




Review of ADS-B Data Usage with the Focus on Data Cleaning

Ruolan Ren ¹, Jingcheng Zhong,¹ Dizhi Guo,¹ Ruixin Wang ,^{*}¹ and Christophe Hurter ²

¹Civil Aviation University of China, Tianjin, China

²Université de Toulouse, ENAC, Toulouse, France

³IPAL Singapore, Singapore

*Corresponding author: ruixin.wang@recherche.enac.fr

(Received: 1 November 2025; Revised: 15 April 2026; Accepted: 18 April 2026; Published: 27 April 2026)

(Editor: Martin Strohmeier; Reviewers: Ramon Dalmau and Enrico Spinielli)

Abstract

Automatic Dependent Surveillance–Broadcast (ADS-B) data have become a vital resource for research on trajectory prediction, conflict detection, and air traffic management. However, due to limitations in data acquisition and transmission, ADS-B datasets often contain missing points, irregular sampling, and anomalies. To ensure usability, researchers typically apply data cleaning and preprocessing, which improve data quality but may alter original characteristics and cause deviations between algorithm outputs and real operational patterns. Existing studies largely focus on individual cleaning methods, lacking systematic and quantitative assessments of their impact on downstream applications. To address this gap, this study systematically investigates the relationship between data cleaning and algorithmic performance in ADS-B analytics. It reviews major ADS-B applications and prevalent cleaning techniques, summarizes typical preprocessing pipelines, and provides guidance for building more robust evaluation frameworks. An AutoEncoder (AE)-based experiment is conducted using three architecturally distinct autoencoders (Fully Connected AE (FC-AE), Long Short-Term Memory AE (LSTM-AE), Gated Recurrent Unit AE (GRU-AE)) across four geographically diverse airport datasets, with trajectories contaminated by Gaussian, drift, spike, and missing data noise to assess the influence of cleaning strategies on trajectory reconstruction. Results indicate that spike noise has the least impact on reconstruction performance across all model and dataset combinations, while the relative sensitivity to Gaussian, drift, and missing data noise varies primarily with dataset characteristics rather than model architecture. These findings suggest that data cleaning priorities should be informed by the specific noise profile of the operational environment rather than applied uniformly.

Keywords: ADS-B; Data Cleaning; Autoencoder; Algorithm Performance

1. Introduction

ADS-B has become a crucial data source for air traffic management, supporting a wide range of applications. Typical studies include trajectory prediction [1, 2], flight phase identification [3], trajectory clustering and modeling [4], safety analysis [5] (e.g., conflict detection and collision risk assessment), and airport operations optimization [6] (e.g., runway occupancy and taxiing analysis). ADS-B has also become a key enabler for environmental studies such as fuel consumption estimation, contrail detection, and large-scale emissions assessment. Despite these advances, ADS-B data quality remains a significant concern. Missing points, irregular sampling, and anomalies complicate processing and may bias analysis. To improve usability, researchers apply cleaning methods such as interpolation, smoothing, outlier removal, and resampling. However, these methods can also dis-

tort the statistical and physical characteristics of trajectories. For example, interpolation can mask subtle variations and increase prediction errors, while outlier removal may discard rare but genuine safety-critical events. Yet, existing studies rarely provide systematic and quantitative analysis of how different strategies influence downstream results. To address this gap, this study systematically evaluates the relationship between data cleaning and algorithmic performance in ADS-B analytics. Section 2 reviews eight major application domains of ADS-B data to contextualize its analytical value. Section 3 summarizes common data cleaning techniques and proposes a generalized preprocessing pipeline integrating detection, interpolation, and smoothing. Section 4 presents an autoencoder-based case study using three architecturally distinct models: a Fully Connected Autoencoder (FC-AE), a Long Short-Term Memory Autoencoder (LSTM-AE), and a Gated Recurrent Unit Autoencoder (GRU-AE). These models are evaluated across four geographically diverse airport datasets, quantifying the impact of four noise types (Gaussian, drift, spike, and missing data) on trajectory reconstruction performance, followed by a discussion of the observed patterns and their implications for cleaning strategy design. Finally, Section 5 concludes the study and outlines directions for future research. These analyses aim to provide a clearer understanding of how data quality shapes learning-based ADS-B algorithms and to inform the design of more robust, context-aware data processing strategies.

2. State of the Art

Early civil aviation primarily relied on primary and secondary radars, which had limitations such as limited detection range, insufficient information accuracy, and delayed updates. These shortcomings not only increased the navigation difficulty for long-distance flights but also posed safety risks [7]. To address these challenges, the aviation industry has gradually developed the ADS-B system through decades of exploration. Relying on the Global Navigation Satellite System (GNSS) and on-board sensors, ADS-B integrates information such as barometric altitude, inertial navigation, and airspeed measurements to generate flight status parameters. It then periodically broadcasts key data like identification codes, position, altitude, speed, and flight intentions via on-board equipment [8].

The development of ADS-B can be traced back to the 1970s. In 2003, the 11th Air Navigation Conference of the International Civil Aviation Organization (ICAO) [9] formally recognized ADS-B as a key surveillance tool for future air traffic management and promoted its standardization and application. After 2010, ADS-B entered the phase of large-scale global application. Countries have successively introduced regulations to promote its widespread use in aviation operations. Meanwhile, the application of space-based ADS-B [10] has enabled real-time, high-precision surveillance of approximately 70% of the world's airspace. Open platforms represented by the OpenSky Network [11] have also provided large-scale ADS-B data resources for academic research.

2.1 Current Usage of ADS-B Data

This section presents an overview of the current use of ADS-B data in the research domain, identifying and organizing clusters of algorithms and application areas. In this study, the collected literature is categorized into eight major domains, spanning from trajectory modeling and operational management to environmental sustainability and cybersecurity. These analyses provide a structured overview of the evolving research landscape surrounding ADS-B applications in aviation.

2.1.1 Paper Selection

To ensure both representativeness and research quality, we focused on journals and conferences with high academic impact in the fields of air traffic management (ATM) and digital aviation. The primary sources include the *Digital Avionics Systems Conference (DASC)*, *SESAR Joint Undertaking Annual Conference*, *Air Traffic Management Seminar (ATM Seminar)*, *International Conference on Research in*

Air Transportation (ICRAT), *Transportation Research Part C: Emerging Technologies*, *IEEE Transactions on Intelligent Transportation Systems*, and the *Journal of Air Transport Management (JATM)*. Literature retrieval was mainly conducted through academic databases such as IEEE Xplore, ScienceDirect, and Elsevier Scopus, as well as publicly available proceedings from the aforementioned conferences.

Considering that large-scale implementation and operational use of ADS-B systems began worldwide around 2012, this year was set as the starting point for the large-scale research phase of ADS-B data. Therefore, this study selected English-language publications issued between 2012 and December 2024 as the objects of analysis. We manually collected research that explicitly utilized real ADS-B flight data from the selected journals and conferences, excluding studies that relied solely on simulated or synthetic datasets. The detailed screening process was as follows:

- **Initial Screening:** Titles and abstracts were reviewed to confirm the study's relevance to the aviation domain, such as airspace optimization, trajectory prediction, or conflict detection and avoidance (DAA).
- **Keyword Filtering:** Only papers containing the term “ADS-B” in the title, abstract, or keywords were retained.
- **Data Authenticity Criterion:** Studies were required to clearly indicate the use of real ADS-B datasets. Papers using only simulated or artificially generated trajectories were excluded.
- **Duplication and Accessibility Review:** Duplicate publications and inaccessible preprints were removed to ensure the reproducibility and verifiability of the results.

After multiple rounds of screening and manual verification, a total of 145 papers were collected, covering representative applications of ADS-B data across diverse research domains. The distribution of the selected studies by source is illustrated in Figure 1.

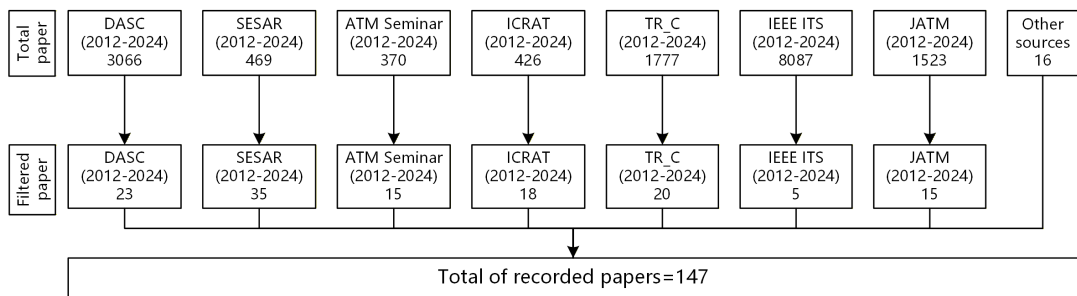


Figure 1. Paper collection flow of the review made on Digital Avionics Systems Conference (DASC), SESAR Joint Undertaking Annual Conference (SESAR), Air Traffic Management Research and Development Seminar (ATM seminar), International Conference on Research in Air Transportation (ICRAT), Transportation Research Part C: Emerging Technologies (TR_C), Journal of Air Transport Management (JATM), and IEEE Transactions on Intelligent Transportation Systems (IEEE Trans. on ITS).

2.1.2 Paper Clustering and Categorization

The screened publications provided a solid data foundation for the categorization and trend analysis of ADS-B applications in this study. To systematically organize the characteristics and focal points of different research directions, we employed a mixed quantitative–qualitative approach for feature extraction and clustering of the collected literature.

Using Excel spreadsheets and reference management tools, the following key features were extracted from each selected conference and journal: Publication year; Source conference or journal; Paper title and author keywords; Application scenario; Research focus or analytical perspective; Methods and algorithmic approaches (e.g., machine learning, optimization models, statistical analysis, sim-

ulation frameworks); Description of the ADS-B dataset used (e.g., public data repositories, airport-specific data, or crowdsourced datasets).

In the preliminary organization stage, the literature was grouped into four broad directions: trajectory prediction, air traffic management, aircraft performance estimation, and environmental sustainability. However, a subsequent semantic and thematic analysis revealed substantial overlaps and hierarchical relationships among these themes. Therefore, this study reclassified the literature through a combined thematic synthesis approach, considering each study's research objectives, data utilization patterns, and the functional role of ADS-B. Finally, ADS-B-related studies were categorized into eight major domains: Trajectory Modeling and Prediction, Operational Optimization and Management, Operational Safety and Surveillance, Aircraft Performance and Efficiency, Data Engineering and Enhancement, Environment and Sustainability, Security and Cybersecurity, Methodology, Simulation, and Policy. This classification framework establishes the structural foundation for the domain-specific analyses presented in the following sections. The classification results, derived from thematic induction and semantic grouping, are summarized in Table 1, illustrating the representative applications and algorithms across the eight research domains. The complete list of clustered and structured references is provided in the supplementary material ("ADS-B_papers.xlsx").

