

Article

Air Traffic Noise Prediction Method Based on Machine Learning Driven by Quick Access Recorder

Zhixing Tang ^{1,*}, Yijie Fan ^{1,*}, Xuanting Chen ¹, Xinyan Shi ¹, Zhaolun Niu ¹, Yuming Zhong ², Meng Jia ³ and Xiaowei Tang ⁴

¹ College of Air Traffic Management, Civil Aviation Flight University of China, Guanghan 618307, China; c1697019649@163.com (X.C.); xinyanshi077@gmail.com (X.S.); nzl0306@163.com (Z.N.)

² China Academy of Civil Aviation Science and Technology, Beijing 100028, China; zhongym@mail.castc.org.cn

³ School of Transportation Engineering, Nanjing Institute of Technology, Nanjing 211167, China; jiameng@njit.edu.cn

⁴ Key Laboratory of Civil Aviation Emergency Science and Technology, Nanjing 211106, China; tangxiaowei@nuaa.edu.cn

* Correspondence: tzx@cafuc.edu.cn (Z.T.); zfyfyj@163.com (Y.F.)

Abstract

Accurate prediction of air traffic noise is critical for advancing environmentally sustainable operations in high density terminal areas. Conventional noise prediction models often exhibit significant limitations due to discrepancies between actual and nominal flight trajectories. To overcome this challenge, this study introduces a probabilistic framework that integrates real air-traffic-flow data to generate realistic flight trajectory distributions. The proposed methodology extracts key operational features—including trajectory distribution probabilities, and essential trajectory operation features—within a machine learning architecture. Furthermore, we develop a dedicated air traffic noise prediction model for clustered flight paths that explicitly incorporates traffic flow patterns, enabling high-fidelity simulation of noise propagation under actual air traffic operation. The framework is validated using a QAR (Quick Access Recorder) dataset from the terminal area of Changsha Huanghua International Airport. Experimental results demonstrate the model's high predictive accuracy for both air traffic noise distribution and its influence, coupled with computational efficiency and practical applicability. The findings indicate that the proposed approach successfully addresses the challenge of predicting air traffic noise from divergent, real-world flight trajectories, offering a robust method for supporting noise-abatement strategies and sustainable aviation-planning initiatives.

Keywords: air traffic noise prediction; probabilistic traffic trajectory; trajectory operation features; machine learning; QAR

Academic Editors: Michael Schultz, Pham Duc Thinh and Alexei Sharpanskykh

Received: 30 December 2025

Revised: 10 February 2026

Accepted: 21 February 2026

Published: 24 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the rapid growth of the aviation industry, aircraft noise around airports has become an increasingly pressing environmental and public-health concern as air transport demand continues to grow. Basner et al. reviewed the state of the science and highlighted that aircraft noise is linked to a broad set of outcomes, which makes quantitative exposure modelling and transparent assumptions central to credible assessments [1]. The World Health Organization has consolidated health evidence and issued policy-

oriented recommendations for transportation noise, further underscoring the need for defensible population-exposure estimation rather than qualitative descriptions [2]. Mechanistic and epidemiological syntheses also indicate that environmental noise can contribute to cardiovascular risk through stress-related pathways, reinforcing that airport-noise modelling should be treated as a health-relevant engineering task rather than a purely acoustic exercise [3].

Importantly, airport noise impact is not determined solely by an “average” sound level. Event-based characteristics and operational variability matter: Janssen et al. showed that the number of aircraft noise events can affect sleep quality, implying that prediction of air traffic noise should explicitly reflect traffic-flow patterns rather than relying on a single mean-track approximation [4]. Saucy et al. further reported associations between night-time aircraft noise and cardiovascular mortality, motivating robust prediction for night operations and cumulative exposure rather than isolated events alone [5]. At the regulatory level, the Environmental Noise Directive establishes a framework for environmental noise assessment and management in Europe, illustrating that standardized and defensible prediction and evaluation is a practical necessity for airports and authorities [6].

From an airport-management perspective, the terminal area is particularly sensitive because community exposure is driven not only by aircraft type and distance but also by air-traffic flow and operational variability. ICAO (International Civil Aviation Organization)’s Balanced Approach explicitly treats operational procedures and their evaluation as one pillar of noise management, which makes credible prediction tools essential for screening and assessing mitigation options [7]. ICAO also provides recommended methods for computing noise contours around airports, offering a standardized basis for contour computation while implicitly requiring that the input representation of procedures and operations be realistic for the local study context [8]. Recent European environmental reporting further emphasizes the continuing relevance of airport noise for sustainability and community impact and motivates methodological improvements that can be applied at scale [9]. In practice, however, published procedure specifications describe nominal intent rather than a statistical representation of how aircraft actually fly, and real trajectories often deviate due to ATC instructions, weather, airline operating policies, and performance management such as thrust settings [10]. As a result, air traffic noise prediction that is driven by traffic flow can be biased if it is built on oversimplified tracks or fixed operational profiles.

Standardized noise methodologies and tools provide consistent computation, but their predictive credibility depends on the operational inputs supplied. AEDT offers an integrated workflow for environmental assessment and supports multiple modelling modes, yet its reliability in a case study remains bounded by the realism of the procedure definition and performance-related inputs [11]. ECAC (European Civil Aviation Conference) Doc.29 provides a widely used standard for computing noise contours around civil airports, but its outputs are similarly sensitive to how track geometry and operational states are represented [12]. Clemente et al. quantified that contour predictions can be highly sensitive to errors in key inputs, arguing that uncertainty in operational inputs can be amplified in output noise levels if not controlled [13]. Zaporozhets and Levchenko further showed that definitions related to noise–power–distance relationships can materially influence single-event calculations, reinforcing that thrust/power representation is a dominant sensitivity for departure noise and cannot be treated as an afterthought [14]. More recently, van der Grift et al. examined how using engine fan speed (N1%) as a proxy input affects noise modelling, again highlighting that converting operational signals into thrust/power inputs is a nontrivial source of modelling uncertainty [15]. These findings collectively indicate that improving air traffic noise prediction requires (i) a statistically

meaningful representation of real traffic trajectories and (ii) a reliable way to provide segment-wise operational states along those trajectories.

To reflect traffic-flow variability and operational realism in terminal-area noise prediction, a growing body of research has explored data-driven representations of real operations. Brooks and Hansman explicitly modelled the effects of flight-track variability on community noise exposure, demonstrating that variability can meaningfully change exposure outcomes and therefore should be incorporated rather than averaged away [16]. Pretto et al. leveraged high-resolution tracking data to reconstruct operational states for airport noise prediction, illustrating the feasibility of data-driven operation reconstruction while highlighting the need to manage parameter uncertainty when reconstructing profiles from surveillance data alone [17]. Crucially, large-scale validation work has begun to quantify how much fidelity is gained when realistic trajectories are used: Rindfleisch et al. conducted a large-scale validation study by pairing a very large set of flight trajectories with measured sound levels for airport arrivals, illustrating both the potential and the practical challenges of validating model outputs against monitoring data at scale [18].

