattern flow prediction and medium-term ETA

The proposed "Bubble" mechanism enables the decomposition of medium-term ETA prediction into two sub-tasks, out-MAS and in-MAS prediction. There is still a missing component for integrating out-MAS and in-MAS prediction. From above, in-MAS prediction

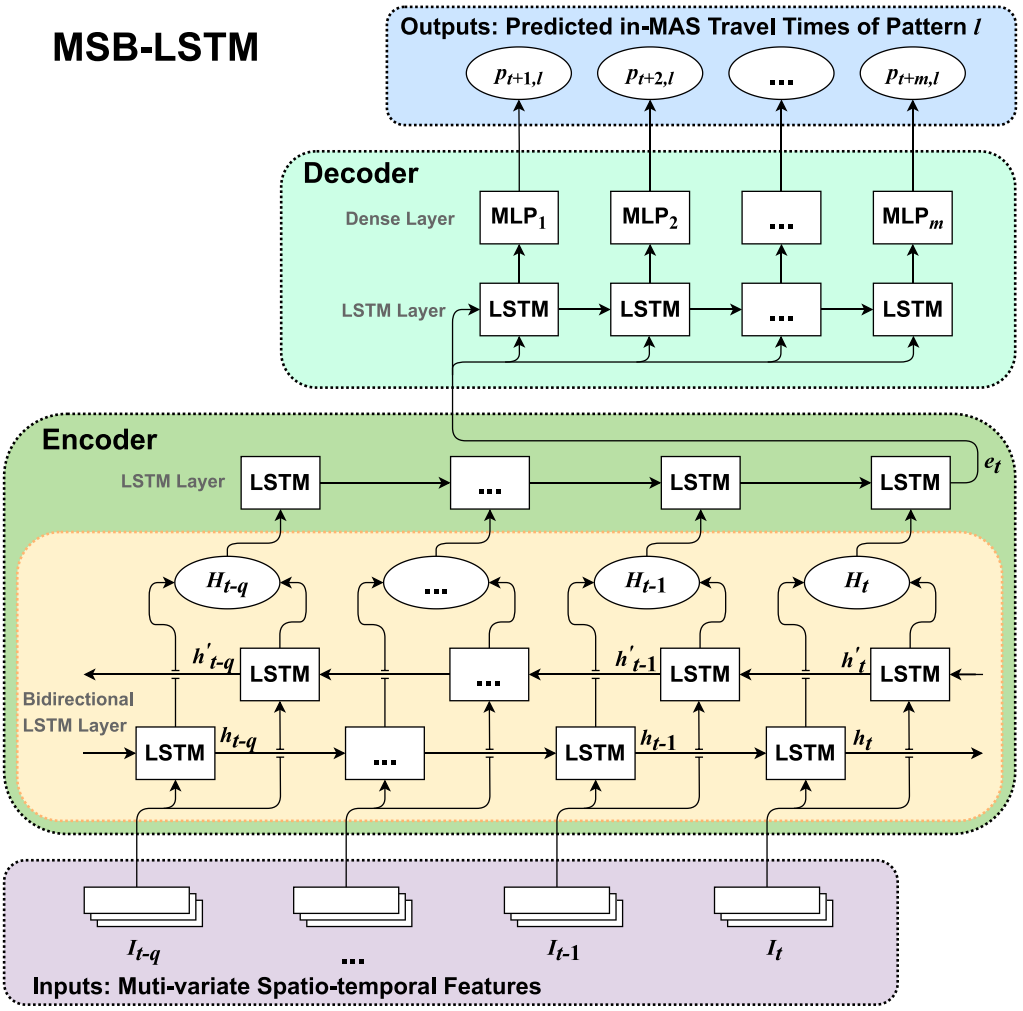



**Fig. 8.** The model of MSB-LSTM for in-MAS prediction.

can make multi-step-ahead predictions of travel times within an MAS if flight trajectory pattern is given. However, the trajectory pattern that a flight may fly is not exactly known when the flight departs. Thus, the component of trajectory pattern prediction is needed for such integration.

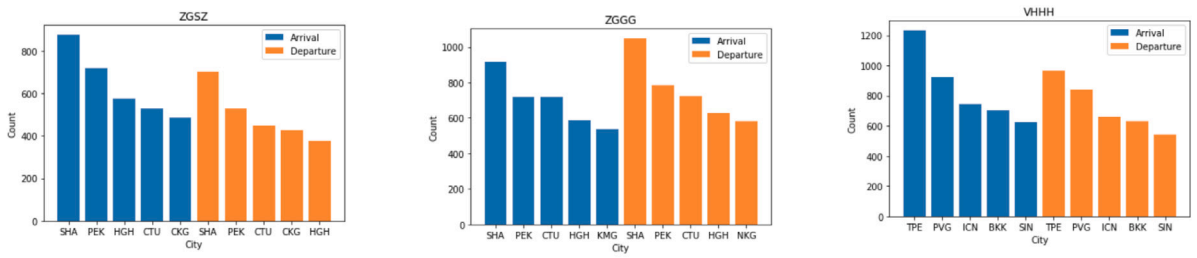
To realize trajectory pattern prediction, the random forest is employed using $[c_{t-q,l}, \ldots, c_{t,l}]$ and $[w_{t,l}, \ldots, w_{t-q,l}]$ for all relevant trajectory patterns $l$ as inputs to obtain the probability distribution of the trajectory pattern that the flight may fly when entering the MAS, which facilitates the calculation of the expected travel time within the MAS. Finally, the medium-term ETA can be obtained by summating predicted out-MAS travel time and expected in-MAS travel time.

## 4. Experimental results

### 4.1. Data set

In this work, since we consider the shared TMAs in GHM-GBA as a whole, we cropped and kept only the spatially nearby airports' interacted region parts. For each airport in GHM-GBA, to guarantee cropped 4D trajectories consisting of the decent stage of the flight as much as possible, we approximately use a circle with 300 km as the radius to cover the airport's TMA. Then, interacted circular regions as the shared TMAs are defined as the rectangle region, as shown in Fig. 1(a). Three points are used to define in-MAS: (1) the right upper corner with the coordination $(25.5°N, 116.6°E)$; (2) the left lower corner with the coordination $(20°N, 111°E)$.

The top eight OD pairs with the heaviest traffic throughput destined to the airports in GHM-GBA in July 2019, as shown in Fig. 9, are collected as the case study in this work, *i.e.*, 785 flights of ZBAA to ZGGG, 1145 flights of ZSSS to ZGGG, 682 flights of ZUUU to ZGGG, 493 flights of ZSHC to ZGSZ, 574 flights ZBAA to ZGSZ, 723 flights ZSSS to ZGSZ, 815 flights of ZSPD to VHHH, and 534 flights of ZBAA to VHHH. To analyze topology and spatial information in the MAS, trajectory data of all flights departing




(a) ZGSZ Traffic Accounts (b) ZGGG Traffic Accounts (c) VHHH Traffic Accounts

**Fig. 9.** Traffic accounts of three major airports in GHM-GBA.

**Table 5**
Performance of out-MAS prediction for 8 OD pairs.

| Departure airport | Arrival airport | MAE (min) | RMSE (min) |
|---|---|---|---|
| ZBAA | VHHH | 2.53 | 3.17 |
| ZBAA | ZGGG | 3.33 | 4.40 |
| ZBAA | ZGSZ | 3.51 | 4.31 |
| ZSPD | VHHH | 2.70 | 3.29 |
| ZSSS | ZGGG | 1.56 | 2.22 |
| ZSSS | ZGSZ | 3.08 | 3.94 |
| ZUUU | ZGGG | 2.31 | 3.17 |
| ZSHC | ZGSZ | 2.70 | 3.56 |

from and arriving at the five airports in GHM-GBA and weather reports of these airports and their surrounding areas in July 2019 are collected and shared via a worldwide network of ADS-B receivers operated by FlightAware. There are 19 651 flights of ZGSZ, 33 801 flights of ZGGG, 30 085 flights of VHHH, 4627 flights of VMMC, and 5789 flights of ZGSD. Each 4D flight trajectory is saved as a CSV file, including general information about the flight, *i.e.*, timestamp, location, course, and speed in kts and mph. In this case, we mainly focus on spatial information of the flight from 4D trajectory data, so timestamp, latitude, longitude, and altitude in feet are information selected. From FlightAware, ADS-B data has various response frequencies, which means different flight segments may have different numbers of ADS-B data points. However, in this study, because we sample the trajectory from a spatial aspect without temporal information, the different response frequency of ADS-B does not impact the sampling result.