Table 1. Summary of ADS-B Application Categories, Main Methods, Data Roles, and Key Data Quality Requirements

Category	Main Methods	Role of ADS-B Data	Key Data Quality Requirements
Trajectory Modeling and Prediction	Clustering (DBSCAN, K-means); LSTM/Transformer prediction; AE feature extraction; DTW ^a ; PCA; Hybrid physical-data models.	Core data source	Positional accuracy, temporal continuity
Operational Optimization and Management	MILP ^a ; Simulated annealing; Heuristic algorithms; KPI ^b -based performance metrics; Historical traffic analysis.	Real-time/historical traffic input	Coverage, completeness, aggregate consistency
Operational Safety and Surveillance	DAA geometric models; Anomaly detection (thresholds, clustering, autoencoder, GMM ^c); Monte Carlo risk evaluation.	Flight monitoring and safety baseline	Inter-aircraft relative accuracy, update rate
Aircraft Performance and Efficiency	Dynamic equation inversion; Maximum likelihood estimation; Bayesian inference; Particle filtering; Regression and neural networks.	Model calibration data source	Altitude and speed accuracy, bias sensitivity
Data Engineering and Enhancement	Kalman filtering; Map-matching; Multisource fusion; Data indexing; Generative models (TimeGAN ^d).	Primary processing object	Raw data quality (processing target)
Environment and Sustainability	Trajectory-based emission estimation; Remote sensing data fusion; Optimal control route planning.	Environmental assessment input	Vertical profile accuracy, trajectory completeness
Security and Cybersecurity	Intrusion detection (ML ^e classifiers); Protocol vulnerability testing; SDR ^f -signal analysis.	Research target	Message-level integrity and authenticity
Methodology, Simulation, and Policy	Open-source simulation platform development; Data standardization; Policy and privacy analysis.	Research infrastructure and policy object	Spatial and temporal coverage

a ML = Machine Learning; DTW = Dynamic Time Warping; KPI = Key Performance Indicator; GMM = Gaussian Mixture Model; MILP = Mixed-Integer Linear Programming; SDR = Software-Defined Radio; TimeGAN = Time-series Generative Adversarial Network.

Trajectory Modeling and Prediction. This domain focuses on modeling and predicting aircraft trajectories based on historical and real-time data. Core tasks include four-dimensional trajectory prediction, ETA estimation, trajectory clustering, and uncertainty quantification. As a core data source, ADS-B provides continuous, high-precision position, velocity, and altitude data that determine model accuracy. Gui et al. [12] proposed a semantic trajectory representation for arrival flight clustering to support airspace design, flow management, and ETA estimation; Wang et al. [13] applied Principal Component Analysis (PCA)-based dimensionality reduction and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering for preprocessing, followed by a Multi-Cell Neural Network (MCNN) for short-term trajectory prediction in terminal maneuvering areas; and Wang et al. [14] integrated clustering-based preprocessing with hybrid MCNN models to improve Estimated Time of Arrival (ETA) prediction accuracy.

Operational Optimization and Management. This domain focuses on improving the overall efficiency of airspace and airport operations, encompassing air traffic flow management, surface operations (taxiing and sequencing), terminal maneuvering area coordination, and airspace structure optimization. ADS-B data provide continuous, fine-grained historical and real-time traffic information, serving as reliable inputs for optimization models and decision-support systems, and enabling accurate performance evaluation and data-driven strategy development. Research in this area often applies Linear Programming, Simulated Annealing, and heuristic algorithms to address sequencing, scheduling, and routing problems. Other studies employ queuing models and Key Performance Indicators (KPIs) for operational assessment or mine historical ADS-B data to identify bottlenecks such as taxiway congestion and sector capacity limits. Basora et al. [15] combined DBSCAN clustering with Random Forest regression for sector occupancy prediction, and Delahaye et al. [16] used hierarchical clustering with Transformer models for flow pattern detection and capacity management.

Operational Safety and Surveillance. This research domain aims to enhance aviation safety and situational awareness through data-driven analysis. It covers conflict detection and resolution (Detect-and-Avoid (DAA) / Airborne Collision Avoidance System (ACAS)), abnormal event detection (e.g., go-arounds, unstable approaches), assessment of collision risk and airspace complexity, and performance evaluation of surveillance systems. As an independent surveillance source, ADS-B data provide continuous and high-precision trajectory and state information, enabling real-time monitoring of aircraft behavior, detection of potential conflicts and anomalies, and quantitative assessment of operational safety. Bonifazi et al. [17] identified unstable approaches and go-arounds using ADS-B data, employing rule-based methods and Gaussian Mixture Models for anomaly detection and integrating runway and weather information for improved accuracy. Rorie et al. [18] conducted the first real-world evaluation of the ACAS Xr airborne collision avoidance system. Zhang et al. [19] investigated conflict-free routing strategies and compared multiple optimization algorithms, while Bao et al. [20] proposed a multi-airport terminal area risk prediction framework to assess inter-airport conflict probabilities.

Aircraft Performance and Efficiency. This research area focuses on deriving aircraft performance parameters from flight data to calibrate or complement existing models such as Base of Aircraft Data (BADA), and to evaluate energy efficiency across aircraft types and flight phases. Key parameters include aircraft mass, drag polar, thrust settings, fuel consumption, and speed profiles. In this context, ADS-B data provide essential flight state information, such as ground speed, vertical rate, and heading—enabling large-scale, fleet-level performance analysis even in the absence of detailed design data. This supports more accurate and data-driven model calibration and validation. Sun et al. [21] developed a probabilistic framework to estimate aerodynamic parameters from operational data; Schultz et al. [22] integrated Flight Data Recorder (FDR) and ADS-B data to model fuel consumption and operational efficiency using machine learning methods; and Alligier et al. [23] predicted aircraft mass and speed intent during climb to enhance physics-based trajectory prediction.

Data Engineering and Enhancement. This category focuses on improving the quality and usability of raw ADS-B data, which form the foundation for subsequent analytical and modeling applications. Key tasks include anomaly detection, missing-value imputation, multi-source data fusion, data compression and indexing, and synthetic data generation. In this domain, ADS-B data themselves are the core subject of engineering—aimed at producing cleaner, more complete, and more interoperable datasets that support trajectory prediction, operational analysis, and safety evaluation. Tabassum et al. [24] conducted long-term statistical analysis to identify anomalies and assess the impact of systematic errors on trajectory accuracy. Wandelt et al. [25] introduced an efficient compression and indexing framework to enable scalable querying and analytics of large-scale ADS-B records. Spinielli et al. [26] developed a reproducible reference trajectory dataset by integrating multiple surveillance sources for performance assessment under the EUROCONTROL Performance Review Commission (PRC) initiative.

Environment and Sustainability. This research area focuses on quantifying the environmental impact of aviation operations and exploring sustainable optimization strategies, including greenhouse gas and pollutant emission assessment, contrail formation detection and avoidance, and noise evaluation. Owing to its wide coverage and high temporal resolution, ADS-B data serve as a crucial source for environmental modeling and validation. For instance, Roosenbrand et al. [27] proposed a method to estimate contrail altitudes using shadows in Landsat satellite imagery, with ADS-B data employed as ground truth for validation. Sun et al. [28] integrated satellite-based and ground-based ADS-B data with wind field information to improve emission estimation and compared actual flight trajectories with optimal routes to quantify excess emissions.

Security and Cybersecurity. This domain focuses on identifying and mitigating cybersecurity threats targeting the ADS-B system itself, such as False Data Injection Attacks (FDIA), signal spoofing, and message tampering, to ensure the integrity and reliability of surveillance information. In this field, the ADS-B protocol, signal, and data link are the direct subjects of vulnerability analysis and protection technology research. For example, Cretin et al. [29] proposed a Domain-Specific Language-based testing framework to evaluate the resilience of Air Traffic Control systems against FDIA, while Khan et al. [30] employed machine learning techniques for ADS-B intrusion detection.

Methodology, Simulation, and Policy. This category provides foundational tools, frameworks, and policy support for aviation research. It includes the development of open-source simulation platforms, advocacy of reproducible research practices, establishment of data standards, and discussion of regulatory and privacy issues related to ADS-B deployment. In this context, ADS-B serves both as input data for constructing realistic scenarios in simulation environments and as a focal topic in advancing data-sharing policies, privacy protection, and industry standards. Mehlitz et al. [31] proposed the Runtime for Airspace Concept Evaluation (RACE) framework for comprehensive airspace data analysis, while Bolic et al. [32] systematically elaborated on the European ATM Open Science Alliance and its Open Performance Data Initiative (OPDI), which aim to foster transparency and open research in the ATM domain.

To illustrate the application domains of ADS-B data, a quantitative analysis was conducted on 147 valid papers according to the eight classification categories, as shown in Figure 2. Since some studies involve multiple domains (e.g., Data Engineering and Aircraft Performance Calculation), they were counted in each relevant category; thus, the total number in the chart exceeds 147.

Overall, Operational Optimization and Management dominates ADS-B research (30.8%), followed by Operational Safety and Surveillance (23.1%) and Trajectory Modeling and Prediction (18.6%). Together, these account for over 70% of all studies, reflecting strong alignment with ADS-B's core functions—enhancing operational efficiency and supporting real-time safety monitoring. In contrast, Security and Cybersecurity (2.6%), Environment and Sustainability (5.1%), and Methodology,

Simulation, and Policy (5.1%) remain less represented, indicating emerging but underexplored areas.

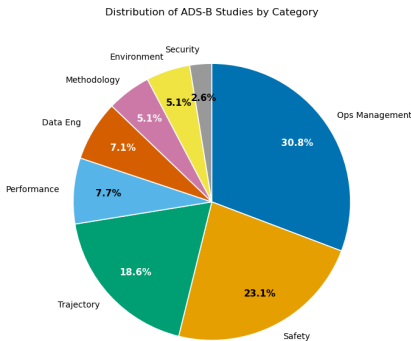


Figure 2. Distribution of reviewed ADS-B studies by application category. Abbreviated labels correspond to the eight categories defined in Table 1.

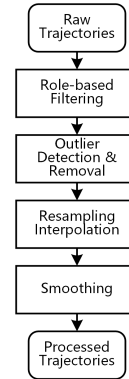


Figure 3. Systematic Cleaning Pipeline of ADS-B Trajectories (Raw to Processed) Based on Literature Survey and Taxonomic Summary.