A key enabling step for traffic-flow-based noise prediction is to reduce large volumes of heterogeneous real trajectories into a compact, statistically meaningful representation. Gariel et al. introduced trajectory clustering for airspace monitoring, providing a foundational framework for extracting representative tracks from surveillance data [19]. Basora and Morio proposed a clustering framework to analyse air-traffic flows, supporting the idea of probabilistic representative trajectories but not directly addressing the parameter-completeness requirements of standardized noise engines [20]. Eerland and Box demonstrated a two-step process that clusters trajectories to extract common trends and then models dispersion probabilistically, and they further showed how representative trajectories can replace large historical datasets while preserving footprint properties [21]. Building on the availability of large-scale ADS-B data, Bhanpato et al. combined trajectory comparison and clustering to identify representative departure procedures and quantify the community-noise impact differences between real operations and defined noise-abatement procedures, directly motivating data-driven traffic-flow representations for noise assessment [22].

In China, related studies have begun to connect procedure analysis with noise impact prediction. Chen proposed a procedure-based noise assessment approach using trajectory clustering and probability assignment; however, simplified spatial assumptions around a center trajectory can limit the representation of complex dispersion patterns, particularly in three dimensions [23]. Hang demonstrated how QAR (Quick Access Recorder) data can support procedure design quality evaluation and mitigation assessment, validating the value of recorder-grade data, but the coupling between clustered trajectories and standardized contour computation still requires a complete and consistent set of operational parameters along representative trajectories [24]. Zhao and Zhang proposed a rapid AEDT (Aviation Environmental Design Tool)-based method for laterally dispersed tracks by predicting noise on a center track and assigning probabilities for deviations; this improves efficiency for traffic-flow-based estimation but remains limited in handling vertical divergence and profile variability [25]. Zhuang et al. integrated departure track simulation and noise influence analysis using a performance model (BADA) and external validation, confirming that noise is strongly influenced by thrust, altitude, and turning strategy – while also implying that inaccurate reconstruction of these parameters can dominate modelling error [26].

While previous studies have provided important insights, a unified framework that predicts air traffic noise from traffic-flow statistics still faces two practical barriers: representing the probabilistic distribution of real 3D trajectories and providing complete, segment-wise operational states required by standardized noise engines. To address these

barriers, this study develops an air traffic noise prediction framework with the following contributions:

(a) Probabilistic traffic trajectories: real QAR trajectories are condensed into representative 3D trajectories and their occurrence probabilities, enabling traffic-flow statistics to be translated into a compact, statistically meaningful trajectory distribution for prediction.

(b) Flight-parameter completion on representative trajectories: key segment-wise operational variables required by standardized noise computation are predicted via a machine-learning module trained on QAR observations, mitigating reliance on nominal profiles and fixed assumptions.

(c) An operational-knowledge-driven prediction pipeline: probabilistic representative trajectories and predicted operational states are integrated with a standardized noise computation engine to deliver computationally efficient air traffic noise prediction suitable for terminal-area scenario screening and operational evaluation.

2. Methodologies

The primary research methodology of this study is divided into two main components: the construction of a trajectory operation knowledge base and noise calculation based on this knowledge base. QAR data is used as the foundational dataset. First, actual flight trajectories are classified, and the resulting trajectory clusters are further grouped through clustering to serve as the basis for noise calculation. At the same time, the distribution probability of each clustered trajectory is determined. Since only actual flight trajectories contain complete flight status parameters required for noise calculation, some parameters in the clustered trajectories remain unknown. Therefore, it is necessary to predict these missing parameters to complete the data required for noise computation. By integrating the distribution probabilities and machine-learned flight parameters of clustered trajectories, the proposed model enhances the accuracy of traditional noise prediction approaches.

Trajectory operation knowledge mining serves as a fundamental basis for terminal area noise analysis, with its core tasks including the extraction of trajectory distribution probabilities and the prediction of flight status parameters. Together, these probabilities and predicted parameters constitute the trajectory operation knowledge system. This study proposes a method for characterizing traffic flow impact based on probabilistic clustered trajectories and utilizes machine learning to accurately predict the flight status parameters of these clustered trajectories as shown in Figure 1. This research approach not only effectively captures the dynamic characteristics of terminal area traffic flow but also lays a solid theoretical foundation for the development of noise prediction models.

2.1. 3D Trajectory Distribution Modeling via Fast-DTW

In the identification of airport aircraft traffic flows, the arrival and departure directions of flight trajectories are considered important input features for the construction of subsequent noise prediction models. To this end, the navigation direction derived from the trajectory heading angle is used as the primary criterion for classification. On this basis, an improved Dynamic Time Warping (DTW) algorithm is introduced to enhance the accuracy of matching between actual flight trajectories and standard flight procedures, thereby achieving precise alignment. The QAR can record a wide range of flight parameters (such as wind speed, latitude and longitude, attitude, thrust, etc.) for all arrival and departure trajectories within the terminal area. Although QAR data has been widely used for fault prediction [27,28], its potential for noise calculation remains underexplored. With its high sampling rate and high accuracy, QAR data is particularly well-suited for trajectory clustering and flight status parameter prediction in aircraft noise modeling.

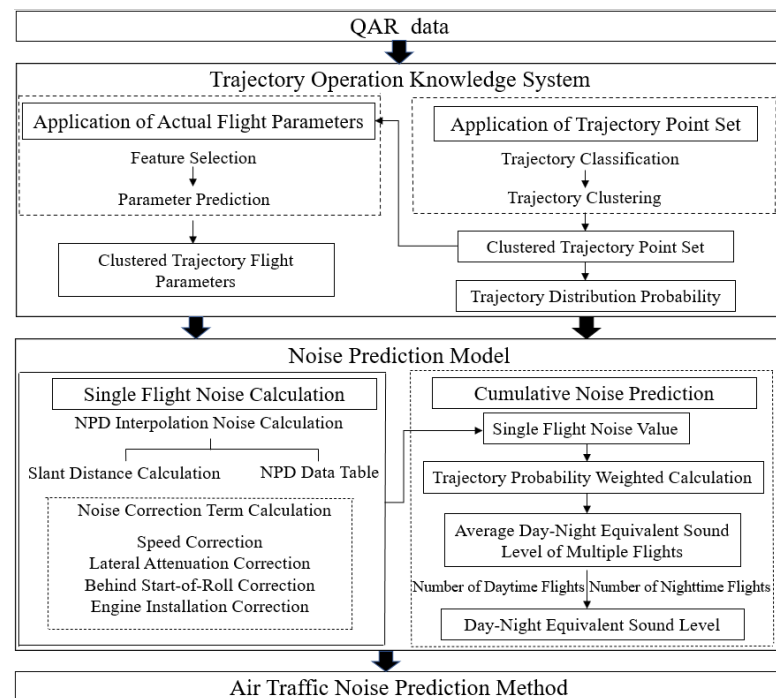


Figure 1. Research methodology roadmap.