### 4.2. Experimental setup

The GHM-GBA is a typical MAS containing five airports, ZGSZ, ZGGG, VHHH, VMMC and ZGSD. A case study of GHM-GBA is conducted here to test the performance of the proposed method. Since ZGGG, ZGSZ and VHHH are the most complex and busy airports due to their geographic locations and ten-million-level throughput, we run experiments on these three airports to demonstrate the performance. All experiments are executed on the Linux cluster (CPU: Intel(R) Xeon(R) E5-2699 v4 @ 2.20 GHz, GPU: NVIDIA GeForce GeForce RTX 2080 Ti). In the following experiments, the performance of the proposed prediction model is evaluated by Mean Absolute Errors (MAE) and Root Mean Square Errors (RMSE),

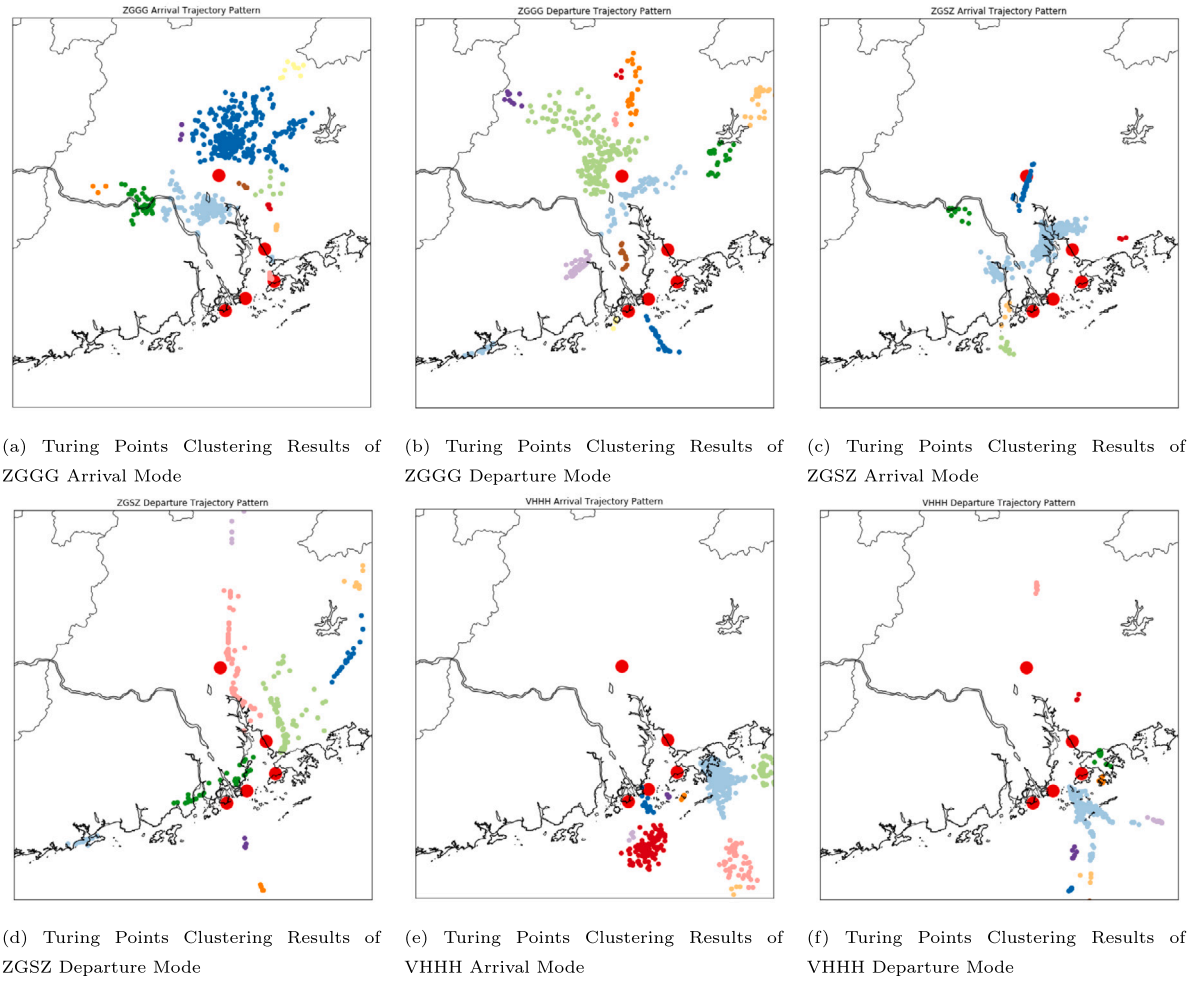
$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|, \quad \mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2},$$

where $n$ is the total number of test cases, $\hat{y}_i$ and $y_i$ are the predicted value and true value of the $i$th case respectively.

The Proposed "Bubble" Mechanism: ARIMA model is used for the out-MAS prediction. For each OD pair in our data set, the time lag order p, the differencing order d, and time lags of moving average q are decided by using "auto_arima" function in Python Libraries. During the training process, we continually change the training set with the fixed length of the time window, *i.e.*, 168 steps in this case, and rolling predict on three steps. For the in-MAS prediction, MSB-LSTM is proposed to make a sequence to sequence predictions. Two layers of Bidirectional LSTM with 32 units are used first. We also stack a regular LSTM with 128 units after Bi-LSTM. Then, stacked fully connected layers are added. Adam optimizer is used with a learning rate of 0.001. Finally, we train the model by minimizing the mean absolute error for 300 epochs with 64 batch sizes.

### 4.3. out-MAS prediction

The data set of the month (July 2019) was used here. Each day(twenty-four hour) is divided into 48 slots identically, each of which is 30 min. $x_t$ represents the average out-MAS travel time for flights which serve the same OD pair and depart in the $t$th time segment.




(a) Turing Points Clustering Results of ZGGG Arrival Mode

(b) Turing Points Clustering Results of ZGGG Departure Mode

(c) Turing Points Clustering Results of ZGSZ Arrival Mode

(d) Turing Points Clustering Results of ZGSZ Departure Mode

(e) Turing Points Clustering Results of VHHH Arrival Mode

(f) Turing Points Clustering Results of VHHH Departure Mode

**Fig. 10.** Turning points clustering results for airports in GHM-GBA.

Augmented Dickey–Fuller (ADF) test was performed to check the stationarity of the time series $\{x_t\}$. Since the original time series $\{x_t\}$ is non-stationary by the ADF test, to apply the ARIMA$(p,d,q)$ model, the suitable parameters of $(p,d,q)$ was carefully selected. For each group of the time series, ADF test, the time plot and the sample ACF of the time series are performed to select the corresponding degree of differencing. Multi-difference may be required unless the differenced data is stationary. After the degree of differencing, $d$, is chosen, the plots of the sample ACF and PACF are used to choose the other two parameters: $p$, the order of the autoregressive model, and $q$, the order of the moving average model. After selecting the parameters of $p,d,q$, the ARIMA$(p,d,q)$ model is trained and used to predict travel time of the flight in the out-MAS stage. The performance of the out-MAS prediction is shown in Table 5.

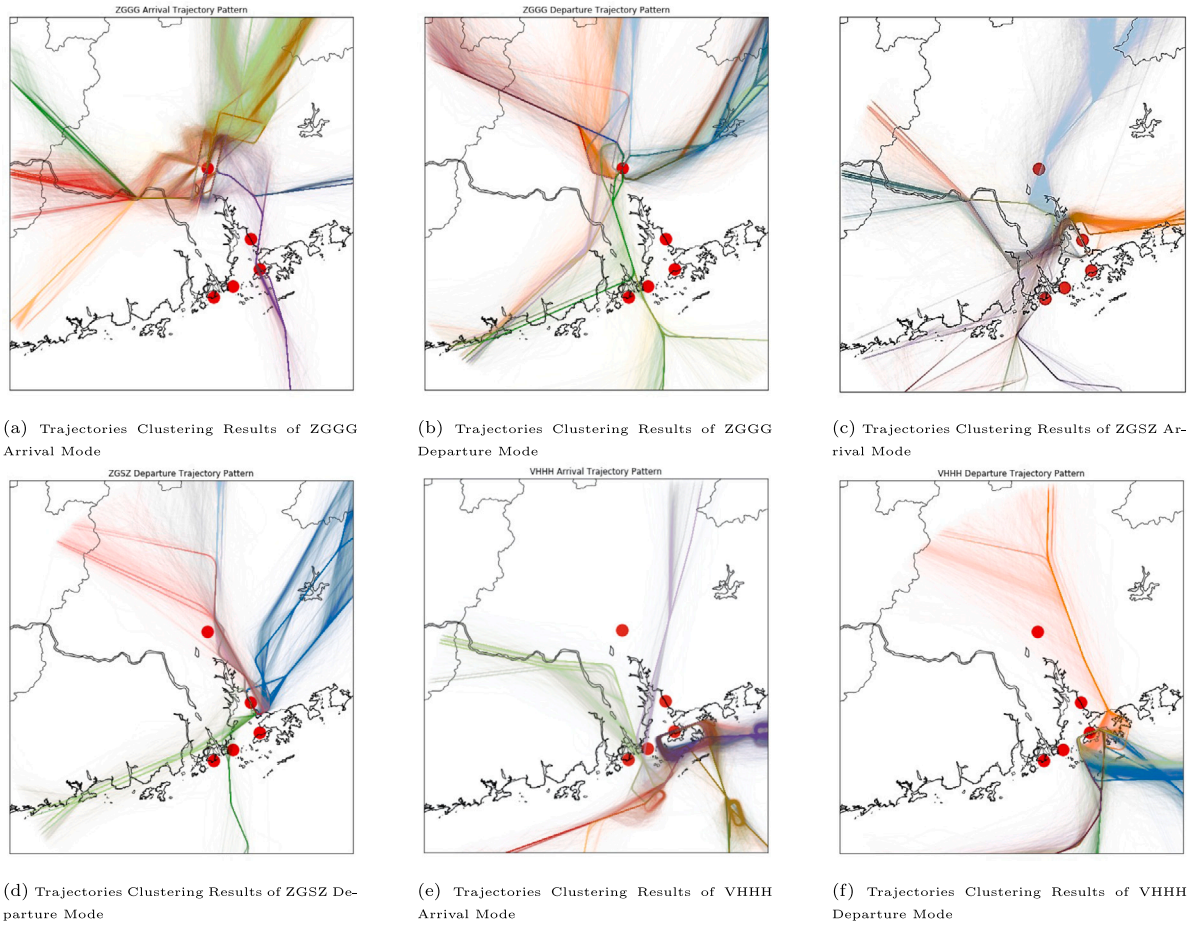
### 4.4. in-MAS trajectory flow pattern clustering

The data-driven hybrid polar sampling is implemented to construct formatted trajectory information vectors $\mathbf{V}_j$ in (4), in which turning points are clustered for each combination of airport and operation mode as shown in Fig. 10. Then, trajectory patterns are identified by applying DBSCAN for the three major airports with ten-million-level throughput in GHM-GBA as shown in Fig. 11.