3. Summary of ADS-B Cleaning Methods and Workflow

3.1 Data Quality Issues in ADS-B Surveillance

Real-world ADS-B data often contain a considerable amount of noise and anomalous errors. Several studies have conducted in-depth analyses of these issues. For instance, Tabassum et al. [24] systematically demonstrate various types of anomalies found in ADS-B messages, while [33] and [34] provide detailed examinations of noise sources and error mechanisms within crowdsourced datasets. These studies collectively indicate that ADS-B data quality is heavily influenced by factors such as hardware performance, signal environment, and network structure, resulting in inconsistencies and unreliability across raw observations.

The quality issues of ADS-B data can generally be characterised along four dimensions. *Completeness* concerns message loss, missing fields, and update interruptions that produce temporal or spatial gaps in recorded trajectories. *Consistency* refers to internal contradictions across data attributes, such as out-of-order timestamps, abrupt altitude jumps, unrealistic speed values, or discrepancies between barometric and geometric altitude, which reflect different measurement definitions rather than instrument error. *Accuracy* addresses systematic deviations between reported values and actual flight states, including positional errors and speed biases arising from quantization, receiver limitations, or environmental interference. *Reliability* captures random noise, falsified messages, and artefacts introduced by multi-source fusion, which are particularly prevalent in open, crowdsourced collection environments.

In summary, errors may occur throughout the collection, transmission, and aggregation of ADS-B data. Without proper treatment, these problems can severely affect the performance of downstream algorithms and compromise the reliability of analytical results. Therefore, systematic data cleaning and quality control are crucial to ensure the usability and accuracy of ADS-B data.

3.2 Data Cleaning Methods

3.2.1 Task-Oriented Filtering

Filtering ADS-B data is typically the initial step in data cleaning. Depending on the research objectives and application scenarios, most studies perform preliminary filtering of raw ADS-B data before

analysis to ensure that the data used are relevant and representative. Common filtering strategies can be broadly categorized into range-based filtering and attribute-based filtering.

Range-Based Filtering This approach primarily selects ADS-B data based on temporal or spatial ranges. Temporal filtering can limit the data to specific seasons, dates, or time periods to match the study timeframe. Spatial filtering focuses on particular routes, airspaces, or airport operations. Additionally, trajectories within specific geographic boundaries (e.g., latitude/longitude ranges or airspace altitude layers) may be extracted to construct local operational networks or airspace models.

Attribute-Based Filtering Beyond temporal and spatial constraints, researchers may remove trajectories that are irrelevant or do not meet task-specific criteria. This type of filtering is often based on flight rules, operational states, aircraft types, or flight phases. For example, Dhief et al. [35] excluded flights operating under Visual Flight Rules (VFR) in a go-around behavior study. Liu et al. [36] filtered trajectories under consistent weather conditions in their study on taxiing optimization at an airport to reduce the influence of weather variations and runway configurations.

3.2.2 Outlier Detection and Removal

After the initial filtering, it is usually necessary to identify and remove outliers in order to ensure the reliability of subsequent analyses. [34] summarized and investigated common methods for outlier detection and handling. Common outlier handling methods include the removal of entire trajectories, local cleaning of individual abnormal points, and automated detection based on clustering or deep learning techniques.

The most common approach is the removal of entire trajectories. Outlier trajectories can arise from various reasons, the most frequent being incomplete data, such as trajectories with too few sampling points to accurately represent the flight process, which need to be excluded. For trajectories that are generally valid but contain a few abnormal points, researchers typically perform local cleaning. Common types of noise include duplicate points and physically impossible “jump points”. Methods like Gaussian filtering or particle filtering are often used to correct these anomalies.

Additionally, DBSCAN is extensively employed for outlier detection and data cleaning. It identifies anomalous points based on local density distribution, effectively separating them from normal trajectories while simultaneously performing clustering. Compared with traditional filtering approaches, DBSCAN provides enhanced adaptability and automation, particularly for trajectory datasets with uneven spatial distributions.

The AE model, as a deep learning-based method, can likewise be applied for anomaly detection. By learning representative patterns of normal trajectories, the AE detects outliers through elevated reconstruction errors. Its ability to capture nonlinear relationships makes it well suited for high-dimensional, time-series ADS-B data, and it can be integrated with clustering or filtering methods to achieve more refined data cleaning.

3.2.3 Interpolation and Resampling

In ADS-B data processing and trajectory reconstruction, interpolation and resampling are two essential preprocessing techniques. Interpolation focuses on repairing missing data points and ensuring trajectory continuity, while resampling aims to unify the temporal or spatial distribution of data, thereby improving the stability of subsequent analysis and model training. Since both techniques are often applied together in practice, they are presented here in an integrated discussion.

(1) Interpolation

The objective of interpolation is to estimate missing values between known points, thereby converting discrete trajectories into continuous and smooth curves. Depending on the fitting principle,

common interpolation methods can be categorized into three groups:

Linear and Polynomial Interpolation This is the most widely used class of interpolation techniques, which assumes that variations between adjacent points follow a linear or low-order polynomial relationship. These methods are computationally efficient and suitable for short time intervals or smooth motion, but their ability to capture nonlinear behavior such as turning or climbing is limited. Representative methods include linear interpolation [37] and polynomial interpolation.

Spline-Based Interpolation Spline methods fit piecewise polynomials with continuity at segment boundaries, ensuring smoothness and stability. Typical examples include linear spline interpolation [38], cubic spline interpolation [39], and piecewise cubic Hermite interpolation [40], which introduces shape-preserving constraints to prevent unrealistic oscillations.

Spatially Adaptive Interpolation This approach ensures consistent spatial resolution along the trajectory, achieving globally uniform point density while preserving geometric accuracy and spatial consistency.

(2) Resampling Methods

Resampling aims to transform irregularly spaced ADS-B data into a unified format suitable for downstream analysis and model input. According to the dimension of unification, resampling techniques can be classified into three categories:

Fixed-Time Interval Resampling This method extracts or generates data points at a fixed temporal interval, ensuring uniform time distribution along the trajectory. It is the most fundamental form of temporal standardization, with sampling intervals ranging from one second [41] to several minutes [42], depending on the temporal resolution required by the study.

Trajectory Feature-Based Resampling This approach resamples according to geometric characteristics such as turning points or curvature changes, thereby reducing redundancy while preserving essential trajectory features. The representative algorithms are the Ramer-Douglas-Peucker (RDP) algorithm [43], which iteratively removes points with distances below a given threshold from the line connecting the start and end points, retaining only key inflection points. This reduces data volume while maintaining the overall geometric structure of the trajectory. Fixed Number of Inputs algorithm: Uses interpolation to map each trajectory into a fixed number of points, ensuring consistent input dimensions for deep learning models such as AE [44] and Transformers [20].

Spatial or Curve-Based Resampling This category emphasizes spatial uniformity and trajectory smoothness. Points are sampled at fixed spatial intervals to ensure consistent spatial density, which is particularly useful for analyses such as airport vicinity trajectory mapping or taxiway path planning, where uneven temporal sampling could otherwise introduce spatial distortions.

3.2.4 Smoothing

After resampling, researchers often apply trajectory smoothing to further suppress noise, reduce trajectory jitter, and preserve the essential motion trend, thereby providing more reliable inputs for subsequent analysis and model construction. According to their underlying principles and computational characteristics, trajectory smoothing methods can generally be categorized into three groups:

Model-Based Filtering Methods These methods rely on state-space or probabilistic estimation models to describe the relationship between the aircraft's true motion states and observational noise, achieving optimal trajectory estimation and smoothing. Representative algorithms include the Kalman Filter [5] and Extended Kalman Filter, which obtain optimal state estimates by minimizing the covariance of recursive estimation errors. Owing to their strong dynamic modeling capability

and physical interpretability, such methods are widely used for aircraft state estimation and altitude smoothing tasks.

Signal-Processing-Based Filtering Methods In this approach, the trajectory is treated as a time-series signal, and digital filters are employed to suppress undesired frequency components, thus achieving trajectory smoothing. Typical examples include the finite impulse response (FIR) low-pass filter [6], the Exponential Moving Average (EMA) algorithm [45], and the bilateral window averaging method [46]. By convolutional or recursive operations, these methods effectively remove oscillatory noise from uniformly sampled trajectories.

Curve-Fitting and Geometric-Statistical Methods These methods approximate the entire trajectory using mathematical curves or geometric-statistical representations to achieve global-level smoothing, producing continuous and geometrically consistent trajectories. For instance, the smoothing cubic spline [47] is a variational fitting technique that balances data fidelity and smoothness through an optimized regularization parameter. The Hough voting algorithm [36], based on the global geometric consistency of trajectories, maps local trajectory features into a parameter space and serves as a common tool for geometric trajectory reconstruction.

3.2.5 ADS-B Data Cleaning Pipeline

Overall, ADS-B data cleaning typically follows a logical progression from macroscopic filtering to microscopic refinement. As illustrated in Figure 3, the cleaning process generally consists of the following steps:

First, raw ADS-B data are filtered by temporal, spatial, and flight-related attributes to extract subsets suitable for analysis. Next, preliminary denoising removes duplicates, eliminates trajectories with excessive missing data, and corrects abrupt anomalies, improving completeness and consistency. Then, interpolation and resampling fill missing values and standardize temporal or spatial resolution, creating a structured basis for algorithmic processing. Finally, trajectory smoothing suppresses residual noise and jitter, highlighting the core motion trends that reflect true flight dynamics.

In recent years, several open-source tools have provided comprehensive support for implementing the above data-cleaning procedures. Among them, the Traffic Library [48] has been widely adopted. It offers ready-to-use implementations for each stage of the workflow, greatly simplifying the pre-processing of ADS-B data. Researchers can transform raw data into high-quality trajectories without developing low-level algorithms manually, which significantly lowers the technical barrier to aviation data processing and allows greater focus on downstream analytical tasks.

It should be noted that the above workflow is derived from our review and comparative analysis of the collected papers. However, the survey reveals significant inconsistencies in how data-cleaning procedures are described across studies. Some works mention only general terms such as “filtering” or “pre-processing” without detailing the specific methods or parameter configurations. This lack of transparency and consistency undermines the reproducibility and comparability of research outcomes and may also affect the reliability of conclusions regarding algorithmic performance.

3.3 Analysis of Data Cleaning Operations

The previously discussed data cleaning procedures can enhance the completeness, consistency, accuracy, and reliability of ADS-B data. However, their impact on downstream algorithms is complex and multifaceted.