To extract representative operational trajectories and their three-dimensional distribution probabilities from QAR data, a two-stage procedure is adopted: (i) procedure-aided corridor-entry classification and trajectory screening, followed by (ii) within-corridor 3D trajectory clustering. Each flight record contains a time-ordered sequence of discrete trajectory points. The Dynamic Time Warping (DTW) algorithm, which employs a dynamic programming approach to align two temporal sequences by minimizing their temporal discrepancies, is capable of estimating the minimal possible distance (i.e., maximal potential similarity) between them. This characteristic makes DTW particularly well-suited for clustering flight trajectories. Dynamic Time Warping (DTW) is used as the trajectory similarity metric [29], and FastDTW is adopted to reduce computational cost while retaining DTW accuracy [30]. For a given flight f , its raw trajectory is denoted by:

$$T^{(f)} = \{(lat_i, lon_i, alt_i)\}_{i=1}^{N_f}, \quad (1)$$

where lat_i , lon_i , and alt_i are the latitude, longitude, and QNH altitude at the i -th recorded point, respectively and N_f denotes the number of recorded trajectory points for flight f . Since departures and arrivals are recorded in opposite temporal directions with respect to the runway reference, departure trajectories are reversed so that both departures and arrivals are represented consistently from the runway vicinity toward the terminal-area boundary. In addition, runway usage is identified from QAR fields, and the corresponding nominal procedure set (SID/STAR) for the detected runway is loaded as a trajectory database:

$$p = \{P^{(m)}\}_{m=1}^M, P^{(m)} = \{(lat_j, lon_j)\}_{j=1}^{L_m}, \quad (2)$$

where each $P^{(m)}$ is a polyline of a published SID/STAR in planform (latitude–longitude), $m \in \{1, \dots, M\}$ indexes the published procedures, M is the number of available procedures under the selected runway and flight status (departure/arrival), j is the point index along a nominal procedure polyline, and L_m is the number of planform points of procedure m . For departures, the nominal polylines are also reversed to match the same direction as the reversed flight trajectories.

2.1.1. Trajectory Classification Based on Direction

A coarse but efficient corridor-entry pre-classification is performed to reduce mismatching among procedures with similar shapes. For each nominal procedure $P^{(m)}$, a procedure direction angle $\theta^{(m)}$ is computed from its first and last planform points:

$$\theta^{(m)} = \text{atan2}(\text{lat}_1^{(m)} - \text{lat}_{L_m}^{(m)}, \text{lon}_1^{(m)} - \text{lon}_{L_m}^{(m)}), \quad (3)$$

where $\text{lat}_1^{(m)}$, $\text{lon}_1^{(m)}$ and $\text{lat}_{L_m}^{(m)}$, $\text{lon}_{L_m}^{(m)}$ denote the latitude/longitude of the first and last points of procedure m , respectively, and atan2 is the two-argument arctangent. Similarly, the flight direction angle $\theta^{(f)}$ is computed from the first and last points of the flight trajectory:

$$\theta^{(f)} = \text{atan2}(\text{lat}_1^{(f)} - \text{lat}_{N_f}^{(f)}, \text{lon}_1^{(f)} - \text{lon}_{N_f}^{(f)}), \quad (4)$$

where $\text{lat}_1^{(f)}$, $\text{lon}_1^{(f)}$ and $\text{lat}_{N_f}^{(f)}$, $\text{lon}_{N_f}^{(f)}$ are the latitude/longitude of the first and last trajectory points of flight f , respectively. Nominal procedures are labeled by corridor-entry identifiers; let c denote a corridor-entry class label and $\theta^{(c)}$ denote the representative direction angle associated with class c . For each flight, the selected corridor-entry class c_{sel} is determined by minimizing the absolute heading difference:

$$c_{\text{sel}} = \arg \min_c |\theta^{(f)} - \theta^{(c)}|, \quad (5)$$

where $\arg \min$ returns the class label that yields the smallest value. This step constrains subsequent matching and clustering to physically consistent traffic-flow directions.

Accordingly, Figure 2 provides a schematic overview of the role of this section. Specifically, flight tracks are distinguished and categorized by evaluating the deviation in heading angle relative to different nominal procedures, thereby enabling a preliminary classification of daily real-world trajectories according to the corridor-entry direction.

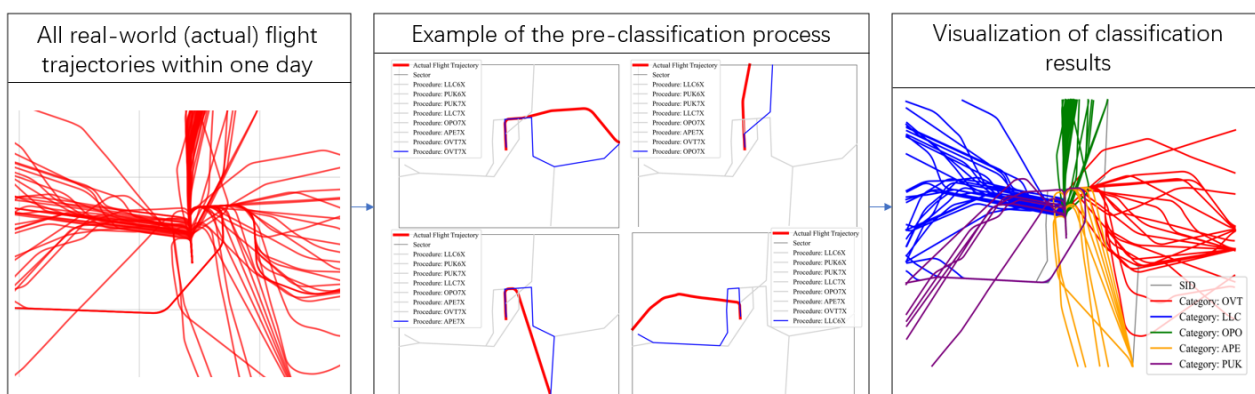


Figure 2. Trajectory classification schematic.