### 4.5. Comparative experiments: Trajectory sampling and clustering analysis

We provide configurations of the comparative sampling methods with two commonly used baselines in trajectory pattern study, including grid-based sampling method, and uniform concentric circle sampling method (Wang et al., 2017, 2018; Ma et al., 2022). At the end, we discuss potential shortcomings compared with the proposed hybrid polar sampling method with figures in detail.

We use the sampling method to provide information on the spatial shape in flight trajectories to clustering algorithms. Therefore, to ensure the difference among trajectories can be computed, the sampling method should first guarantee the same dimension for




(a) Trajectories Clustering Results of ZGGG Arrival Mode

(b) Trajectories Clustering Results of ZGGG Departure Mode

(c) Trajectories Clustering Results of ZGSZ Arrival Mode

(d) Trajectories Clustering Results of ZGSZ Departure Mode

(e) Trajectories Clustering Results of VHHH Arrival Mode

(f) Trajectories Clustering Results of VHHH Departure Mode

**Fig. 11.** Trajectory clustering results for airports in GHM-GBA.

each trajectory. Then, sampling methods that contain fewer samples with more rich information are better. For the flight entering MAS or TMAs, we sample the flight from three major aspects in the hybrid polar sampling method, i.e., where it is from, where it makes turns, and the regular path information.

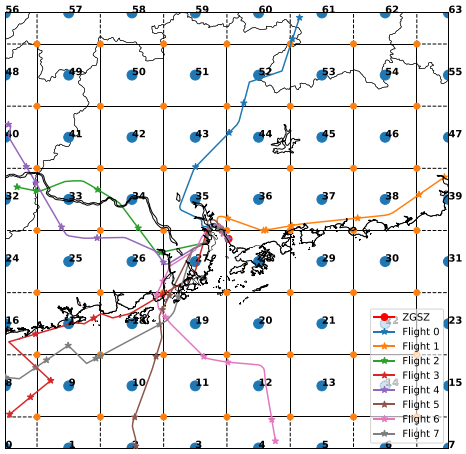
Except entry and landing locations used in previous work, our method is the first to consider flight behaviors in trajectory sampling and clustering, as shown in Table 3. Using the point merge system as the special case, the proposed hybrid polar sampling can be clearly distinguished from other methods used in prior research.

The point merge system in EUROCONTROL is based on a specific P-RNAV route structure, including the merge point and the sequencing legs equidistant from this point. In the proposed hybrid polar sampling method, one of the critical reasons we provide turning-point clustering as the information of where flights possibly make turns is to identify regions of the merge point and the corresponding turning locations for each sequencing leg in the point merge system. If the point merge system is used, even though we cannot precisely sample turning points of sequencing legs to the same merging point, the turning-point clustering limits those points in the fixed region, as shown in Fig. 10(c). As the main drawback in previous research (Ma et al., 2022; Wang et al., 2017, 2018, 2020b; Hong and Lee, 2015), the temporal sampling method (Hong and Lee, 2015) and the spatial sampling method (Ma et al., 2022; Wang et al., 2017, 2018, 2020b) without considering flight behaviors are inappropriate to this case.

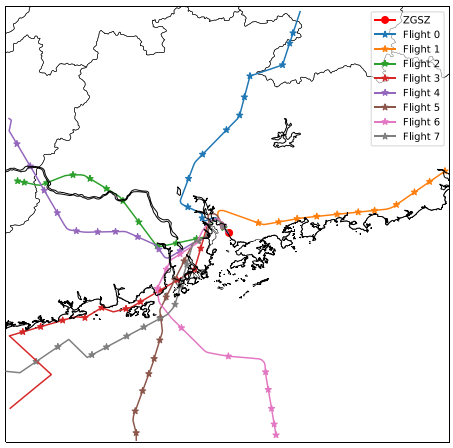
Second, to clearly illustrate the advantages of the proposed hybrid polar sampling, we provide comparative experiments of sampling methods with corresponding clustering results. Flights to ZGSZ in July 2019 are used as the dataset in this study. We randomly select eight flight trajectories as samples to illustrate how different sampling methods work.

We exhibit configurations of different sampling methods as follows:

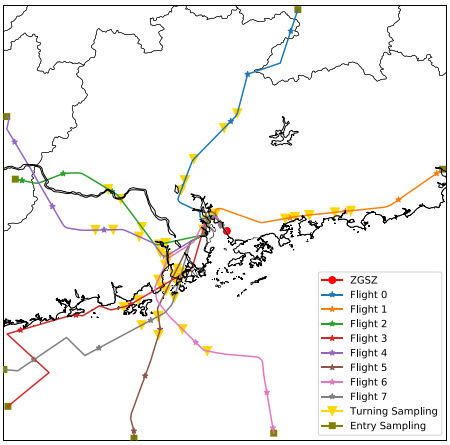
- Grid-based sampling method
  - We discretize the defined TMAs in this study to 64 grids. As shown in Fig. 12(a), each flight is separated into connected segments. We collect points that cross the grid as samples of this 4D trajectory.
- Uniform concentric circle sampling method




(a) The Grid Based Sampling Results of ZGSZ




(b) The uniform concentric circle sampling Results




(c) The Hybrid Polar Sampling Results

**Fig. 12.** Comparative sampling results.

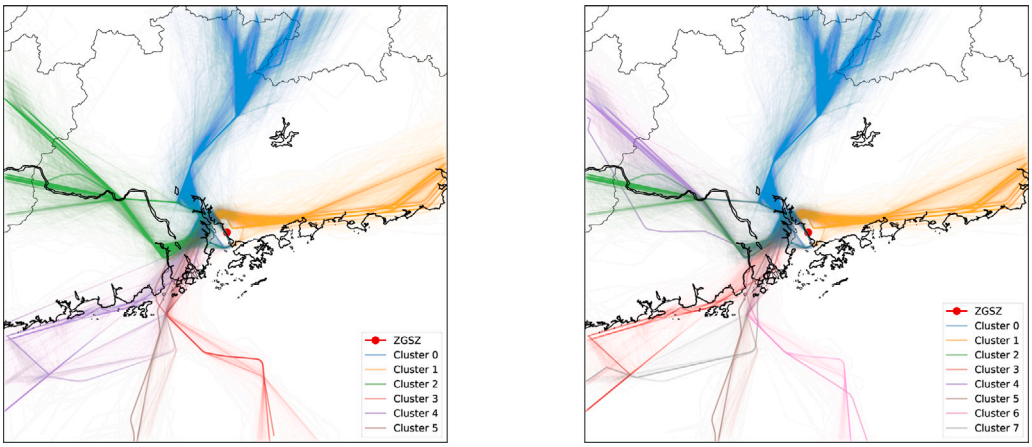
- To compare with the proposed hybrid polar sampling method, we limit the number of samples in each flight trajectory. In this case, As shown in Fig. 12(b), points in the flight are sampled uniformly using 12 concentric circles.

• The proposed hybrid polar sampling method

- The proposed hybrid polar sampling method samples 12 points in total from the flight trajectory, including one entering location, six turning-point samplings, and five concentric circles sampling.

#### *4.5.1. Grid-based sampling method*

As we mentioned, the priority that the sampling method should guarantee is to provide the same dimension to sampled flights; otherwise, the clustering algorithm cannot compute distance among samples. However, as shown in Fig. 12(a), without data interpolations, the grid-based sampling method cannot guarantee dimensions after sampling because flights pass through different amounts of grids. Moreover, the grid-based sampling trades points of the flight same, which means it has a high probability of missing important points, like the turning point.