Outlier Removal: Removing outliers reduces noise and enhances data quality, but misclassifications may eliminate legitimate maneuvers, such as temporary avoidance, lowering the accuracy of flight pattern modeling. Excessive removal of marginal data can also reduce sample size and limit the rep-

representativeness of rare conditions, such as adverse weather or operations in remote airspace. Among the methods discussed, DBSCAN can misclassify sparse but normal points in unevenly distributed trajectories, while AE-based methods depend on sufficient, high-quality normal data; biases in the training set may shift detection thresholds, causing normal trajectories to be wrongly flagged as outliers.

Interpolation and Resampling: These techniques fill missing points and standardize temporal resolution to improve trajectory continuity and comparability. However, excessive or improper interpolation may obscure genuine micro-maneuvers, while resampling can introduce artificial variations in speed and acceleration. Such distortions affect models that rely on short-term motion states, reducing their ability to capture key dynamics and increasing prediction errors. Moreover, interpolation may blur aircraft proximity, weakening conflict detection and risk identification.

Smoothing: These filters effectively suppress high-frequency jitter in positional data, producing smoother trajectories and facilitating the calculation of derived features such as heading and curvature. Nevertheless, smoothing can weaken sharp trajectory characteristics, such as the precise onset and recovery points of turns, which may negatively affect maneuver-based anomaly detection (e.g., go-around identification) and flight phase classification models.

Overall, ADS-B data cleaning is a crucial step in improving data integrity and reliability. However, as the above limitations indicate, over-cleaning may remove essential flight characteristics, while insufficient cleaning may fail to meet the quality requirements for algorithmic processing. Both extremes can adversely affect downstream analysis and model performance.

Therefore, in practical applications, researchers need to balance data fidelity and usability when designing preprocessing pipelines. In the following chapter, the influence of different cleaning strategies on algorithm performance will be quantitatively evaluated through experiments.

4. Use Case: Investigation of the Noise Impact on Autoencoder Algorithms

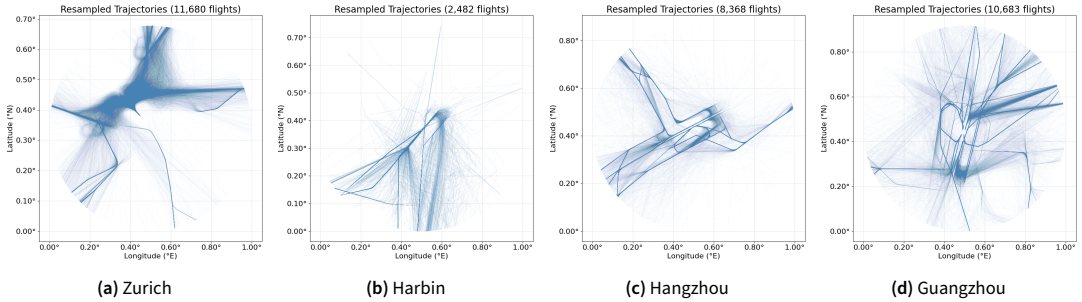
This section presents a complete case study that quantifies how four common ADS-B noise types – Gaussian noise, drift, spike, and missing data – affect the trajectory reconstruction performance of autoencoder (AE) models, and thereby informs how data-cleaning efforts should be prioritised across noise types rather than applied uniformly.

Datasets. Experiments are conducted on four terminal-area datasets summarised in Table 2. Zurich Airport (LSZH) is a major European hub with complex approach procedures; Harbin Taiping (ZYHB) is a mid-scale Northeast Asian airport with low traffic density; Guangzhou Baiyun (ZGGG) is one of China's highest-density hubs; and Hangzhou Xiaoshan (ZSHC) is a high-growth regional hub adjacent to the Shanghai TMA, where overlapping sector coordination produces characteristic lateral deviations. The three Chinese datasets share a common data source distinct from the European set, each covering two weeks (28 Jul – 3 Aug and 21–28 Dec 2025); the Zurich dataset spans 1 Oct – 30 Nov 2019. All records contain timestamp, latitude, longitude, altitude, ground speed, heading, callsign, and ICAO 24-bit address.

Preprocessing. Raw ADS-B tracks were filtered to retain only flights operating within 20 NM of the target airport. Continuous streams were segmented into individual flights at ground-dwell gaps exceeding 600 s, after which duplicate timestamps, missing position/altitude records, position jumps above 1 km, and segments with fewer than 50 reports were removed. Each surviving trajectory was resampled to $T = 100$ points by arc-length-parameterised linear interpolation. Coordinates were normalised to $[0, 1]$ using a shared range factor that preserves geographic aspect ratio, then standardised with a StandardScaler before model input. Table 2 reports post-preprocessing trajectory counts; Figure 4 visualises the four resulting datasets.

Table 2. Overview of the four ADS-B datasets used.

	Zurich (LSZH)	Harbin (ZYHB)	Hangzhou (ZSHC)	Guangzhou (ZGGG)
Region	Europe	Northeast Asia	East China	South China
Trajectories	11,680	2,482	8,368	10,683
Time span	2 months	2 weeks	2 weeks	2 weeks
Points per trajectory	200	1,088	1,322	1,084
Spatial extent (lat × lon)	218 × 147 km	285 × 200 km	232 × 200 km	218 × 200 km

**Figure 4.** Preprocessed ADS-B trajectory visualizations for each dataset.

4.1 Autoencoder Model

Table 3. Architecture and training configurations of the three autoencoder models.

	FC-AE	LSTM-AE	GRU-AE
Encoder	Linear(200→128→64→32), ReLU	2-layer LSTM (hidden=128) + FC(128→32)	2-layer GRU (hidden=128) + FC(128→32)
Decoder	Linear(32→64→128→200), ReLU	FC(32→128) + 2-layer LSTM (hidden=128) + FC(128→2)	FC(32→128) + 2-layer GRU (hidden=128) + FC(128→2)
Latent dim	32	32	32
Parameters	72.3 K	472.5 K	356.5 K
Optimizer	Adam	Adam	Adam
Learning rate	1×10^{-3}	1×10^{-3}	1×10^{-3}
Epochs	200	200	200
Batch size	64	64	64
Loss	MSE	MSE	MSE

We evaluate three autoencoder variants that differ in how they process trajectory sequences: a fully connected autoencoder (FC-AE), an LSTM-based autoencoder (LSTM-AE), and a GRU-based autoencoder (GRU-AE). Their configurations are summarised in Table 3.

The FC-AE treats each trajectory as a flattened vector and compresses it through stacked linear layers with ReLU activations. Because the input is flattened, it has no explicit notion of temporal order and relies entirely on global spatial structure, serving as a lightweight baseline. The LSTM-AE and GRU-AE instead process the trajectory as an ordered sequence of coordinate pairs, using 2-layer recurrent encoders to produce a context vector that is then projected into the latent space; the decoders reverse this process, reconstructing the trajectory point by point. Including both variants allows us to examine whether the additional gating mechanism of the LSTM provides any advantage

over the more parameter-efficient GRU for trajectories of this length.

The latent dimension is set to $z = 32$ for all three architectures. Prior autoencoder-based studies on aircraft trajectories have typically adopted latent dimensions in the range of 64, notably Olive and Basora [49] on en-route track angles ($d = 150$, $z = 64$) and Krauth et al. [50] on terminal-area trajectories at Zurich airport ($z = 64$). We adopt a smaller value to reduce training time; preliminary experiments with $z = 64$ produced no qualitative change in the noise sensitivity results reported in Section 4.4.

All three models share an identical training protocol: Adam optimiser, learning rate of 1×10^{-3} , batch size of 64, and 200 training epochs, so that any performance differences observed in the subsequent experiments can be attributed to architectural properties rather than training configuration. The mean squared error between the input and reconstructed trajectory is used as the loss function, consistent with the continuous-valued nature of the coordinate data.

4.2 Baseline Generation

Each trained AE model identifies trajectories most similar to clean data as the baseline. It reconstructs all normalized samples and ranks them by reconstruction error, selecting those with the lowest errors. Trajectories that are highly reconstructable, which show minimal error, are considered the most representative and clean within the dataset. As shown in Figure 5, the baseline trajectories exhibit high smoothness and spatial consistency, conforming to the physical laws of real flight paths. In contrast, high-error samples, shown in Figure 6, often contain data anomalies or noisy points. Using AE reconstruction error as the selection criterion enables automatic identification of high-quality trajectories without manual thresholds or interpolation. This method helps avoid errors or inappropriate parameter settings that can arise during manual preprocessing, thereby producing a statistically sound and model-adaptive baseline dataset. In total, 100 high-quality trajectories were selected as the baseline dataset for the experiment.

Since the three architectures weight different trajectory features, the FC-AE captures global spatial structure while the LSTM-AE and GRU-AE are additionally sensitive to temporal regularity – each model selects its own set of 100 baseline trajectories independently. This model-adaptive selection avoids manual thresholds or interpolation and ensures that each architecture is tested on the trajectories it considers cleanest.

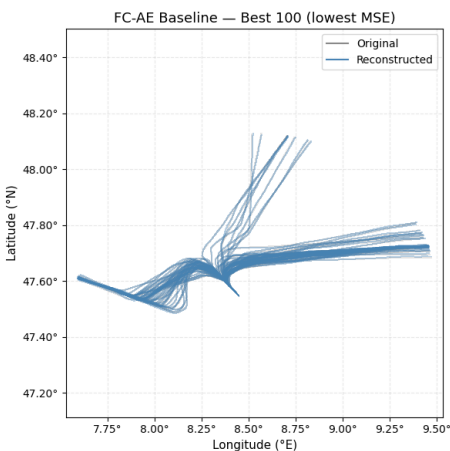


Figure 5. The best 100 trajectories of Zurich airport reconstructed by trained AE model.

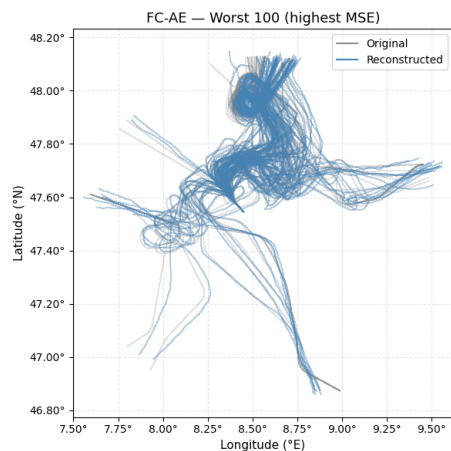


Figure 6. The worst 100 trajectories of Zurich airport reconstructed by trained AE model.

4.3 Noise Injection

To systematically analyze the impact of different noise types on AE reconstruction performance, four typical artificial noise sources were injected into the baseline trajectories: Gaussian noise, drift noise, spike noise, and missing data. These correspond to the most common sources of error in ADS-B data and collectively simulate disturbances that occur during data acquisition, transmission, and decoding. The experiment procedure is shown in Figure 7. All noise is applied in the normalised $[0, 1]$ coordinate space.