2.1.2. Procedure Matching and Similarity Scoring

Within the selected corridor-entry class c_{sel} , the flight trajectory is matched against each nominal procedure $P^{(m)} \in p_{c_{\text{sel}}}$, where $p_{c_{\text{sel}}} \subseteq p$ is the subset of procedures belonging to corridor class c_{sel} . Matching is performed using the Fast Dynamic Time Warping (Fast-DTW) algorithm. Let $\mathbf{X} = \{x_i\}_{i=1}^n$ and $\mathbf{Y} = \{y_j\}_{j=1}^\ell$ be two point sequences, where n and ℓ denote their lengths. In our procedure matching stage, $x_i = (\text{lat}_i, \text{lon}_i)$ and $y_j = (\text{lat}'_j, \text{lon}'_j)$ are 2D planform points (latitude–longitude). The Fast-DTW distance is defined as [31]:

$$D_{\text{FDTW}}(\mathbf{X}, \mathbf{Y}) = \min_w \sum_{(i,j) \in w} d(x_i, y_j), \quad (6)$$

where $D_{\text{FDTW}}(\mathbf{X}, \mathbf{Y})$ is the Fast-DTW distance, $\mathcal{W}$ is a valid warping path, and $d(x_i, y_j)$ is the point-wise Euclidean distance between the two 2D points.

For each flight f , the Fast-DTW distances $\{D_{\text{FDTW}}(\mathbf{P}^{(m)}, \mathbf{T}_{2D}^{(f)})\}_{m \in p_{c_{\text{sel}}}}$ are computed, where $\mathbf{T}_{2D}^{(f)} = \{(lat_i, lon_i)\}_{i=1}^{N_f}$ is the planform projection of the flight trajectory. The best-matching nominal procedure index is selected as

$$m_{\text{best}} = \arg \min_{m \in p_{c_{\text{sel}}}} D_{\text{FDTW}}(\mathbf{P}^{(m)}, \mathbf{T}_{2D}^{(f)}). \quad (7)$$

To provide an interpretable match-quality metric, the distance is mapped to a similarity score through linear normalization:

$$S_m^{(f)} = 1 - \frac{D_{\text{FDTW}}(\mathbf{P}^{(m)}, \mathbf{T}_{2D}^{(f)})}{\max_{u \in p} D_{\text{FDTW}}(\mathbf{P}^{(u)}, \mathbf{T}_{2D}^{(f)})}, \quad (8)$$

where $S_m^{(f)} \in [0, 1]$ is the similarity score of flight f with respect to procedure m , and the denominator uses the maximum Fast-DTW distance over all procedures $u \in p$ for scaling. Flights with sufficiently high similarity (e.g., $S_{m_{\text{best}}}^{(f)} \geq 0.8$) are treated as successfully matched to a nominal procedure and retained as “near” samples, while flights with low similarity are treated as operational deviations. Importantly, both near and other samples remain assigned to their corridor-entry class c_{sel} for subsequent within-corridor clustering.

Building on the above rationale, evaluating the trajectory-to-trajectory similarity within each group enables a secondary subdivision of trajectories under the initial classification, thereby providing a sound basis for subsequent clustering. The example shown in Figure 3 offers an intuitive illustration of the role of this step.

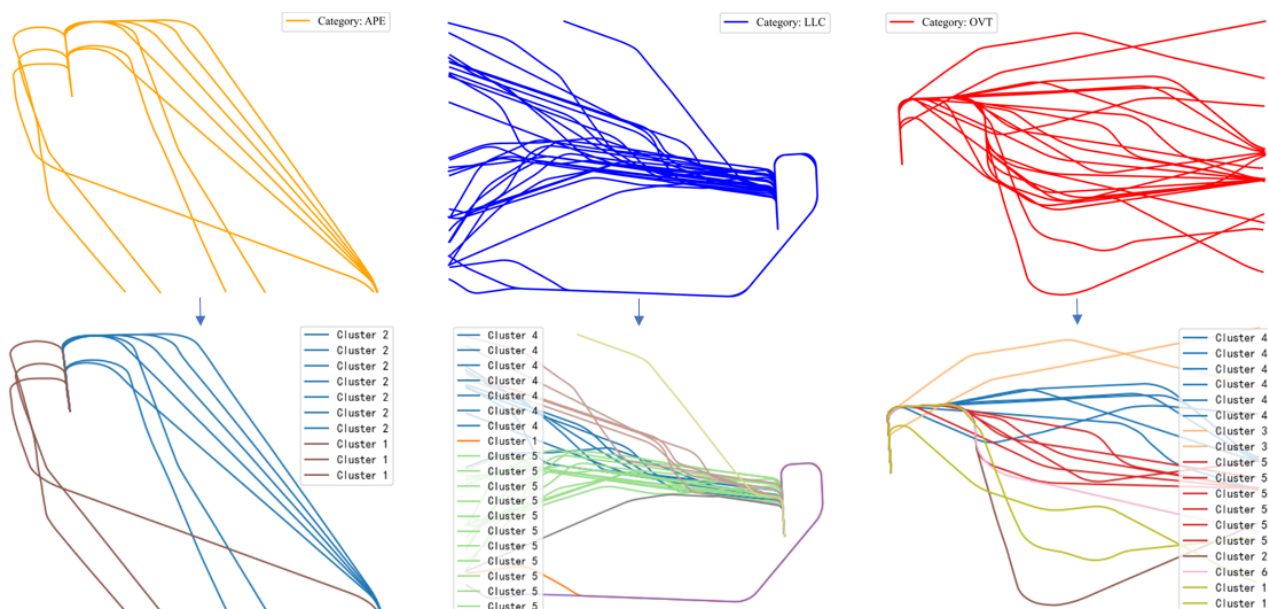


Figure 3. Reclassification of similar trajectories.

2.1.3. 3D Trajectory Clustering via Normalized Fast-DTW

To capture the realistic variability of operational trajectories within each corridor-entry class, a second-stage clustering is performed on the 3D trajectories $\mathbf{T}^{(f)} = \{(lat_i, lon_i, alt_i)\}$. Because latitude/longitude and altitude have different scales and units, each trajectory is normalized using Min–Max scaling before distance computation:

$$\tilde{\mathbf{T}}^{(f)} = \text{MinMax}(\mathbf{T}^{(f)}), \quad (9)$$

where $\tilde{\mathbf{T}}^{(f)}$ denotes the normalized 3D trajectory of flight f . For a given corridor class c containing N_c trajectories (where N_c is the number of flights assigned to corridor class c), we compute pairwise Fast-DTW distances on normalized 3D sequences:

$$\Delta_{pq} = D_{\text{FDTW}}(\tilde{\mathbf{T}}^{(p)}, \tilde{\mathbf{T}}^{(q)}), p, q \in \{1, \dots, N_c\}, \quad (10)$$

where p and q index the trajectories within the same corridor class, and Δ_{pq} is the (symmetric) distance between trajectories p and q . These distances form a symmetric distance matrix $\Delta \in \mathbb{R}^{N_c \times N_c}$.