(a) Trajectory Pattern Clustering of ZGSZ: Uniform Concentric Circle Sampling

(b) Trajectory Pattern Clustering of ZGSZ: The Hybrid Polar Sampling

**Fig. 13.** Comparative clustering analysis.

#### 4.5.2. *Uniform concentric circle sampling method*

Compared with the proposed sampling method, even though the uniform sampling method uses the same number of samples for flights, it treats all points in the flight the same, regardless of whether they are turning points. In this case, this uniform concentric circle sampling method cannot identify where flights make turns, which is however an important information regarding flight behaviors. In contrast, in the proposed hybrid polar sampling method, we first conduct turning-point clustering to identify the regions where flights to the same airport more likely make turns, which faciliates more sampling in turning areas to better capture flight behavior information. Therefore, comparing Figs. 13(a) and 13(b), by using the proposed method that samples points from the three defined aspects, we can more accurately and specifically identify clusters to represent different flight behaviors. For instance, for flights from the west to ZGSZ, the proposed hybrid polar sampling classifies two clusters to reflect the importance of entry point and turning points in characterizing flight behavior.

#### 4.5.3. *The proposed hybrid polar sampling method*

In the proposed hybrid polar sampling method, we sample the flight from three different information, i.e., where it is from, where the flight possibly turns, and the regular path information. Therefore, we can provide a weighted sampling vector to represent the flight.

In Fig. 12(c), each sampled flight contains information on entering location, possible turning locations of flights to the same airport, and regular path information. Therefore, if two flights are from the same entry, and turning sampling points are closed, the computed distance in the clustering algorithm will be small. Since the hybrid polar sampling method provides the weighted features to DBSCAN, more accurate trajectory flow patterns are clustered in Fig. 13(b).

### 4.6. *Trajectory flow pattern prediction*

We apply Random Forest (RF) to obtain the probability of trajectory patterns that a flight may fly inside an MAS accordingly for each OD pair. To fully use the constructed feature of each cluster in the in-MAS prediction model, the RF model can reuse the same information to classify the flight pattern from clustered trajectory flow patterns. In the prediction work of Omar and Hasanen, they claimed significant characteristics of RF as the major powerful algorithms, including abilities on dealing with high dimensional data, difficult correlations, and interactions (Madeeh and Abdullah, 2021). In our problem, the feature matrix we constructed in the in-MAS has various correlations among each feature, for instance, the weather condition and time-dependent numbers of flights in the airspace. Moreover, features that are provided to RF model are collected from in-MAS prediction model that at least contain information of two clusters, which means the feature matrix has high dimensions, *i.e.*, $[c_{t-q,l}, \dots, c_{t,l}]$ and $[w_{t,l}, \dots, w_{t-q,l}]$ for all relevant trajectory patterns $l$ in $q$ time steps as inputs. Furthermore, to deal with the class imbalance issue commonly encountered in the multi-class classification problem, we use the receiver operating characteristic curve (ROC) and area under the curve (AUC) to adjust the classification model. The implicit goal of AUC is to deal with situations with a skewed sample distribution, and to avoid over-fit for one particular class, *i.e.*, the higher the AUC is, the better ability the model has in classifying samples into different classes. For example, for flights from Beijing to Guangzhou, there are two possible trajectory patterns as identified in the clustering analysis. As shown in Fig. 14, the score of AUC for the OD pair from Beijing to Guangzhou is 0.84, which demonstrates that the classification model has a good ability to predict the two patterns using historical traffic information during the intervals.

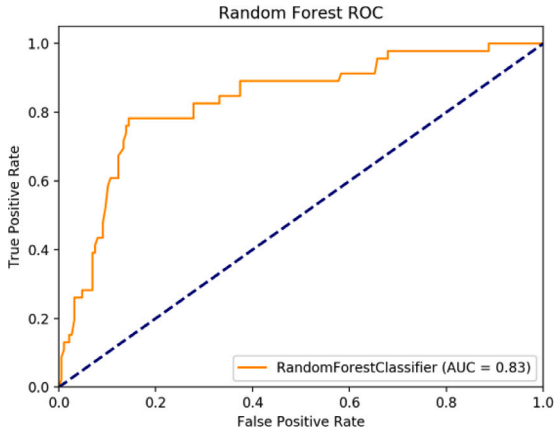



**Fig. 14.** ROC-AUC Plot: Flights from ZBAA to ZGGG.

**Table 6**
Performance of in-MAS prediction for trajectory patterns.

| Trajectory patterns | | Step 1: 0-1 h Prediction | | Step 2: 1-2 h Prediction | |
|---|---|---|---|---|---|
| Entering gate | Arrival airport | MAE (min) | RMSE (min) | MAE (min) | RMSE (min) |
| East9 | ZGSZ | 2.21 | 3.05 | 4.34 | 5.17 |
| East55 | ZGSZ | 1.34 | 3.13 | 2.79 | 4.06 |
| North0 | ZGSZ | 2.42 | 3.67 | 2.72 | 4.08 |
| Northwest3 | ZGSG | 3.80 | 5.58 | 4.56 | 6.38 |
| Northeast2 | ZGGG | 3.39 | 5.14 | 3.97 | 6.23 |
| Northeast7 | ZGGG | 4.19 | 7.06 | 6.78 | 8.34 |
| Northeast8 | ZGGG | 1.40 | 3.67 | 1.98 | 4.18 |
| Northeast9 | ZGGG | 2.15 | 5.21 | 3.09 | 6.11 |
| West5 | ZGGG | 3.02 | 4.41 | 3.90 | 5.84 |
| East1 | VHHH | 3.66 | 5.63 | 5.48 | 6.34 |
| East5 | VHHH | 4.87 | 7.17 | 6.12 | 8.69 |
| North7 | VHHH | 5.98 | 8.12 | 7.02 | 9.04 |

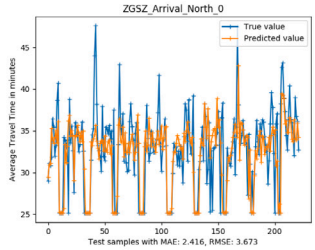
### 4.7. in-MAS prediction

To realize in-MAS prediction, the MSB-LSTM is conducted using spatio-temporal features as inputs, which are constructed based on trajectory pattern clustering. Each day (24 h) was divided into 24 segments identically. The length of time windows $q$ in inputs is set as 3 and time steps $m$ in outputs are set as 2 for the proposed sequence-to-sequence prediction model. Namely, MSB-LSTM receives spatio-temporal information from air traffic networks in GHM-GBA during the past 4 h $[I_{t-3}, I_{t-2}, I_{t-1}, I_t]$ to make the multi-step prediction of average travel time of the trajectory pattern $l$, $[p_{t+1,l}, p_{t+2,l}]$.

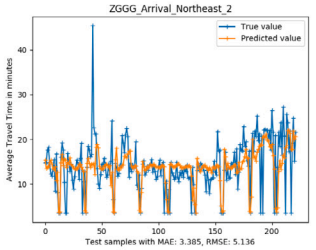
As shown in Table 6, the MSB-LSTM can well predict the in-MAS travel time for all trajectory patterns relevant to medium-term ETA prediction in terms of MAE and RMSE. Each trajectory pattern $l$ is labeled as the pair of entering gate and destination airport, such as "From North0 to ZGSZ", "From Northwest2 to ZGGG", "From East5 to VHHH", *i.e.*, the three trajectory patterns shown in Fig. 15 (The entering gate of each trajectory pattern is named after the geographic location where the flights of the corresponding trajectory pattern enter GHM-GBA, such as "East9", "East55", "North0", "Northwest3", "Northeast2", "Northeast7", "Northeast8", "Northeast9", "West5", "East1", "East5" and "North7"). In this case study, 2 hour-steps prediction enable the flight to select the corresponding predicted travel time based on results from trajectory pattern prediction and arrival time at the margin of GHM-GBA by in-MAS prediction. The detailed prediction results of those trajectory patterns are demonstrated in Fig. 15, in which the tendency and seasonality in the real cases can be closely captured by the predicted values for both the cases with regular daily patterns, such as "From North0 to ZGSZ" and "Northeast2 to ZGGG", shown in Figs. 15(a) and 15(b), and the cases with irregular daily patterns, such as "East5 to VHHH", shown in Fig. 15(c). To verify the performance of the proposed prediction model, we provide prediction plots of all trajectory patterns in Appendix.

### 4.8. Baselines

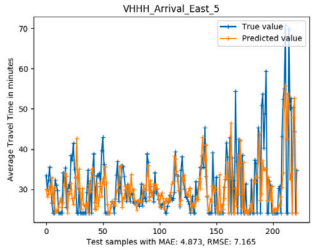
In this study, we compare the proposed prediction framework with four baseline models, *i.e.*, ARIMA, SVR, XGboost Regressor, and Bi-LSTM. For ARIMA and SVR, only temporal information of travel time of OD pairs is converted to inputs. For Bi-LSTM and XGBoost Regressor, traffic information of the OD pair and temporal information of travel time are converted to inputs without information and analysis of the in-MAS in detail. Benchmarks demonstrate the performance of the proposed framework from the constructed spatio-temporal features and the effect of the "Bubble" mechanism.