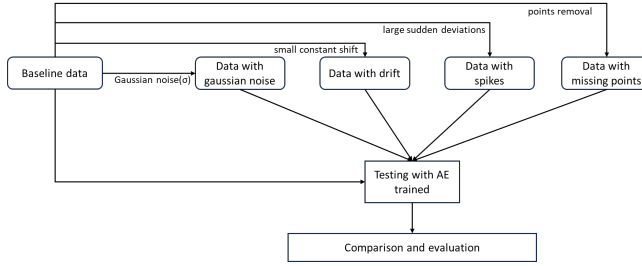


Figure 7. Experimental flowchart for evaluating reconstruction performance of the AE model.

Gaussian noise simulates random measurement errors in ADS-B signals during reception or localization, like GNSS range errors, multipath effects, or receiver thermal noise, resulting in small position fluctuations. It is generated by adding Gaussian perturbations to each coordinate of the baseline trajectories. The standard deviation σ is swept from 0 to 0.005, which corresponds to approximately 0.6–1.4 km in physical distance depending on the airport’s spatial extent, covering the range from nominal GPS accuracy up to and beyond the 300 m horizontal position limit specified by current surveillance standards [51]. Drift noise simulates systematic positioning deviations caused by GNSS reference drift, sensor bias, or timestamp misalignment, and is implemented by adding a fixed spatial offset to the trajectory. The offset magnitude follows the same sweep range as Gaussian noise (σ from 0 to 0.005). Spike noise simulates sporadic position jumps caused by data encoding issues or transponder faults, and is implemented by perturbing random trajectory points with sharp amplitude variations. Since the frequency of such jumps is not precisely quantified in the existing literature, the spike probability is swept over a broad range from 0 to 0.30 to capture both rare and frequent outlier scenarios. Missing data simulates the loss of position reports due to signal dropout or message corruption. It is implemented by removing a specified proportion of trajectory points and reconstructing the sequence through linear interpolation. The missing rate is swept from 0 to 0.50, encompassing the loss rates of 20–50% reported for ground-based ADS-B reception under varying traffic densities [11].

The sweep design for all four noise types deliberately extends beyond typical real-world magnitudes to reveal the full degradation profile of each architecture.

4.4 Metrics and Results

We use the root mean squared error (RMSE) to quantify the reconstruction quality of each autoencoder under noisy input conditions. For a trajectory consisting of T timesteps, the RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T [(\hat{x}_t - x_t)^2 + (\hat{y}_t - y_t)^2]} \quad (1)$$

where (x_t, y_t) denotes the normalised coordinates of the original baseline trajectory at timestep t , and (\hat{x}_t, \hat{y}_t) denotes the autoencoder’s reconstructed output. For each noise type and intensity level, noise is injected into 100 baseline trajectories and the corrupted trajectories are passed through the trained autoencoder. The reported RMSE is the median over the 100 trajectories, with shaded regions indicating the interquartile range (IQR).

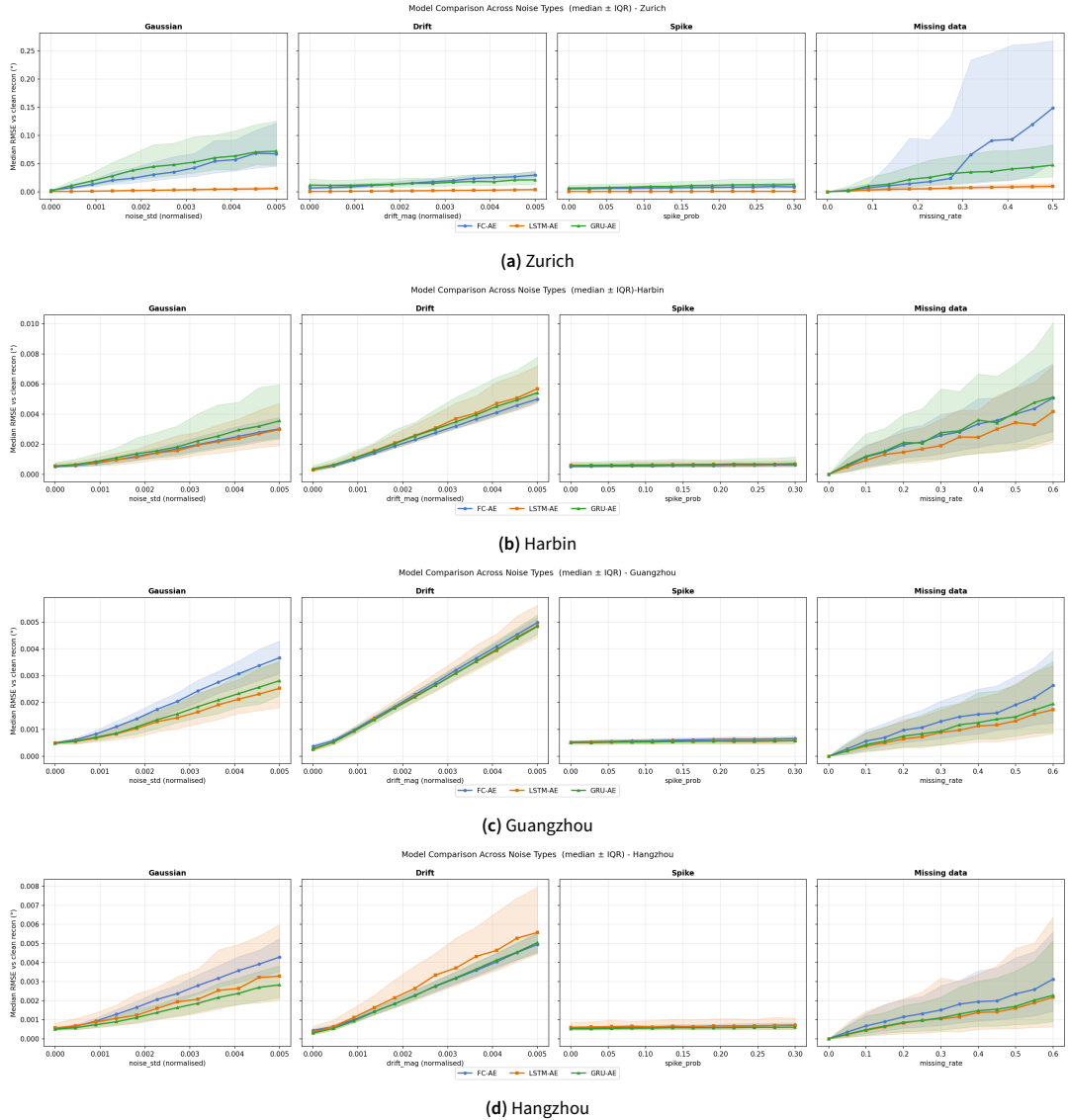


Figure 8. Median RMSE (with IQR shading) of three autoencoder architectures across four noise types for each airport dataset. Each row corresponds to one dataset; columns correspond to Gaussian, drift, spike, and missing data noise, respectively.

Figure 8 presents the experimental results for all four airport datasets. Each subfigure displays the RMSE trends of the three models across the four noise types in a single row. Because the three architectures differ considerably in absolute reconstruction precision—LSTM-AE typically achieves substantially lower RMSE than FC-AE and GRU-AE—the degradation trend of certain models may

appear visually compressed when plotted on a shared vertical axis. the appendix figures present each architecture separately, where the degradation trends are individually visible.

Both Gaussian and drift noise types produce a monotonically increasing RMSE across all model–dataset combinations. Gaussian noise, as a random perturbation applied independently to each point, results in a relatively wide IQR that reflects inter-trajectory variability. Drift noise, as a constant spatial offset applied uniformly to the entire trajectory, produces a tighter RMSE distribution with a steeper and more consistent degradation curve.

Spike noise consistently produces the lowest RMSE across all twelve model–dataset combinations. Even at the highest spike probability, the RMSE remains well below the levels observed for the other three noise types at moderate intensities. This suggests that the bottleneck structure of the autoencoder favours learning global trajectory patterns over local anomalies, conferring inherent robustness to sparse point-wise perturbations.

Missing data exhibits a nonlinear degradation profile: RMSE increases gradually at low missing rates and accelerates markedly beyond approximately 30%. This noise type also produces the largest IQR at high missing rates, indicating substantial variation in how individual trajectories respond to data loss.

Spike noise ranks last in sensitivity across all twelve model–dataset combinations, making it the only fully stable finding. The relative ranking among the remaining three noise types—Gaussian, drift, and missing data—varies by dataset, driven primarily by airport spatial characteristics rather than model architecture. Within a given dataset, all three architectures produce concordant sensitivity rankings, indicating that noise sensitivity patterns are jointly determined by data properties and noise geometry, with model architecture playing a comparatively limited role.

To verify the statistical reliability of the observed trends, we apply the Wilcoxon signed-rank test ($\alpha = 0.05$) to all pairs of adjacent noise levels. This non-parametric paired test is appropriate for the non-normally distributed paired samples in this experiment. The test results are consistent with the visual trends: drift noise achieves statistical significance at nearly all adjacent-level pairs; Gaussian noise is significant at low-to-moderate intensities, with some model–dataset combinations showing saturation at higher levels; spike noise yields a large number of non-significant pairs, consistent with its flat RMSE curve; and missing data shows weak significance at low missing rates but near-universal significance at high rates, reflecting its nonlinear degradation profile. Full test results are provided in the appendix tables.

4.5 Discussion

The above results carry three practical implications for ADS-B data cleaning. First, different noise types have fundamentally different effects on autoencoder reconstruction quality, and cleaning resources should be allocated accordingly rather than uniformly. Spike noise has the least impact on reconstruction across all combinations, indicating that the bottleneck structure of the autoencoder inherently filters sparse local anomalies; in workflows where autoencoders serve as the downstream analytical tool, detailed detection and removal of position jumps is not a priority. Drift noise, by contrast, causes significant reconstruction degradation even at small magnitudes, and systematic spatial offsets should therefore be treated as the primary cleaning target. The nonlinear threshold effect observed for missing data suggests that low levels of data loss can be tolerated through simple interpolation, but beyond a certain proportion the distortion introduced by interpolation escalates rapidly, at which point more sophisticated imputation strategies or outright trajectory exclusion become necessary.

Second, the relative sensitivity ranking among Gaussian, drift, and missing data noise varies across

datasets, driven primarily by airport spatial characteristics. This implies that no universal cleaning priority scheme exists. For a new airport or airspace, researchers should first profile the noise characteristics of the target data before determining the cleaning strategy, rather than adopting a fixed pipeline from other studies.