Hierarchical agglomerative clustering is conducted using the average linkage criterion. Let A and B be two clusters, and let n_A and n_B denote the numbers of trajectories in clusters A and B , respectively. The inter-cluster distance is defined as

$$D(A, B) = \frac{1}{n_A n_B} \cdot \sum_{p \in A} \sum_{q \in B} \Delta_{pq} \quad (11)$$

Starting from singleton clusters, the algorithm iteratively merges the two closest clusters until the linkage distance exceeds a preset maximum distance threshold $d_{\max}$, where $d_{\max}$ is the clustering cut-off parameter controlling the granularity of within-corridor patterns.

2.1.4. Representative Trajectories and Occurrence Probabilities

For each resulting cluster k within corridor class c , a representative (clustered) trajectory $\bar{\mathbf{T}}_{c,k}$ is generated. Let $\mathcal{C}_{c,k}$ denote the set of flights (trajectories) belonging to cluster k within corridor c , and let $n_{c,k} = |\mathcal{C}_{c,k}|$ be the number of trajectories in that cluster. To ensure a consistent point-wise average, trajectories in the same cluster are truncated to the minimum length

$$L_{c,k} = \min_{\mathbf{T}^{(f)} \in \mathcal{C}_{c,k}} N_f, \quad (12)$$

and the truncated trajectory for flight f is $\mathbf{T}_{\text{tr}}^{(f)} = \{(lat_i, lon_i, alt_i)\}_{i=1}^{L_{c,k}}$. The representative trajectory is then obtained by point-wise averaging:

$$\bar{\mathbf{T}}_{c,k} = \frac{1}{n_{c,k}} \sum_{\mathbf{T}^{(f)} \in \mathcal{C}_{c,k}} \mathbf{T}_{\text{tr}}^{(f)}. \quad (13)$$

Finally, the three-dimensional distribution probability of each representative trajectory is estimated by the relative frequency of original trajectories assigned to that cluster:

$$p_{c,k} = \frac{n_{c,k}}{\sum_u n_{c,u}}, \quad (14)$$

where $p_{c,k}$ is the probability weight of cluster k within corridor class c , and the denominator sums over all clusters u in corridor c . These probabilities are subsequently used as weighting factors in the probabilistic noise calculation, enabling cumulative noise estimation that explicitly accounts for air traffic flow and operational variability.

By applying a two-step process of clustering followed by hierarchical grouping to actual flight trajectories, the three-dimensional distribution probabilities of clustered trajectories can be effectively extracted. This approach provides essential support for noise analysis under the influence of air traffic flow. The overall research framework for trajectory distribution probability is illustrated in Figure 4.

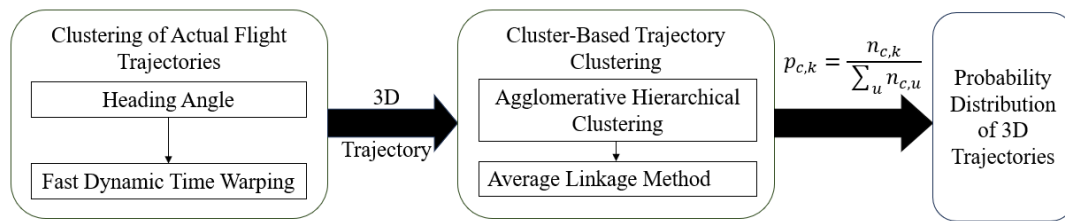


Figure 4. Framework for trajectory distribution probability research.

2.2. Trajectory Operation Features Prediction

Due to the missing flight state parameters in clustered trajectories, this study proposes a parameter prediction method based on actual flight QAR data. Specifically, an ensemble learning approach combining Random Forest and XGBoost is adopted. This ensemble strategy leverages the strengths of both models, enhancing prediction accuracy and robustness. The input features of the model include key parameters such as ground speed, climb angle, pitch angle, and engine power, which comprehensively reflect changes in flight state and provide reliable data support for noise calculation. Therefore, we construct a data-driven parameter completion module based on QAR observations, consisting of (i) equal-distance trajectory resampling and segment-wise feature extraction, and (ii) supervised learning for multi-parameter prediction with an ensemble of Random Forest and XGBoost.

2.2.1. Segment-Level Feature Extraction via Equal-Distance Resampling

After removing points within a predefined latitude–longitude rectangle to mitigate local non-representative effects, the along-track distance between two consecutive trajectory points $p_{t-1}(lat_{t-1}, lon_{t-1})$ and $p_t(lat_t, lon_t)$ is computed using the Haversine formula:

$$d_t = 2R \arcsin \left(\sqrt{\sin^2 \frac{\Delta lat_t}{2} + \cos lat_{t-1} \cos lat_t \sin^2 \frac{\Delta lon_t}{2}} \right), \quad (15)$$

where R is the Earth radius (m), lat_t and lon_t denote latitude and longitude in radians, $\Delta lat_t = lat_t - lat_{t-1}$, $\Delta lon_t = lon_t - lon_{t-1}$, and d_t is the geodesic distance between consecutive points (m). A segment is formed once the cumulative along-track distance reaches a preset threshold D_0 (m; $D_0 = 5000$ m in the implementation). This step is an equal-distance discretization of each trajectory: it groups consecutive recorded points into ~5 km path-length segments so that segment-wise statistics represent comparable local flight states, independent of the original time sampling or speed variations.

For each segment s , we compute aggregated descriptors and define the input vector

$$x_s = [D_s, m_s, A_s, E_s, C_s], \quad (16)$$

where D_s is the segment distance (m), m_s is the segment-mean aircraft mass (kg) converted from recorded gross weight, A_s is the aircraft type category, E_s is the engine type category, and C_s is the corridor class obtained from a predefined identifier-to-corridor mapping. The multi-output target vector is

$$y_s = [v_{tg,s}, \varphi_s, \gamma_s, P_s^*], \quad (17)$$

where $v_{tg,s}$ is the segment-mean ground speed, φ_s is the segment-mean bank angle, γ_s is the segment-mean climb angle, and P_s^* is a propulsion-intensity proxy.