(a) Travel time prediction: From North0 to ZGSZ




(b) Travel time prediction: From Northeast2 to ZGGG




(c) Travel time prediction: From East5 to VHHH

**Fig. 15.** Examples of Step 1 prediction: Exhibition of in-MAS travel time prediction of trajectory patterns.

- ARIMA
  - As one of the most used time series prediction models, three parameters have to be defined. For each OD pair in our dataset, the time lag order $p$, the differencing order $d$, and time lags of moving average $q$ are decided by using "auto_arima" function in Python Libraries. In addition, we also use Bayesian Information Criterion (BIC) to verify parameters selected by "auto_arima". To guarantee the fairness of comparison and the dynamic training process of ARIMA, we continually change the training set with the fixed length of the time window, *i.e.*, 168 steps in this case, and rolling predict on three steps.
- SVR
  - Comparing with traditional regression models, SVR has tolerated deviation $\epsilon$ in the loss calculation. Two parameters, *i.e.*, the penalty tern $C$ 0.1 and the deviation $\epsilon$ 0.1, with the linear kernel are selected. The tested regions of the penalty tern $C$ and the deviation $\epsilon$ is from 0.1 to 0.5 with step 0.05.
- XGBoost Regressor
  - As the supervised learning model, we provide the same inputs with the benchmark Bi-LSTM. *n_estimator* is selected as 1000 from the region of 500 to 1500 with step 100, and others use default tuning parameters.
- Bi-LSTM
  - In this model, we use the grid search hyper-parameters tuning algorithm on the training. Units' regions in Bi-LSTM and LSTM are provided as $[16, 32, 64, 128]$. In addition, we set up the learning rate in this test as $[1e-4, 5e-4, 1e-3, 5e-3]$. Specifically, Bi-LSTM contains two layers of LSTM with opposite recurrent directions. Each layer contains 64 units. We also stack a regular LSTM with 128 units after Bi-LSTM. Then, stacked fully connected layers are added. Adam optimizer is used with a learning rate of 0.001. We train the model by minimizing the mean absolute error for 300 epochs with 64 batch sizes.

### *4.9. Experimental comparison and the case study: Prediction of medium-term ETA for eight OD pairs*

The case study of medium-term ETA prediction is conducted for the 8 busiest OD pairs destined for the major airports in GHM-GBA to demonstrate the performance of the "Bubble" mechanism, in which out-MAS prediction and in-MAS prediction are separately carried out and integrated afterward. For the example of ZBAA to ZGSZ, out-MAS prediction takes flight departure times at ZBAA and historical travel time in the past 2 h as inputs of ARIMA to predict the corresponding travel time of the segment, which is the air route to the margin of GHM-GBA. In this case, the departure time of the flight at ZBAA stands for the current time slot $t$, and in-MAS prediction constructs spatio-temporal features, $[I_{t-3}, I_{t-2}, I_{t-1}, I_t]$ as inputs to make multi-step-ahead prediction through MSB-LSTM for possible trajectory patterns selected by historical flights serving ZBAA to ZGSZ. Trajectory pattern prediction uses the result from out-MAS prediction to indicate the probability distribution for clustered trajectory patterns of the OD pair at the time when the flight enters the shared TMAs. After that, the weighted travel time of the flight is carried out by in-MAS predictions and the clustered patterns' probability distribution in the trajectory pattern prediction. Finally, with results from three main components in the "Bubble" mechanism, the medium-term ETA of the flight can be represented by accumulating prediction results together.

We compare the performance of the proposed "Bubble" mechanism versus commonly used baselines, including Bi-LSTM, XGboost Regressor, ARIMA, and SVR, shown in Table 7. Temporal information about OD pairs' travel time provides features to ARIMA and SVR. For Bi-LSTM and XGboost Regressor, features of the OD pair include the number of flights as traffic information and the corresponding temporal information of the travel time.

As shown in Table 7, for the first type of comparison, the MAE of the proposed "Bubble" mechanism is 20% to 50% better than the ARIMA and SVR. Using the history of travel time, ARIMA and SVR are the worst baselines compared with others. Only

**Table 7**
Performance experimental comparison of different methods : Prediction of medium-term ETA.

| OD Pairs | | The "Bubble" mechanism | | Bi-LSTM | | XGboost regressor | | ARIMA | | SVR | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Origin | Destination | RMSE (min) | MAE (min) | RMSE (min) | MAE (min) | RMSE (min) | MAE (min) | RMSE (min) | MAE (min) | RMSE (min) | MAE (min) |
| ZBAA | ZGSZ | **6.17** | **4.70** | 9.09 | 6.29 | 8.12 | 5.60 | 10.21 | 8.37 | 9.77 | 6.72 |
| ZSHC | ZGSZ | **4.19** | **3.32** | 5.82 | 4.41 | 6.72 | 5.31 | 7.03 | 5.79 | 8.34 | 6.12 |
| ZSSS | ZGSZ | **5.80** | **4.55** | 6.72 | 5.54 | 6.30 | 5.04 | 8.37 | 6.77 | 7.12 | 5.83 |
| ZBAA | ZGGG | **8.07** | **6.60** | 10.45 | 7.51 | 12.55 | 10.10 | 14.32 | 12.44 | 13.74 | 10.82 |
| ZSSS | ZGGG | **7.99** | **6.51** | 9.57 | 7.21 | 10.34 | 9.12 | 11.33 | 8.45 | 12.11 | 8.13 |
| ZUUU | ZGGG | **6.09** | **4.65** | 7.98 | 6.09 | 6.33 | 5.23 | 11.02 | 7.67 | 9.37 | 6.17 |
| ZBAA | VHHH | **10.12** | **8.18** | 15.03 | 11.23 | 14.31 | 12.36 | 17.05 | 15.77 | 15.83 | 13.22 |
| ZSPD | VHHH | **7.55** | **5.76** | 13.05 | 10.24 | 11.75 | 8.92 | 15.07 | 12.21 | 14.12 | 11.34 |

**Table 8**
Comparative experiments: Effects of the MAS in The "Bubble" mechanism.

| OD Pairs | | The "Bubble" mechanism | | The "Bubble" mechanism without MAS analysis | |
|---|---|---|---|---|---|
| Origin | Destination | RMSE (min) | MAE (min) | RMSE (min) | MAE (min) |
| ZBAA | ZGGG | **8.07** | **6.60** | 10.12 | 7.44 |
| ZBAA | ZGSZ | **6.17** | **4.70** | 7.67 | 5.12 |
| ZBAA | VHHH | **10.12** | **8.18** | 13.12 | 9.83 |

the historical end-to-end travel time (temporal information) utilized in the two baseline methods is insufficient to produce a good ETA prediction. For the second type of comparison, the proposed framework is 10% to 40% better than the Bi-LSTM and XGboost Regressor. Compared with the proposed "Bubble" mechanism, there is no spatial information behind air traffic networks from the in-MAS analysis. Therefore, the effect of high travel time variance at MAS and dynamic trajectory patterns in the shared TMAs cannot be reflected by features in Bi-LSTM and XGboost Regressor. Namely, the proposed "Bubble" mechanism can help focus on the complexities of the in-MAS stage and capture MAS's spatial and temporal characteristics to make a better prediction.

From the comparative experiments in Table 7, the proposed "Bubble" mechanism demonstrates the importance of spatial information behind the shared TMAs of GHM-GBA. To further analyze and verify how the MAS affects prediction problems using the proposed framework, we simulate three tested OD pairs in this study as the single airport case by ignoring effects from MAS. In Table 8, we provide a comparative experiment based on the proposed "Bubble" mechanism without MAS analysis.

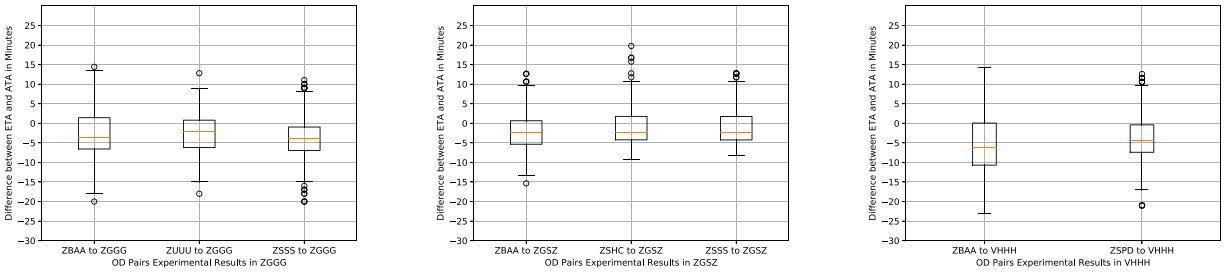
In the proposed $I_{t,l}$, $c_{t,l}$ and $r_{t,l}$ are features related to traffic conditions of the entire MAS. Therefore, we ignore those two features and use ZBAA to ZGSZ, ZGGG, and VHHH as case studies under the simulated single airport case. To keep fairness, we use the same configuration of prediction neural networks in the in-MAS prediction. As shown in Table 8, since we try to ignore effects from MAS which is simulated as the single airport system, prediction results of the "Bubble" Mechanism without MAS Analysis is worse than the proposed method, but they are still better than other baselines in Table 7 because of spatial information in the proposed "Bubble". For this comparative experiment, compared with others only considering OD pairs, the "Bubble" mechanism without MAS analysis demonstrates the importance of the "Bubble" and the effects of spatial dependencies of the shared TMAs. Moreover, Compared with the proposed "Bubble" mechanism, the predicted results of three tested OD pairs without features of neighbor airports and the MAS reduce performance of MAE from around 8% to 16%.