Third, although the three architectures differ in absolute RMSE, with LSTM-AE consistently achieving the lowest error due to its sequential modelling capacity, they produce concordant sensitivity rankings within each dataset. That is, model architecture determines the absolute level of noise tolerance, but the relative sensitivity pattern across noise types remains stable. This consistency across three structurally distinct architectures provides a degree of architectural generalisability.

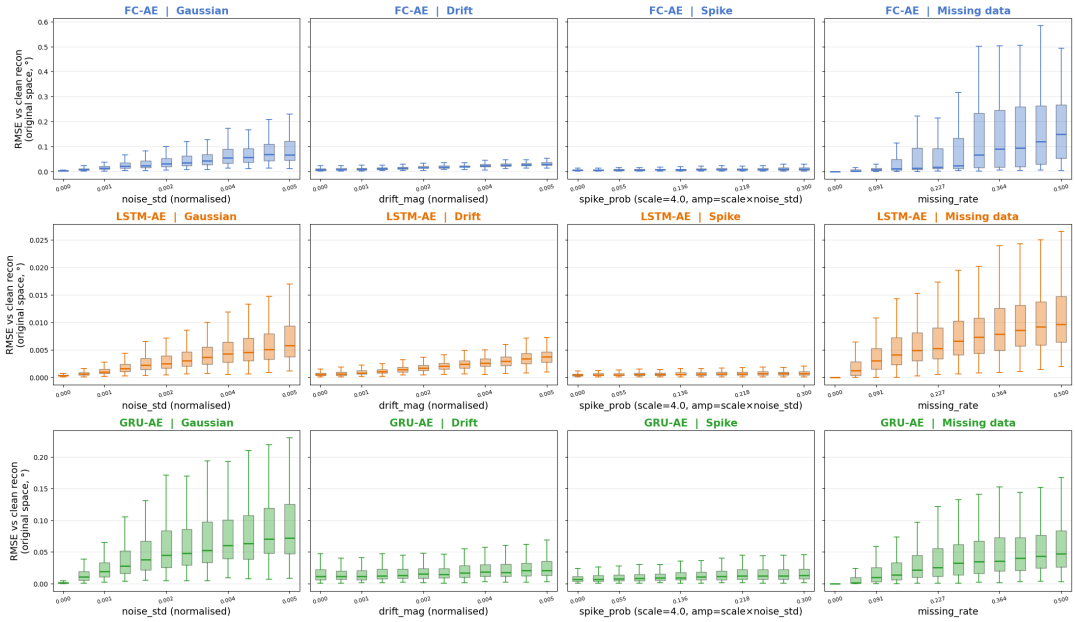
5. Conclusion

In this paper, we reviewed key application domains of ADS-B data, from trajectory modeling and prediction to methodology, simulation, and policy. We outlined the major cleaning procedures, including outlier removal, interpolation, resampling, and smoothing. Using an autoencoder-based case study, we quantitatively assessed how four noise types (Gaussian, drift, spike, and missing data) affect trajectory reconstruction across three architectures (FC-AE, LSTM-AE, GRU-AE) and four airport datasets (Zurich, Hangzhou, Guangzhou, Harbin). The results show that spike noise has the least impact on reconstruction across all model and dataset combinations, while drift noise causes the most consistent degradation even at small magnitudes. Missing data exhibits a nonlinear threshold effect, with reconstruction quality deteriorating rapidly beyond moderate loss rates. These findings indicate that, for AE-based reconstruction models, data cleaning should prioritise the correction of systematic spatial offsets and the management of high-rate data loss, rather than the exhaustive removal of local outliers. Furthermore, the relative sensitivity ranking among noise types varies across datasets but remains consistent across architectures, suggesting that cleaning strategies should be adapted to the spatial characteristics of the target airspace rather than applied as a fixed pipeline.

Several directions remain for future work. The evaluation could be extended beyond the autoencoder family to examine whether the observed noise sensitivity patterns generalise to other reconstruction paradigms such as transformer-based models or classical filtering methods. The autoencoder architectures themselves could also be refined with mechanisms that improve robustness to the noise types found most harmful in this study. Finally, complementing the current synthetic noise framework with validation against real-world ADS-B error profiles would further bridge the gap between controlled experiments and operational conditions.

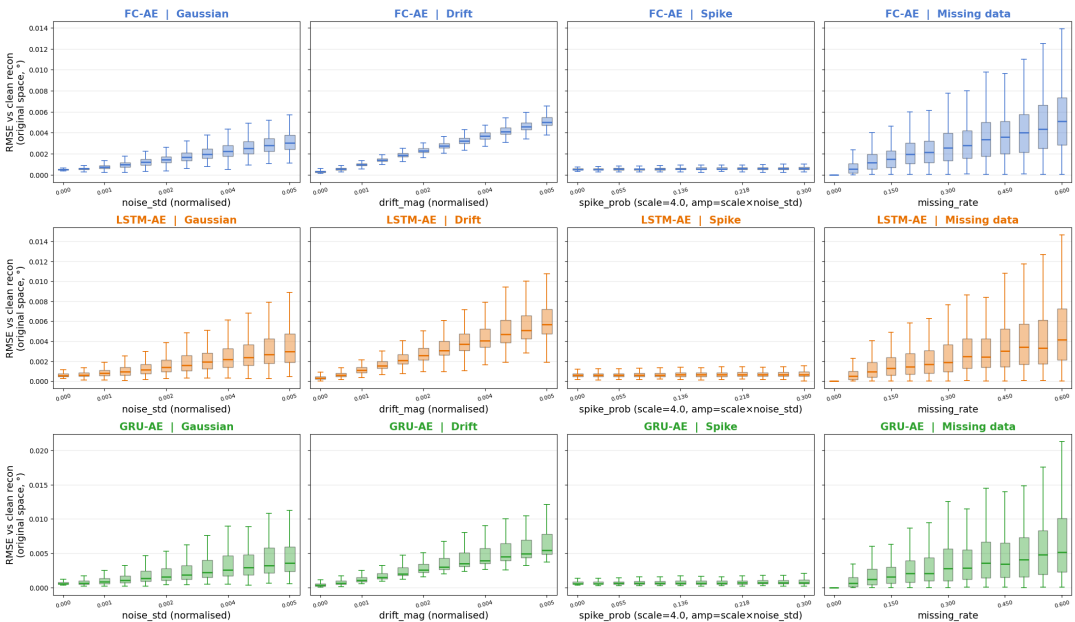
Appendix 1. Supplementary figures

Impact of Different Noise Types on Model Reconstruction — FC-AE & LSTM-AE & GRU-AE
 Mixed noise: base $\sigma=0.0005$, drift=0.0005, spike_prob=0.2 | spike amp = $4.0 \times \text{noise_std}$



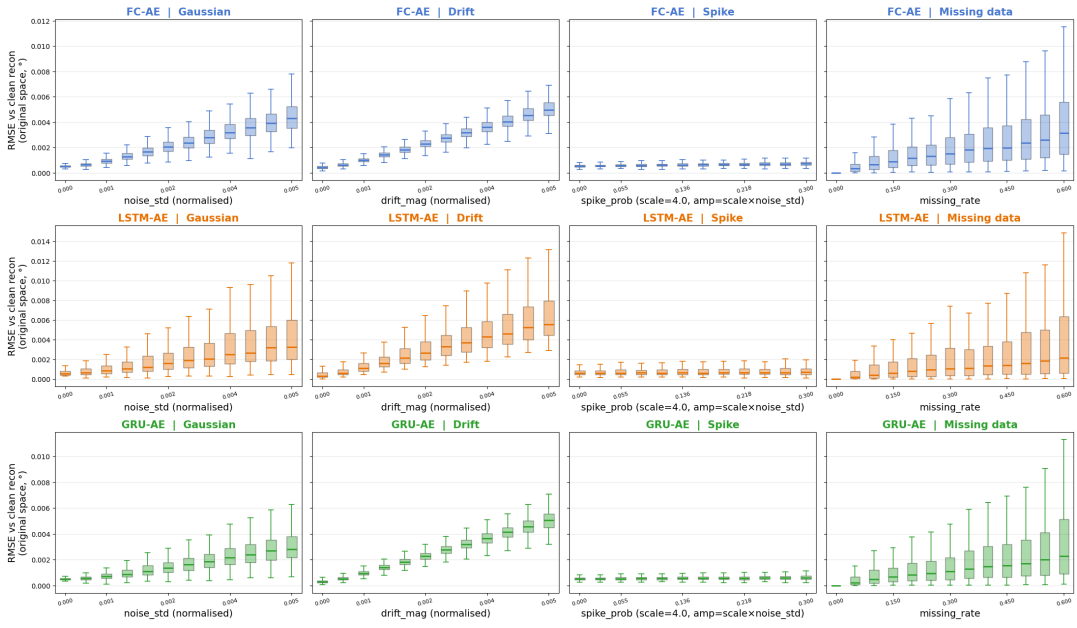
(a) Zurich (LSZH)

Impact of Different Noise Types on Model Reconstruction — FC-AE & LSTM-AE & GRU-AE
 Mixed noise: base $\sigma=0.0005$, drift=0.0005, spike_prob=0.2 | spike amp = $4.0 \times \text{noise_std}$



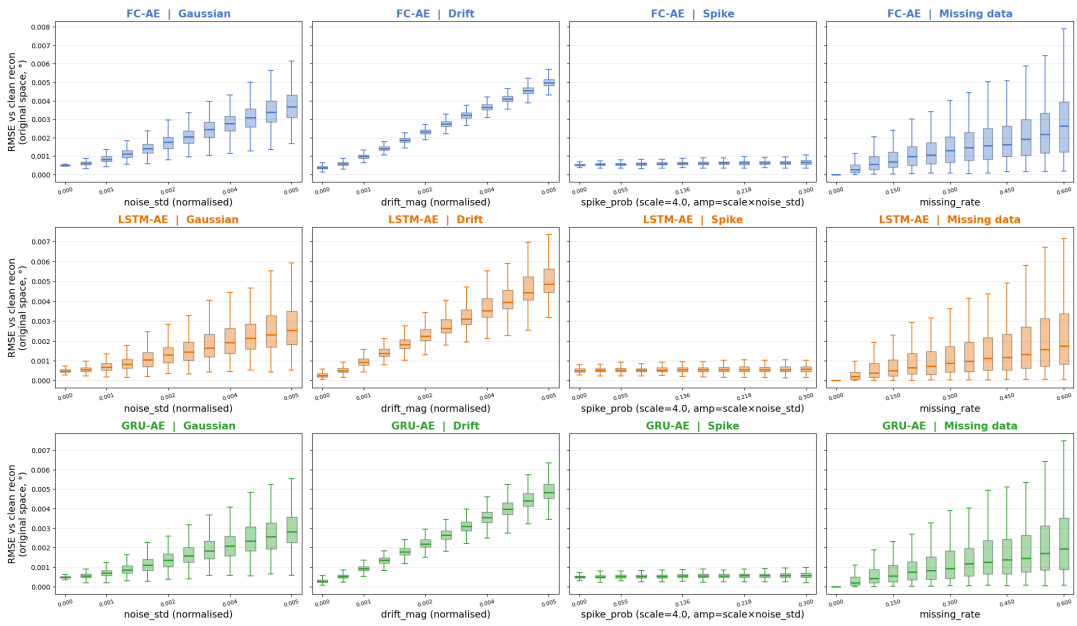
(b) Harbin (ZYHB)

Impact of Different Noise Types on Model Reconstruction — FC-AE & LSTM-AE & GRU-AE
 Mixed noise: base $\sigma=0.0005$, drift=0.0005, spike_prob=0.2 | spike amp = $4.0 \times \text{noise_std}$



(c) Hangzhou (ZSHC)

Impact of Different Noise Types on Model Reconstruction — FC-AE & LSTM-AE & GRU-AE
 Mixed noise: base $\sigma=0.0005$, drift=0.0005, spike_prob=0.2 | spike amp = $4.0 \times \text{noise_std}$



(d) Guangzhou (ZGGG)

Figure 9. RMSE vs. clean reconstruction across four noise types, three autoencoder architectures, and four datasets.