To ensure consistency with the subsequent noise-model interpolation inputs, the propulsion proxy is represented by a thrust-setting variable that is directly compatible with the NPD tables. Specifically, the propulsion intensity (thrust setting) is obtained via an

“ $N1$ (temperature-corrected to $N1_c$) $\rightarrow$ linear mapping to the NPD ‘Power Setting’ domain $\rightarrow$ clipping” procedure to yield P_s^* (lbf/engine). The ECAC Doc 29 NPD methodology uses power/thrust setting together with distance as key inputs for noise computation and interpolation; mapping segment-level $N1_c$ onto the NPD thrust axis therefore provides thrust inputs that are consistent with the downstream NPD interpolation process [12]. In addition, prior work has estimated $N1$ /power setting as a required input for noise calculations and has modelled take-off thrust settings accordingly, supporting the feasibility of using $N1$ -type indicators as a surrogate for thrust setting [32].

For each segment s , the segment-mean engine rotational-speed parameter $N1_s$ is first computed from QAR records by averaging across engines (e.g., for the A320, averaging $N11$ and $N12$). To account for the influence of ambient temperature on $N1$, $N1_s$ is corrected using the static air temperature (SAT) when available, yielding the corrected speed $N1_{c,s}$:

$$\theta_s = \frac{T_s}{288.15}, T_s = SAT_s + 273.15, N1_{c,s} = \frac{N1_s}{\sqrt{\theta_s}}, \quad (18)$$

where θ_s is the non-dimensional temperature ratio for segment s ; T_s is the segment-mean static air temperature in Kelvin (K); SAT_s is the segment-mean static air temperature in degrees Celsius ($^{\circ}\text{C}$) obtained from the QAR SAT field; 288.15 K is the ISA sea-level reference temperature; $N1_s$ is the segment-mean $N1$ (%), i.e., the low-pressure rotor speed expressed as a percentage; and $N1_{c,s}$ is the temperature-corrected $N1$ (%). If SAT_s is unavailable, $N1_{c,s}$ is set to $N1_s$.

To convert $N1_{c,s}$ into a thrust setting that can be directly used for NPD interpolation, dataset-wide quantiles of $N1_c$ are computed to mitigate the influence of outliers. Let $N1_{c,\text{low}}$ and $N1_{c,\text{high}}$ denote the lower and upper quantiles of $N1_c$ over the full dataset. The 5th and 95th percentiles are used as robust mapping endpoints: trimming 5% at each tail reduces the leverage of a small number of abnormal or transient observations on the endpoints, yielding a more stable linear mapping that better represents the typical operational envelope [33,34]. The corrected speed $N1_{c,s}$ is then linearly mapped to the thrust-setting domain $[P_{\min}, P_{\max}]$ defined by the “Power Setting” column of the aircraft-specific NPD table (lbf/engine):

$$r_s = \frac{N1_{c,s} - N1_{c,\text{low}}}{N1_{c,\text{high}} - N1_{c,\text{low}}}, P_s^* = P_{\min} + r_s(P_{\max} - P_{\min}), \quad (19)$$

and subsequently clipped to ensure that the resulting thrust setting strictly lies within the NPD domain:

$$P_s^* \leftarrow \min(\max(P_s, P_{\min}), P_{\max}). \quad (20)$$

Here, r_s is a dimensionless normalized mapping coefficient; $N1_{c,\text{low}}$ and $N1_{c,\text{high}}$ are the 5th and 95th percentiles of $N1_c$ (%) over the full dataset; $P_{\min}$ and $P_{\max}$ are the minimum and maximum thrust-setting values (lbf/engine) in the NPD “Power Setting” column, defining the admissible thrust domain; and P_s^* is the segment s thrust setting (lbf/engine), used as the power component of the multi-output target vector. This construction ensures that the predicted thrust setting is physically consistent with the NPD thrust axis and bounded within its valid range, enabling direct use in subsequent NPD-based noise interpolation.

Before training, numerical predictors (D_s, m_s) are Min–Max normalized, while categorical predictors (A_s, E_s, C_s) are one-hot encoded. The targets are also normalized during training and inverse-transformed for physical interpretation. Samples with evidently invalid states are removed to enhance stability.

2.2.2. Flight-State Prediction via RF–XGBoost Ensemble Regression

We employ two complementary tree-based models—Random Forest (RF) and XGBoost (XGB)—to learn the mapping from $\tilde{\mathbf{x}}_s$ to $\tilde{\mathbf{y}}_s$, where $\tilde{\cdot}$ indicates normalized variables.

For Random Forest regression, the predicted multi-output vector is obtained by averaging the outputs of B trees [35]:

$$\hat{\mathbf{y}}_s^{RF} = \frac{1}{B} \sum_{b=1}^B \mathbf{T}_b(\tilde{\mathbf{x}}_s), \quad (21)$$

where $\mathbf{T}_b(\cdot)$ is the b -th regression tree. Each tree is grown by selecting splits that maximize the reduction in within-node variance

For XGBoost regression, the prediction is represented as an additive model of K trees [36]:

$$\hat{\mathbf{y}}_s^{XGB} = \sum_{k=1}^K \mathbf{f}_k(\tilde{\mathbf{x}}_s), \mathbf{f}_k \in \mathcal{F}, \quad (22)$$

where $\mathcal{F}$ denotes the space of regression trees, and $\mathbf{f}_k(\cdot)$ produces a vector output for the four targets. The regularized objective minimized by XGB is

$$\mathcal{L} = \sum_{i=1}^n \ell + \sum_{k=1}^K \Omega(\mathbf{f}_k), \quad (23)$$

where n is the number of training segments, $\ell(\cdot)$ is the loss, and $\Omega(\cdot)$ penalizes model complexity.

Finally, we fuse RF and XGB by a linear ensemble:

$$\hat{\mathbf{y}}_s = \omega \hat{\mathbf{y}}_s^{RF} + (1 - \omega) \hat{\mathbf{y}}_s^{XGB}, \omega \in [0,1], \quad (24)$$

where ω is the ensemble weight. In the implementation, $\omega = 0.5$. Averaging reduces generalization error when the two models exhibit non-identical and imperfectly correlated errors, and provides a stable prediction for subsequent modelling [37].

Diverse sample sets are generated using bootstrap sampling, and a random feature selection mechanism is applied to ensure model diversity and generalization capability. Specifically, the XGBoost and Random Forest models are trained independently. Then, their prediction results are integrated using an averaging method, resulting in the final prediction output of the fused XGBoost–Random Forest ensemble model.

2.3. Operational-Knowledge-Driven Trajectory Noise Precicting

Based on trajectory operational knowledge a noise prediction model is developed to account for the influence of air traffic flow. In this study, the Day-Night Equivalent Sound Level is selected as the evaluation metric for airport noise. Using trajectory operational knowledge derived from QAR data, the Noise–Power–Distance (NPD) relationship is obtained. The aircraft noise levels are then calculated following the methodology outlined in Document 29 of the European Civil Aviation Conference (ECAC) [12].