The prediction error distributions of the 8 busiest OD pairs are depicted in Fig. 16 using box-plot, and histograms in Appendix A.3. The majority of cases have prediction errors of almost less than 10 min, and 15 min is usually used as the threshold to determine whether a flight is delayed by CAAC in China and FAA in US. To look more closely at the prediction errors, we provide VHHH as an example to exhibit the detailed prediction values in Fig. 17. It further demonstrates the proposed "Bubble" mechanism can capture the tendency and seasonality of travel times in each OD pair when predicting medium-term ETA. To verify prediction results, prediction plots of other 6 OD pairs are exhibited in Appendix.

## 5. Conclusion

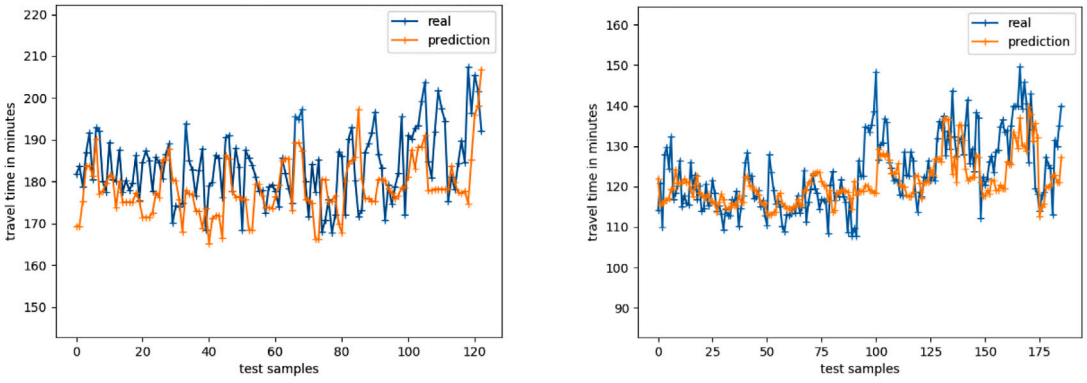
Flight ETA prediction is a challenging task for a multi-airport system like GHM-GBA containing five major airports that are closely located and tightly coupled with each other through shared and limited TMA. It becomes even harder when dealing with the medium horizon. In this paper, a "Bubble" mechanism is developed to efficiently and effectively make medium-term ETA prediction in an MAS.

Through the "Bubble" mechanism, the medium-term ETA prediction can be decomposed into two sub-tasks: out-MAS prediction and in-MAS prediction. The out-MAS prediction is realized by the ARIMA model because of relatively regular en-route traffic conditions. The in-MAS prediction can focus on dealing with the complex network structure and vagaries of traffic conditions within an MAS. A novel spatio-temporal feature vector is constructed based on trajectory pattern classification with the help of a data-driven hybrid polar sampling, which is developed to efficiently and effectively capture the spatial and temporal characteristics of flight trajectories. A sequence-to-sequence prediction model, MSB-LSTM, is further developed to make multi-step-ahead predictions of in-MAS travel times for each trajectory pattern. The out-MAS prediction and in-MAS prediction are finally integrated via trajectory




(a) medium-term ETA of ZGGG (b) medium-term ETA of ZGSZ (c) medium-term ETA of VHHH

**Fig. 16.** Box-plots: Medium-term ETA of airports in GHM-GBA.




(a) Medium-term ETA: ZBAA to VHHH (b) Medium-term: ZSPD to VHHH

**Fig. 17.** VHHH as examples: Exhibition of medium-term ETA prediction.

pattern prediction to predict medium-term ETA. A case study of GHM-GBA, a typical MAS in China, is conducted to demonstrate the usage and promising performance of our proposed method in comparison to four popular baseline methods, which can facilitate improving traffic flow management.

We will study the flight delay propagation scheme for an MAS and accordingly construct a discrete event simulation platform. A hybrid model will be developed based on the simulation platform and the proposed learning model to further improve prediction accuracy. Furthermore, since a promising prediction model enables a good model predictive control, we consider developing a trajectory planning method with the help of the hybrid model for enhancing the air traffic management for an MAS.

## CRediT authorship contribution statement

**Lechen Wang:** Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft. **Jianfeng Mao:** Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing. **Lishuai Li:** Writing – review & editing. **Xuechun Li:** Data curation, Visualization, Investigation, Writing – original draft. **Yilei Tu:** Visualization, Investigation.

## Data availability

The authors do not have permission to share data.

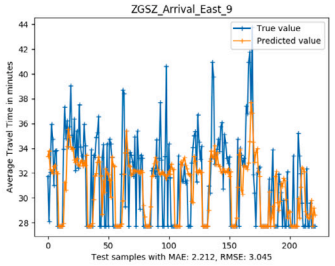
## Acknowledgments

This work was supported in part by National Natural Science Foundation of China under grant U1733102, in part by the Guangdong Provincial Key Laboratory of Big Data Computing, The Chinese University of Hong Kong, Shenzhen under grant B10120210117, in part by Shenzhen Science and Technology Innovation Committee under grant ZDSYS20170725140921348, in part by CUHK-Shenzhen under grant PF.01.000404, and in part by the Development and Reform Commission of Shenzhen Municipality.

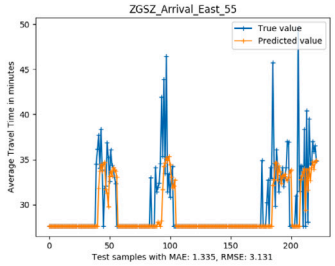
## Appendix

### A.1. Plots of in-MAS travel time prediction of trajectory patterns

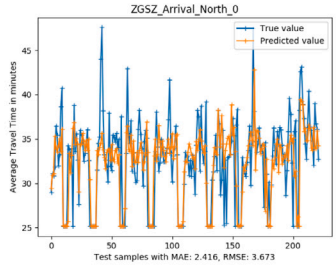
See Fig. 18.