Appendix 2. Supplementary tables

Table 4. Pairwise Wilcoxon signed-rank tests on ΔRMSE for each dataset. $\Delta\text{RMSE} = \text{RMSE}(\text{max noise}) - \text{RMSE}(\text{no noise})$; G = Gaussian, D = Drift, S = Spike, M = Missing; “Tol.” = noise type with smaller ΔRMSE ; * $p < 0.05$.

(a) Zurich (LSZH)						(b) Harbin (ZYHB)					
Model	Pair	ΔA	ΔB	p	Tol.	Model	Pair	ΔA	ΔB	p	Tol.
FC	G–D	.0635	.0227	2.5e-66	* D	FC	G–D	.0025	.0047	1.7e-80	* G
	G–S	.0635	.0034	1.4e-82	* S		G–S	.0025	.0001	1.3e-83	* S
	G–M	.0635	.1488	4.3e-23	* G		G–M	.0025	.0051	6.2e-47	* G
	D–S	.0227	.0034	1.1e-70	* S		D–S	.0047	.0001	1.3e-83	* S
	D–M	.0227	.1488	2.7e-68	* D		D–M	.0047	.0051	1.0e-02	* D
	S–M	.0034	.1488	6.7e-82	* S		S–M	.0001	.0051	1.4e-83	* S
LSTM	G–D	.0054	.0031	5.5e-40	* D	LSTM	G–D	.0023	.0054	2.3e-54	* G
	G–S	.0054	.0003	1.5e-83	* S		G–S	.0023	.0001	1.0e-82	* S
	G–M	.0054	.0096	6.1e-34	* G		G–M	.0023	.0042	4.0e-22	* G
	D–S	.0031	.0003	4.4e-75	* S		D–S	.0054	.0001	1.3e-83	* S
	D–M	.0031	.0096	5.2e-75	* D		D–M	.0054	.0042	2.6e-08	* M
	S–M	.0003	.0096	1.8e-82	* S		S–M	.0001	.0042	3.4e-83	* S
GRU	G–D	.0704	.0089	1.6e-74	* D	GRU	G–D	.0029	.0051	1.7e-36	* G
	G–S	.0704	.0062	9.6e-79	* S		G–S	.0029	.0001	1.8e-80	* S
	G–M	.0704	.0477	7.1e-17	* M		G–M	.0029	.0051	1.4e-17	* G
	D–S	.0089	.0062	2.3e-02	* S		D–S	.0051	.0001	1.3e-83	* S
	D–M	.0089	.0477	1.0e-59	* D		D–M	.0051	.0051	4.3e-02	* D
	S–M	.0062	.0477	6.2e-71	* S		S–M	.0001	.0051	3.3e-83	* S
(c) Hangzhou (ZSHC)						(d) Guangzhou (ZGGG)					
Model	Pair	ΔA	ΔB	p	Tol.	Model	Pair	ΔA	ΔB	p	Tol.
FC	G–D	.0038	.0045	1.4e-22	* G	FC	G–D	.0032	.0046	5.7e-74	* G
	G–S	.0038	.0002	1.3e-83	* S		G–S	.0032	.0001	1.3e-83	* S
	G–M	.0038	.0031	2.4e-03	* M		G–M	.0032	.0026	5.6e-07	* M
	D–S	.0045	.0002	1.3e-83	* S		D–S	.0046	.0001	1.3e-83	* S
	D–M	.0045	.0031	1.1e-12	* M		D–M	.0046	.0026	6.6e-44	* M
	S–M	.0002	.0031	3.5e-83	* S		S–M	.0001	.0026	1.4e-83	* S
LSTM	G–D	.0027	.0051	6.1e-34	* G	LSTM	G–D	.0020	.0046	1.0e-73	* G
	G–S	.0027	.0001	1.3e-80	* S		G–S	.0020	.0001	1.4e-83	* S
	G–M	.0027	.0022	8.9e-02	—		G–M	.0020	.0017	2.0e-01	—
	D–S	.0051	.0001	3.9e-82	* S		D–S	.0046	.0001	1.3e-83	* S
	D–M	.0051	.0022	3.1e-21	* M		D–M	.0046	.0017	2.0e-44	* M
	S–M	.0001	.0022	1.3e-80	* S		S–M	.0001	.0017	2.4e-83	* S
GRU	G–D	.0023	.0047	2.5e-76	* G	GRU	G–D	.0023	.0045	4.4e-80	* G
	G–S	.0023	.0001	1.4e-83	* S		G–S	.0023	.0001	1.3e-83	* S
	G–M	.0023	.0023	6.3e-04	* M		G–M	.0023	.0019	1.5e-02	* M
	D–S	.0047	.0001	1.3e-83	* S		D–S	.0045	.0001	1.3e-83	* S
	D–M	.0047	.0023	1.9e-22	* M		D–M	.0045	.0019	7.0e-59	* M
	S–M	.0001	.0023	2.6e-83	* S		S–M	.0001	.0019	2.1e-83	* S

Table 5. Mapping of normalised Gaussian noise parameters to real-world distances for each dataset.

Airport	max_range (°)	Lat extent	Lon extent	$\sigma = 0.003$	$\sigma = 0.005$
Hangzhou	2.0812°	232 km	200 km	601–695 m	1001–1158 m
Guangzhou	1.9597°	218 km	200 km	601–654 m	1001–1091 m
Harbin	2.5628°	285 km	200 km	599–856 m	998–1426 m
Zurich	1.9574°	218 km	147 km	442–654 m	737–1089 m

Author contributions

All authors contributed to the overall content of the paper. Ruolan Ren, Jingcheng Zhong, and Dizhi Guo performed the literature review and writing. Christophe Hurter provided the main idea, coding, and supervision. Ruixin Wang contributed to reviewing and editing the paper.

Funding statement

This research was funded by the National Natural Science Foundation of China (No. 72301278) and the Tianjin Key Research and Development Program, China (No. 25YFXTZH00070).

Open data statement

The Zurich (LSZH) ADS-B dataset is obtained from the OpenSky Network [11].

Reproducibility statement

The implementation code for the experiments, figures, and tables presented in this paper is openly available at <https://github.com/Rowan4399/Use-case-investigation-of-the-noise-impact-on-Auto-Encoder-algorithm>.

References

- [1] Yuejingyan Wang, Liang Zhao, Yuyang Jia, and Kaiquan Cai. “A Performance Learning Method for Aircraft Trajectory Modeling”. In: *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*. IEEE, 2021, pp. 1–7.
- [2] Samantha J. Corrado, Tejas G. Puranik, Olivia Pinon Fischer, and Dimitri N. Mavris. “A Clustering-Based Quantitative Analysis of the Interdependent Relationship Between Spatial and Energy Anomalies in ADS-B Trajectory Data”. In: *Transportation Research Part C: Emerging Technologies* 131 (2021), p. 103331.
- [3] Junzi Sun, Joost Ellerbroek, and Jacco Hoekstra. “Large-scale flight phase identification from ads-b data using machine learning methods”. In: *7th International Conference on Research in Air Transportation*. 2016, pp. 1–8.
- [4] Chunyao Ma, Sameer Alam, Qing Cai, and Daniel Delahaye. “A Machine Learned Traffic Flow Coordination Framework for Flow-Centric Airspace”. In: *SESAR Innovations Days*. SESAR. Séville (SP), Spain, Nov. 2023. URL: <https://enac.hal.science/hal-04314653>.
- [5] Fei Lu, Zichen Chen, and Huiyu Chen. “Lateral collision risk assessment of parallel routes in ocean area based on space-based ADS-B”. In: *Transportation Research Part C: Emerging Technologies* 124 (2021), p. 102970. ISSN: 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2021.102970>. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X21000085>.