2.3.1. Single-Event Noise Computation Using Predicted Flight States

The single-event noise at a ground receptor is computed using the ECAC Doc.29 framework, where the Sound Exposure Level (SEL) of one operation is evaluated through NPD (Noise–Power–Distance) based lookup/interpolation combined with standard operational and geometric corrections. In order to apply the Doc.29 engine to real operations, this study uses the outputs of the flight-parameter prediction module to provide the segment-wise operational state required by the noise calculation, thereby reducing reliance

on nominal assumptions about aircraft behavior. The workflow of single-event noise calculation based on predicted flight parameters is shown in the Figure 5 below.

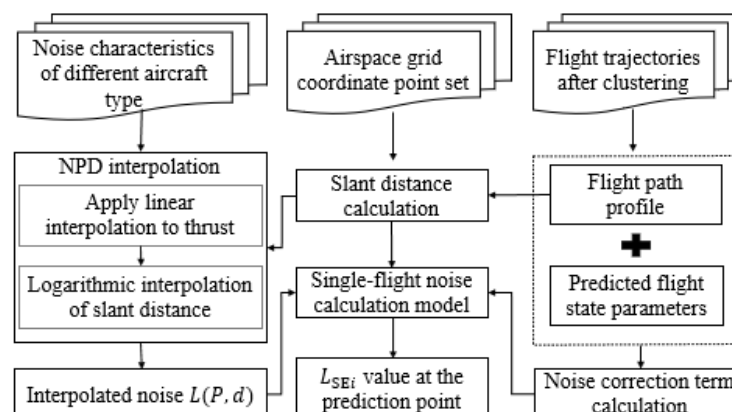


Figure 5. Flowchart of noise calculation.

A complete single-event computation requires four categories of inputs. First, an aircraft/engine-specific acoustic dataset is needed, i.e., NPD tables that relate reference noise levels to an engine power/thrust descriptor and slant distance under standard conditions. Second, the flight-path geometry must be available as a discretized 3D trajectory, including (lat, lon) and altitude along the path, so that the relative geometry between the aircraft and each receptor can be determined. Third, the calculation requires segment-level operational variables, which in this work are provided by the prediction module: ground speed v_{tg} , climb angle γ , bank angle ϕ , and a propulsion-related variable serving as an engine thrust proxy. These variables jointly define the operating condition of each trajectory segment for NPD interpolation and for the correction structure. Fourth, the receptor definition and reference geometry are required, including the receptor location and the runway reference direction used to determine lateral offset and directional conditions in the correction terms.

In implementation, the 3D trajectory is divided into short consecutive segments so that each segment can be approximated as locally uniform. For each segment–receptor pair, the slant distance and associated geometric quantities are computed from the segment endpoints, altitude, and receptor coordinates. The predicted propulsion proxy is mapped to a consistent power/thrust descriptor compatible with the selected NPD dataset, after which the base noise level is obtained by NPD interpolation in the power/thrust and distance dimensions. The base level is then adjusted using the Doc.29 correction structure to account for operational and geometric effects, including speed-related correction (driven by v_{tg}), lateral attenuation (driven by receptor–track geometry), installation/directivity effects, and directional handling associated with the reference takeoff direction when applicable. Finally, the single-event L_{SEi} at the receptor is obtained by energetically accumulating the corrected contributions over all segments along the trajectory.

2.3.2. Cumulative Noise Estimation via Probability Weighting

A cumulative noise event refers to the aggregated noise impact generated by multiple aircraft operations over a specific period of time, taking into account factors such as the number of flights and their temporal distribution. The proposed model integrates trajectory operational knowledge with a fundamental noise calculation framework to construct a representative terminal traffic flow scenario. The detailed integration methodology is illustrated in Figure 6.

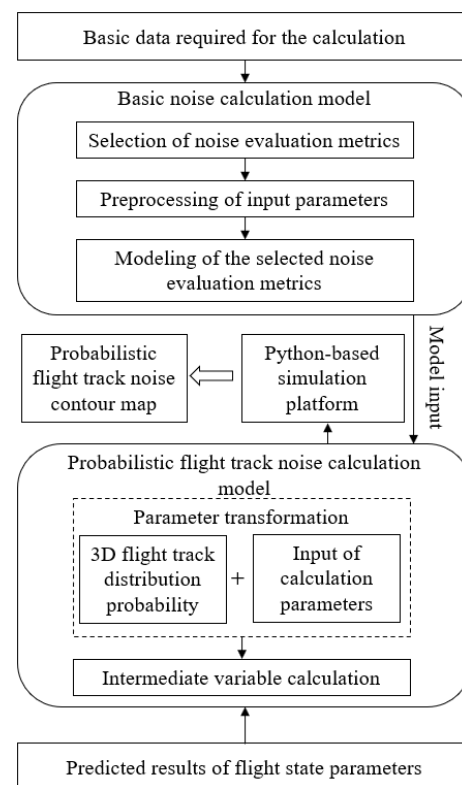


Figure 6. Flowchart of noise calculation modeling based on trajectory operation knowledge.

By summing all significant sound exposure levels from aircraft activities, the total noise exposure level at a given observation point can be determined. The calculation is expressed by the following formula:

$$L_{dn} = \overline{L_{SE}} + 10 \lg(N_d + 10N_n) - 49.4, \quad (25)$$

In this equation, N_d denotes the number of daytime flights (06:00–22:00), and N_n denotes the number of nighttime flights (22:00–06:00). $\overline{L_{SE}}$ represents the average sound exposure level equivalent for N flights during both day and night, measured in decibels (dB).

$$\overline{L_{SE}} = 10 \lg \left[\frac{1}{N} \sum_{i=1}^N 10^{0.1 L_{SEi}} \right], \quad (26)$$

$$L_{SEi, corrected} = L_{SEi}(P, d) + \Delta v - \Lambda(\beta, l) + \Delta L + \Delta \varphi. \quad (27)$$

In this equation, $L_{SEi, corrected}$ is the actual sound exposure level of a single flight event, measured in decibels (dB); $L_{SEi}(P, d)$ represents the interpolated noise level from the NPD (Noise–Power–Distance) curves, also in dB; P is the engine thrust, measured in kilonewtons (kN); d is the shortest distance from the ground calculation point to the flight trajectory, in meters (m); Δv is the speed correction factor; $\Lambda(\beta, l)$ is the lateral attenuation factor, where β is the elevation angle of the observation point relative to the aircraft trajectory ($^\circ$), and l is the perpendicular distance from the observation point to the aircraft ground track, in meters (m); ΔL is the correction factor applied for observation points located behind the aircraft takeoff position; and $\Delta \varphi$ is the engine installation position correction.