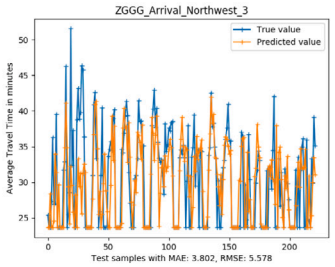

(a) Travel time prediction: From East9 to ZGSZ



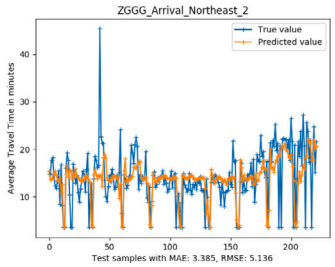

(b) Travel time prediction: From East55 to ZGSZ



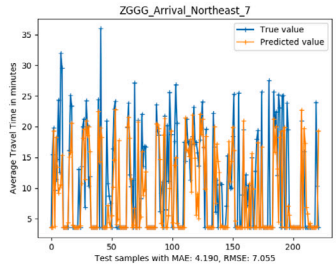

(c) Travel time prediction: From North0 to ZGSZ



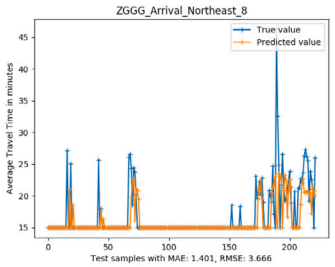

(d) Travel time prediction: From Northwest3 to ZGGG



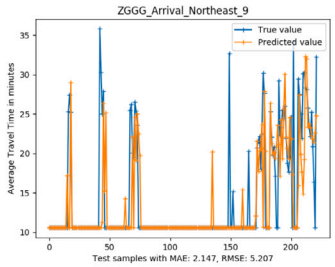

(e) Travel time prediction: From Northeast2 to ZGGG



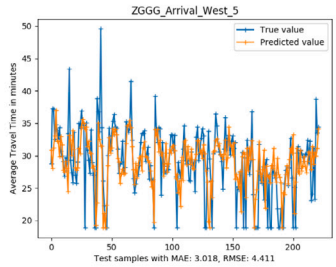

(f) Travel time prediction: From Northeast7 to ZGGG



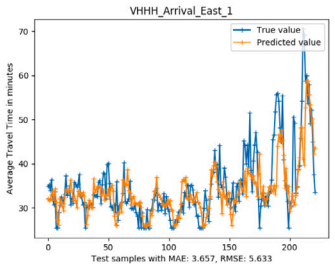

(g) Travel time prediction: From Northeast8 to ZGGG



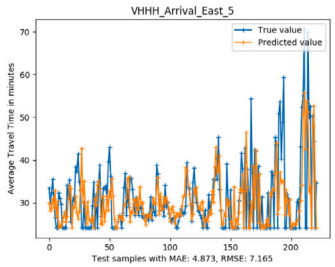

(h) Travel time prediction: From Northeast9 to ZGGG



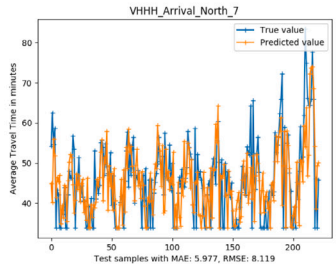

(i) Travel time prediction: From West5 to ZGGG




(j) Travel time prediction: From East1 to VHHH

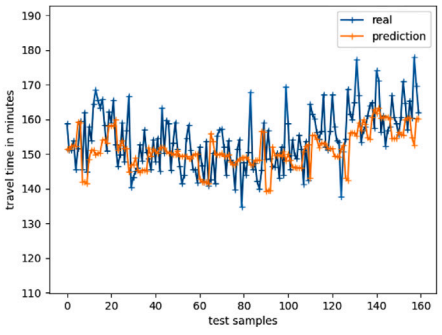



(k) Travel time prediction: From East5 to VHHH

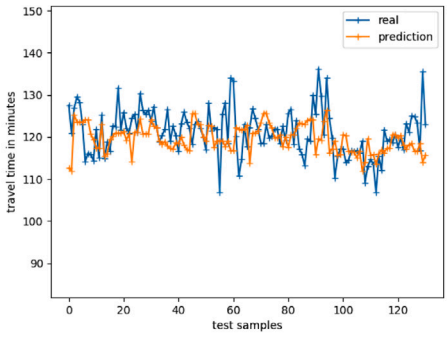



(l) Travel time prediction: From North7 to VHHH

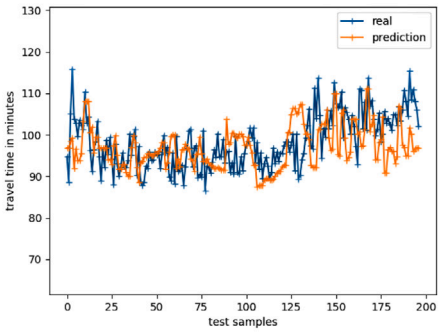
**Fig. 18.** Prediction in-MAS travel times of trajectory patterns.



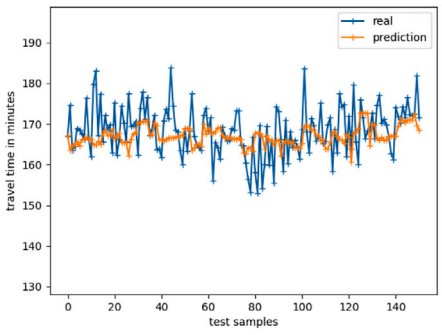

(a) ZBAA to ZGGG



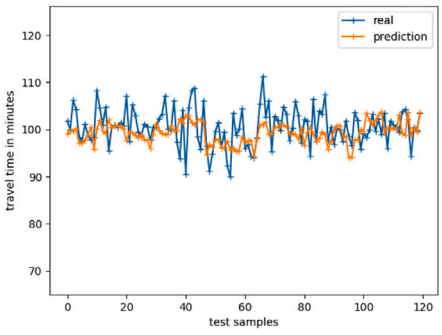

(b) ZUUU to ZGGG



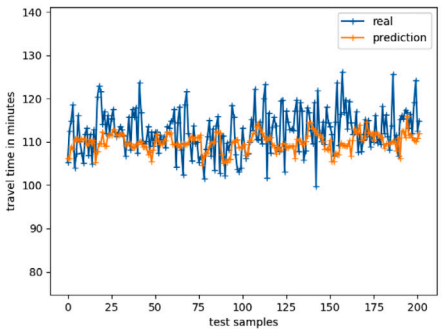

(c) ZSSS to ZGGG



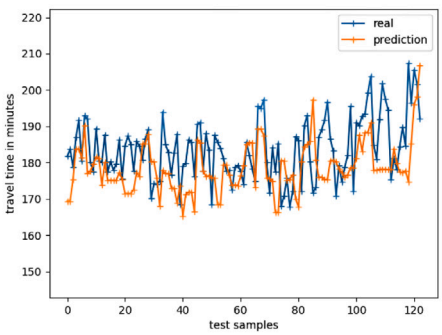

(d) ZBAA to ZGSZ



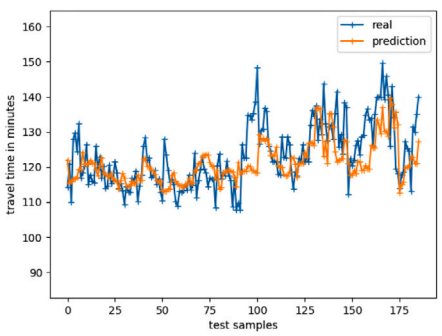

(e) ZSHC to ZGSZ




(f) ZSSS to ZGSZ

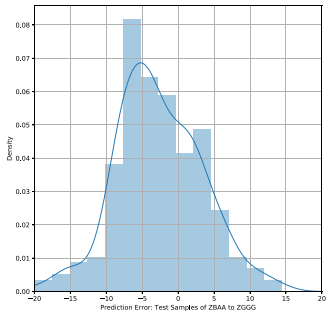



(g) ZBAA to VHHH

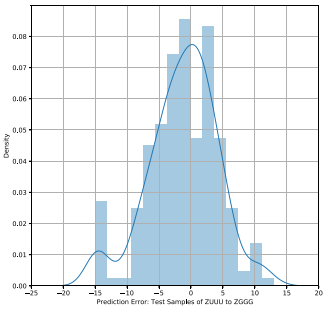



(h) ZSPD to VHHH

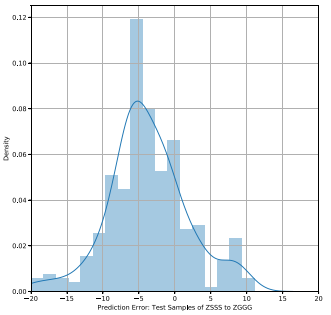
**Fig. 19.** medium-term ETA prediction for 8 OD pairs.



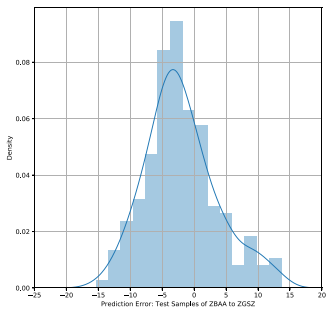

(a) ZBAA to ZGGG



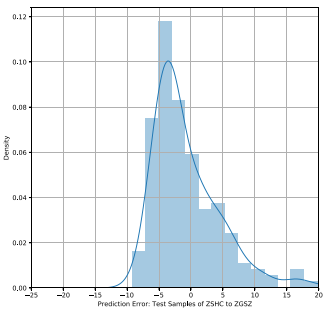

(b) ZUUU to ZGGG



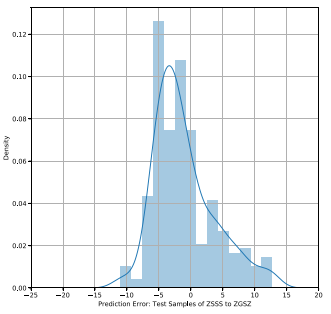

(c) ZSSS to ZGGG



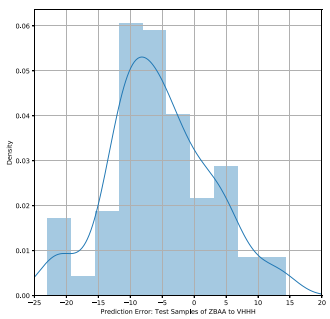

(d) ZBAA to ZGSZ



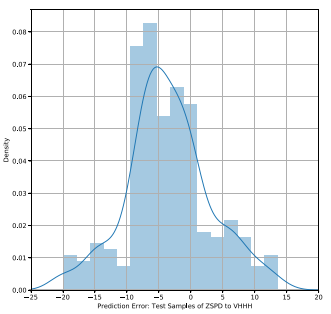

(e) ZSHC to ZGSZ




(f) ZSSS to ZGSZ




(g) ZBAA to VHHH




(h) ZSPD to VHHH

**Fig. 20.** Histograms: medium-term ETA prediction for 8 OD pairs.

### *A.2. Plots of prediction of medium-term ETA*

See Fig. 19.

### *A.3. Histograms of predictions of medium-term ETA*

See Fig. 20.

## References

Atkins, S., Capozzi, B., Hinkey, J., Idris, H., Kaiser, K., 2011. Investigating the nature of and methods for managing metroplex operations. NASA Metroplex NRA Project Report, NASA/CR–2011-216413.

Ayhan, S., Costas, P., Samet, H., 2018. Predicting estimated time of arrival for commercial flights. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 33–42.