- [6] Andrew M Churchill and Michael Bloem. “Clustering aircraft trajectories on the airport surface”. In: *Proceedings of the 13th USA/Europe Air Traffic Management Research and Development Seminar, Chicago, IL, USA*. 2019, pp. 10–13.
- [7] International Civil Aviation Organization. *Destruction of Korean Air Lines Boeing 747 on 31 August 1983 (Report)*. Ref. LE 4/19.4 - 93/68. Montreal, 1993.
- [8] Junzi Sun. *The 1090 Megahertz Riddle: A Guide to Decoding Mode S and ADS-B Signals*. 2nd ed. TU Delft OPEN Publishing, 2021. ISBN: 978-94-6366-402-8. DOI: 10.34641/mg.11.
- [9] International Civil Aviation Organization. “Eleventh Air Navigation Conference”. In: *Eleventh Air Navigation Conference*. AN-Conf/11-WP/190. Montreal: ICAO, 2003.
- [10] Victor Monzonis Melero, Juan V Balbastre, and Alex Ganau. “SATERA Baseline: Identifying the Challenges of Space-Based Multilateration Systems”. In: *2024 SESAR Innovation days* (2024).
- [11] Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, and Matthias Wilhelm. “Bringing up OpenSky: A large-scale ADS-B sensor network for research”. In: *IPSN-14 proceedings of the 13th international symposium on information processing in sensor networks*. IEEE. 2014, pp. 83–94.
- [12] Gui Xuhao, Zhang Junfeng, and Peng Zihan. “Trajectory clustering for arrival aircraft via new trajectory representation”. In: *Journal of Systems Engineering and Electronics* 32.2 (2021), pp. 473–486.
- [13] Zhengyi Wang, Man Liang, and Daniel Delahaye. “Short-term 4D Trajectory Prediction Using Machine Learning Methods”. In: *7th SESAR Innovation Days (SID 2017)*. 2017.
- [14] Zhengyi Wang, Man Liang, and Daniel Delahaye. “A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area”. In: *Transportation Research Part C: Emerging Technologies* 95 (2018), pp. 280–294.
- [15] Luis Basora, Valentin Courchelle, Judicaël Bedouet, and Thomas Dubot. “Occupancy peak estimation from sector geometry and traffic flow data”. In: *Proceedings of the 8th SESAR Innovation Days, Salzburg, Austria* (2018), pp. 3–7.
- [16] Daniel Delahaye, Chunyao Ma, Sameer Alam, and Qing Cai. “Air traffic flow representation and prediction using transformer in flow-centric airspace”. In: *SESAR Innovation Days*. 2022.
- [17] Alberto Bonifazi, Junzi Sun, Gerben van Baren, and Jacco Hoekstra. “Modeling and Detecting Anomalous Safety Events in Approach Flights Using ADS-B Data”. In: *Proceedings of the Fourteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2021), Virtual Event*. 2021, pp. 20–23.
- [18] R Conrad Rorie and Casey L Smith. “Detect and Avoid and Collision Avoidance Flight Test Results with ACAS Xr”. In: *2024 AIAA DATC/IEEE 43rd Digital Avionics Systems Conference (DASC)*. IEEE. 2024, pp. 01–10.
- [19] Yi Zhang, Sheng Zhang, Yicheng Zhang, and Yifang Yin. “A study of TMA aircraft conflict-free routing and operation: With mixed integer linear programming, multi-agent path finding, and metaheuristic-based neighborhood search”. In: *IEEE Transactions on Intelligent Transportation Systems* 25.10 (2024), pp. 13976–13990.
- [20] Jie Bao, Zijie Zhang, Junfeng Zhang, Yixuan Chen, and Xuhao Gui. “Exploring the conflict risk characteristics of air weaving sections in Metroplex terminal areas with flight trajectory data and adaptive graph spatial-temporal transformer”. In: *Journal of Air Transport Management* 120 (2024), p. 102667.
- [21] Junzi Sun, Jacco M Hoekstra, and Joost Ellerbroek. “Aircraft drag polar estimation based on a stochastic hierarchical model”. In: *Eighth SESAR innovation days* (2018), pp. 1–8.
- [22] Chenyu Huang and Xiaoyue Cheng. “Estimation of aircraft fuel consumption by modeling flight data from avionics systems”. In: *Journal of Air Transport Management* 99 (2022), p. 102181.
- [23] Richard Alligier. “Predictive distribution of mass and speed profile to improve aircraft climb prediction”. In: *Journal of Air Transportation* 28.3 (2020), pp. 114–123.

- [24] Asma Tabassum, Nicholas Allen, and William Semke. “ADS-B message contents evaluation and breakdown of anomalies”. In: *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*. IEEE. 2017, pp. 1–8.
- [25] Sebastian Wandelt, Xiaoqian Sun, and Hartmut Fricke. “Ads-bi: Compressed indexing of ads-b data”. In: *IEEE Transactions on Intelligent Transportation Systems* 19.12 (2018), pp. 3795–3806.
- [26] Enrico Spinielli, Rainer Koelle, Massimiliano Zanin, and Seddik Belkoura. “Initial implementation of reference trajectories for performance review”. In: *Proceedings of the 7th SESAR innovation days, Belgrade (Serbia)* (2017).
- [27] Esther Roosenbrand, Junzi Sun, and Jacco Hoekstra. “Contrail Altitude Estimation Based on Shadows Detected in Landsat Imagery”. In: *Proceedings of the 13th SESAR Innovation Days, Sevilla, Spain*. 2023.
- [28] Junzi Sun, Aidana Tassanbi, Piotrek Obojski, and Philip Plantholt. “Evaluating Transatlantic Flight Emissions and Inefficiencies Using Space-Based ADS-B Data”. In: *Proceedings of the 13th SESAR Innovation Days, Sevilla, Spain*. 2023, pp. 27–30.
- [29] Aymeric Cretin, Bruno Legeard, Fabien Peureux, and Alexandre Vernotte. “Increasing the resilience of ATC systems against false data injection attacks using DSL-based testing”. In: *International Conference on Research in Air Transportation*. 2018.
- [30] Suleman Khan, Joakim Thorn, Alex Wahlgren, and Andrei Gurtov. “Intrusion detection in automatic dependent surveillance-broadcast (ADS-B) with machine learning”. In: *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*. IEEE. 2021, pp. 1–10.
- [31] Peter Mehltitz, Dimitra Giannakopoulou, and Nastaran Shafiei. “Analyzing airspace data with race”. In: *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*. IEEE. 2019, pp. 1–10.
- [32] Tatjana Bolic, Andrew Cook, Rainer Koelle, Enrico Spinielli, Quinten Goens, and Martin Strohmeier. “Roadmap for a European open science alliance for ATM research”. In: *European Journal of Transport and Infrastructure Research* 24.4 (2024), pp. 18–40.
- [33] Matthias Schäfer, Martin Strohmeier, Matthew Smith, Markus Fuchs, Vincent Lenders, and Ivan Martinovic. “OpenSky Report 2018: Assessing the Integrity of Crowdsourced Mode S and ADS-B Data”. In: *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. 2018, pp. 1–9. doi: 10.1109/DASC.2018.8569833.
- [34] Xavier Olive, Jan Krummen, Benoit Figuet, and Richard Alligier. “Filtering Techniques for ADS-b Trajectory Preprocessing”. In: *Journal of Open Aviation Science* 2.2 (Mar. 2025). doi: 10.59490/joas.2024.7882.
- [35] Imen Dhief, Sameer Alam, Chan Chea Mean, and Nimrod Lilith. “A Tree-based Machine Learning Model for Go-around Detection and Prediction”. In: *Proceedings 11th SESAR Innovation Days* (2021).
- [36] Yingli Liu, Minghua Hu, Jianan Yin, Jiaming Su, and Peiran Qiao. “Adaptive airport taxiing rule management: Design, assessment, and configuration”. In: *Transportation Research Part C: Emerging Technologies* 163 (2024), p. 104652. ISSN: 0968-090X. doi: <https://doi.org/10.1016/j.trc.2024.104652>. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X24001736>.
- [37] Martin Lindner, Thomas Zeh, Hannes Braßel, and Hartmut Fricke. “Aircraft performance-optimized departure flights using traffic flow funnels”. In: *Proceedings of the Fourteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2021), Virtual Event*. 2021, pp. 20–23.
- [38] Junzi Sun, Joost Ellerbroek, and Jacco M Hoekstra. “WRAP: An open-source kinematic aircraft performance model”. In: *Transportation Research Part C: Emerging Technologies* 98 (2019), pp. 118–138.
- [39] Hesam Shafienya and Amelia C Regan. “4D flight trajectory prediction using a hybrid Deep Learning prediction method based on ADS-B technology: A case study of Hartsfield–Jackson

- Atlanta International Airport (ATL)". In: *Transportation Research Part C: Emerging Technologies* 144 (2022), p. 103878.
- [40] Seokbin Yoon and Keumjin Lee. "Improving Aircraft Trajectory Prediction Accuracy with Over-sampling Technique". In: *2023 IEEE/AIAA 42nd Digital Avionics Systems Conference (DASC)*. 2023, pp. 1–6. doi: 10.1109/DASC58513.2023.10311324.
- [41] Zhengyi Wang, Man Liang, and Daniel Delahaye. "Automated data-driven prediction on aircraft Estimated Time of Arrival". In: *Journal of Air Transport Management* 88 (2020), p. 101840. issn: 0969-6997. doi: <https://doi.org/10.1016/j.jairtraman.2020.101840>. url: <https://www.sciencedirect.com/science/article/pii/S0969699719304429>.
- [42] RW Vos, Junzi Sun, and JM Hoekstra. "A Transformer-based Trajectory Prediction Model to Support Air Traffic Demand Forecasting". In: (2024).
- [43] Michael Schultz, Judith Rosenow, and Xavier Olive. "Data-driven airport management enabled by operational milestones derived from ADS-B messages". In: *Journal of Air Transport Management* 99 (2022), p. 102164.
- [44] Xavier Olive, Jeremy Grignard, Thomas Dubot, and Julie Saint-Lot. "Detecting controllers' actions in past mode S data by autoencoder-based anomaly detection". In: *SID 2018, 8th SESAR Innovation Days*. 2018.
- [45] Xinting Zhu, Ning Hong, Fang He, Yu Lin, Lishuai Li, and Xiaowen Fu. "Predicting aircraft trajectory uncertainties for terminal airspace design evaluation". In: *Journal of Air Transport Management* 113 (2023), p. 102473. issn: 0969-6997. doi: <https://doi.org/10.1016/j.jairtraman.2023.102473>. url: <https://www.sciencedirect.com/science/article/pii/S0969699723001163>.
- [46] Zouhair Mahboubi and Mykel J. Kochenderfer. "Learning Traffic Patterns at Small Airports From Flight Tracks". In: *IEEE Transactions on Intelligent Transportation Systems* 18.4 (2017), pp. 917–926. doi: 10.1109/TITS.2016.2598064.
- [47] Richard Alligier and David Gianazza. "Learning aircraft operational factors to improve aircraft climb prediction: A large scale multi-airport study". In: *Transportation research part C: emerging technologies* 96 (2018), pp. 72–95.
- [48] Xavier Olive. "a toolbox for processing and analysing air traffic data". In: *Journal of Open Source Software* 4.39 (2019), pp. 1518–1.
- [49] Xavier Olive, Luis Basora, Benoit Viry, and Richard Alligier. "Deep Trajectory Clustering with Autoencoders". In: *ICRAT 2020, 9th International Conference for Research in Air Transportation*. ICRAT 2020. Tampa, United States, Sept. 2020. url: <https://enac.hal.science/hal-02916241>.
- [50] Timothé Krauth, Adrien Lafage, Jérôme Morio, Xavier Olive, and Manuel Waltert. "Deep Generative Modelling of Aircraft Trajectories in Terminal Maneuvering Areas". In: *Machine Learning with Applications* 11 (2023), p. 100446. issn: 2666-8270.
- [51] NATS. *General Aviation: ADS-B / GPS Trial Results*. Tech. rep. Version 2.1. National Air Traffic Services, Dec. 2015.