Ayhan, S., Samet, H., 2015. Diclerge: Divide-cluster-merge framework for clustering aircraft trajectories. In: Proceedings of the 8th ACM SIGSPATIAL International Workshop on Computational Transportation Science. pp. 7–14.
Bai, X., Weitz, L.A., Priess, S., 2016. Evaluating the impact of estimated time of arrival accuracy on interval management performance. In: AIAA Guidance, Navigation, and Control Conference.
Billings, D., Yang, J.-S., 2006. Application of the ARIMA models to urban roadway travel time prediction - A case study. In: 2006 IEEE International Conference on Systems, Man and Cybernetics. vol. 3, pp. 2529–2534.
Bonnefoy, P.A., 2008. Scalability of the air transportation system and development of multi-airport systems: A worldwide perspective (Ph.D. thesis). Massachusetts Institute of Technology.
Chen, M.-Y., Chiang, H.-S., Yang, K.-J., 2022. Constructing cooperative intelligent transport systems for travel time prediction with deep learning approaches. IEEE Trans. Intell. Transp. Syst. 23 (9), 16590–16599.
De Neufville, R., 1995. Management of multi-airport systems: A development strategy. J. Air Transp. Manage. 2 (2), 99–110.
Dion, F., Rakha, H., 2006. Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. Transp. Res. B 40 (9), 745–766.
Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD96). pp. 226–231.
Gariel, M., Srivastava, A.N., Feron, E., 2011. Trajectory clustering and an application to airspace monitoring. IEEE Trans. Intell. Transp. Syst. 12 (4), 1511–1524.
Glina, Y., Jordan, R., Ishutkina, M., 2012. A tree-based ensemble method for the prediction and uncertainty quantification of aircraft landing times. In: American Meteorological Society–10th Conference on Aritificial Intelligence Applications To Environmental Science, New Orleans, LA.
He, P., Jiang, G., Lam, S.-K., Sun, Y., 2020. Learning heterogeneous traffic patterns for travel time prediction of bus journeys. Inform. Sci. 512, 1394–1406.
He, P., Jiang, G., Lam, S.-K., Tang, D., 2019. Travel-time prediction of bus journey with multiple bus trips. IEEE Trans. Intell. Transp. Syst. 20 (11), 4192–4205.
Hong, S., Lee, K., 2015. Trajectory prediction for vectored area navigation arrivals. J. Aerosp. Inf. Syst. 12 (7), 490–502.
Hou, Z., Li, X., 2016. Repeatability and similarity of freeway traffic flow and long-term prediction under big data. IEEE Trans. Intell. Transp. Syst. 17 (6), 1786–1796.
Huang, H., Roy, K., Tomlin, C., 2007. Probabilistic estimation of state-dependent hybrid mode transitions for aircraft arrival time prediction. In: AIAA Guidance, Navigation and Control Conference and Exhibit.
Innamaa, S., 2005. Short-term prediction of travel time using neural networks on an interurban highway. Transportation 32 (6), 649–669.
Jang, J., 2016. Outlier filtering algorithm for travel time estimation using dedicated short-range communications probes on rural highways. IET Intell. Transp. Syst. 10 (6), 453–460.
Jenelius, E., Koutsopoulos, H.N., 2013. Travel time estimation for urban road networks using low frequency probe vehicle data. Transp. Res. B 53, 64–81.
Jie, Y., Hui, C., Xingyu, L., Xuhui, W., 2019. Research on estimated time of arrival prediction based upon ADS-B and spatiotemporal analysis. In: 2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology. ICCASIT, pp. 630–634.
Krozel, J., Lee, C., Mitchell, J., 1999. Estimating time of arrival in heavy weather conditions. In: Guidance, Navigation, and Control Conference and Exhibit.
Levy, B., Rappaport, D., 2007. Arrival time estimation (ETA) from on-final to gate. In: 7th AIAA ATIO Conf, 2nd CEIAT Int'L Conf on Innov and Integr in Aero Sciences, 17th LTA Systems Tech Conf; Followed By 2nd TEOS Forum.
Li, C.-S., Chen, M.-C., 2014. A data mining based approach for travel time prediction in freeway with non-recurrent congestion. Neurocomputing 133, 74–83.
Lin, D.-J., Chen, M.-Y., Chiang, H.-S., Sharma, P.K., 2022. Intelligent traffic accident prediction model for internet of vehicles with deep learning approach. IEEE Trans. Intell. Transp. Syst. 23 (3), 2340–2349.
Ma, Y., Du, W., Chen, J., Zhang, Y., Lv, Y., Cao, X., 2022. A spatiotemporal neural network model for estimated-time-of-arrival prediction of flights in a Terminal Maneuvering Area. IEEE Intell. Transp. Syst. Mag. 2–15.
Madeeh, O.D., Abdullah, H.S., 2021. An efficient prediction model based on machine learning techniques for prediction of the stock market. In: Journal of Physics: Conference Series. vol. 1804, (1), IOP Publishing, 012008.
Mendes-Moreira, J., Jorge, A.M., Freire de Sousa, J., Soares, C., 2015. Improving the accuracy of long-term travel time prediction using heterogeneous ensembles. Neurocomputing 150, 428–439.
Mueller, T., Sorensen, J., Couluris, G., 2002. Strategic aircraft trajectory prediction uncertainty and statistical sector traffic load modeling. In: AIAA Guidance, Navigation, and Control Conference and Exhibit.
Murça, M.C.R., Hansman, R.J., 2019. Identification, characterization, and prediction of traffic flow patterns in multi-airport systems. IEEE Trans. Intell. Transp. Syst. 20 (5), 1683–1696.
Murça, M.C.R., Hansman, R.J., Li, L., Ren, P., 2018. Flight trajectory data analytics for characterization of air traffic flows: A comparative analysis of terminal area operations between New York, Hong Kong and Sao Paulo. Transp. Res. C 97, 324–347.
Petersen, N.C., Rodrigues, F., Pereira, F.C., 2019. Multi-output bus travel time prediction with convolutional LSTM neural network. Expert Syst. Appl. 120, 426–435.
Ramanujam, V., Balakrishnan, H., 2009. Estimation of arrival-departure capacity tradeoffs in multi-airport systems. In: Proceedings of the 48h IEEE Conference on Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conference. pp. 2534–2540.
Ren, P., Li, L., 2018. Characterizing air traffic networks via large-scale aircraft tracking data: A comparison between China and the US networks. J. Air Transp. Manage. 67, 181–196.
Roy, K., Levy, B., Tomlin, C., 2006. Target tracking and estimated time of arrival (ETA) prediction for arrival aircraft. In: AIAA Guidance, Navigation, and Control Conference and Exhibit.
Salamanis, A., Kehagias, D.D., Filelis-Papadopoulos, C.K., Tzovaras, D., Gravvanis, G.A., 2016. Managing spatial graph dependencies in large volumes of traffic data for travel-time prediction. IEEE Trans. Intell. Transp. Syst. 17 (6), 1678–1687.
Sidiropoulos, S., Majumdar, A., Han, K., 2018. A framework for the optimization of terminal airspace operations in multi-airport systems. Transp. Res. B 110, 160–187.
Stefan, A., Athitsos, V., Das, G., 2012. The move-split-merge metric for time series. IEEE Trans. Knowl. Data Eng. 25 (6), 1425–1438.
Strottmann Kern, C., de Medeiros, I.P., Yoneyama, T., 2015. Data-driven aircraft estimated time of arrival prediction. In: 2015 Annual IEEE Systems Conference (SysCon) Proceedings. pp. 727–733.
Takács, G., 2014. Predicting flight arrival times with a multistage model. In: 2014 IEEE International Conference on Big Data (Big Data). pp. 78–84.
Trivedi, S., Pardos, Z.A., Heffernan, N.T., 2015. The utility of clustering in prediction tasks. arXiv preprint arXiv:1509.06163.
Wang, L., Li, X., Mao, J., 2020a. Integrating ARIMA and bidirectional LSTM to predict ETA in multi-airport systems. In: 2020 Integrated Communications Navigation and Surveillance Conference. ICNS, pp. 3F2–1–3F2–15.
Wang, Z., Liang, M., Delahaye, D., 2017. Short-term 4d trajectory prediction using machine learning methods. In: Proc. SID. pp. 1–10.
Wang, Z., Liang, M., Delahaye, D., 2018. A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area. Transp. Res. C 95, 280–294.
Wang, Z., Liang, M., Delahaye, D., 2020b. Automated data-driven prediction on aircraft Estimated Time of Arrival. J. Air Transp. Manage. 88, 101840.
Wang, P., Wanke, C., Wieland, F., 2004. Modeling time and space metering of flights in the National Airspace System. In: Proceedings of the 2004 Winter Simulation Conference, 2004.. vol. 2, pp. 1299–1304.

Wang, Y., Zhang, Y., 2021. Prediction of runway configurations and airport acceptance rates for multi-airport system using gridded weather forecast. Transp. Res. C 125, 103049.
Wei, J., Lee, J., Hwang, I., 2015. Estimated time of arrival prediction based on state-dependent transition hybrid estimation algorithm. In: AIAA Guidance, Navigation, and Control Conference.
Wu, C.-H., Ho, J.-M., Lee, D., 2004. Travel-time prediction with support vector regression. IEEE Trans. Intell. Transp. Syst. 5 (4), 276–281.
Zahid Reza, R., Pulugurtha, S.S., 2019. Forecasting short-term relative changes in travel time on a freeway. Case Stud. Transp. Policy 7 (2), 205–217.
Zhang, A., Liu, Q., Zhang, T., 2022. Spatial–temporal attention fusion for traffic speed prediction. Soft Comput. 26 (2), 695–707.
Zheng, L., Yang, J., Chen, L., Sun, D., Liu, W., 2020. Dynamic spatial-temporal feature optimization with ERI big data for short-term traffic flow prediction. Neurocomputing 412, 339–350.
Zhu, G., Matthews, C., Wei, P., Lorch, M., Chakravarty, S., 2018. En route flight time prediction under convective weather events. In: 2018 Aviation Technology, Integration, and Operations Conference.