Assume that a cluster of actual flight trajectories yields n clustered trajectories after clustering, hence the corresponding probabilities denoted as $p_{c,k,1}, p_{c,k,2}, \dots, p_{c,k,3}$, as computed by Equation (14). Let the individual clustered trajectories be represented by $z_1, z_2, \dots, z_n$. Then, the single-event noise impact model at a given grid center point within the terminal area, considering the probabilistic distribution of the trajectories, is expressed as follows:

$$L_{SEi}^* = L(z_1)p_{c,k,1} + L(z_2)p_{c,k,2} + \dots + L(z_n)p_{c,k,n} \quad (28)$$

In this equation, L_{SEi}^* represents the single-event noise level at a terminal area grid center point under the influence of a given trajectory cluster, measured in decibels (dB); $L(z_i)$ denotes the single-event noise ($L_{SEi,corrected}$) contribution from the i th trajectory within the cluster, in dB; and $p_{c,k,n}$ denotes the probability associated with the trajectory.

By enhancing the basic noise model, the proposed approach avoids the computational complexity of evaluating noise levels for multiple individual trajectories as required in traditional methods. The introduction of probabilistic clustered trajectories effectively reduces the influence of infrequent and distant flight paths, while increasing the weighting of trajectories that have significant impacts on observation points. As a result, the accuracy of noise prediction is significantly improved.

3. Results

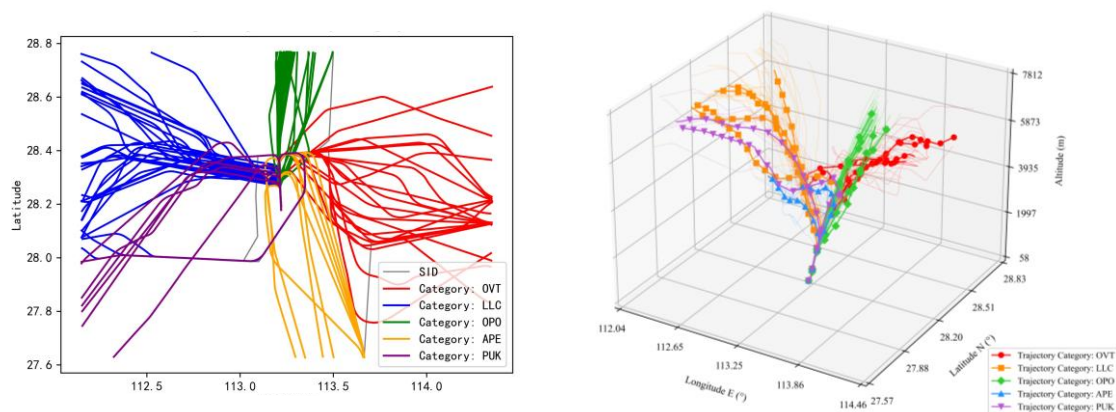
QAR data from Changsha Huanghua International Airport on a single day were utilized to establish the noise analysis model. The modeling process involved sequentially performing trajectory operational knowledge mining and noise calculation modeling. Subsequently, noise contour maps corresponding to clustered trajectories at various corridor entry points in the terminal area of Changsha Airport were generated through simulation. Finally, a comparative analysis was conducted to verify the effectiveness of the proposed method.

3.1. Trajectory Clustering Results

To reduce computational complexity, each trajectory in the QAR dataset is segmented and resampled. The latitude and longitude information of trajectory points is converted into radians to facilitate distance calculation. The terminal area boundary and standard arrival/departure procedures are incorporated for heading angle matching. The latitude, longitude, and altitude data of trajectory points are organized and fitted to reconstruct operational trajectories. By comparing the actual flight direction of each trajectory with its corresponding nominal path, an initial classification is performed based on different corridor entry points. To improve classification accuracy and avoid misclassification caused by similar heading angles, the Fast Dynamic Time Warping (Fast-DTW) algorithm is employed. Prior to clustering, the latitude, longitude, and altitude data are normalized. During clustering, the Euclidean distance between each pair of trajectories is calculated using the Fast-DTW algorithm, and a distance matrix is constructed to store all pairwise distances. A maximum distance threshold is selected through parameter tuning, and hierarchical clustering is performed using the average linkage method. At each iteration, the two closest clusters are merged until all trajectories in the same direction are grouped into clusters representing distinct representative trajectories. After clustering, the results are denormalized to produce the final clustered trajectories. A reasonableness check was also conducted for the clustered trajectories, showing that the maximum climb angle is 11.181° , the mean climb angle is 3.501° , and the maximum bank angle during turns is 22.681° . These values comply with the requirements specified by the flight procedures.

Due to the varying number of actual flight trajectories in different directions, each cluster yields a different number of representative clustered trajectories. The clustering results are shown in Figure 7. The classifications are labeled as OVT, LLC, OPO, APE, and PUK (Corridor entry direction name). Figure 7a presents the classification results based on corridor entry points, where the grey polylines denote the standard departure procedures. Figure 7b illustrates the three-dimensional clustering results incorporating altitude information. In Figure 7b, darker-colored trajectories represent the clustered

(representative) trajectories, whereas lighter-colored trajectories correspond to the actual flight trajectories within each respective category.



(a) Schematic Diagram of 2D Trajectory Classification (b) Schematic Diagram of 3D Trajectory Clustering

Figure 7. Schematic Diagram of Trajectory Clustering.

The number of actual flight trajectories corresponding to each clustered trajectory was counted. Based on this, the flight probability of each clustered trajectory was calculated using Equation (14). The results are presented in Table 1.

Table 1. Distribution Probability of Clustered Trajectories.

Classification Category	Clustered Flight Path	Flight Path Probability	Classification Category	Clustered Flight Path	Flight Path Probability
OVT	4	0.2083	OPO	2	0.5833
OVT	5	0.0833	OPO	3	0.1250
OVT	2	0.5833	OPO	1	0.1250
OVT	3	0.0416	OPO	4	0.1458
OVT	1	0.0833	OPO	5	0.0208
LLC	5	0.4000	APE	1	0.7000
LLC	1	0.1000	APE	2	0.3000
LLC	2	0.3000	PUK	1	0.8571
LLC	4	0.1666	PUK	2	0.1428
LLC	3	0.0333			

3.2. Flight-State Parameter Prediction Results

To predict flight state parameters, feature extraction and coordinate point sampling were performed on the QAR dataset, with parameters such as climb angle and engine thrust explicitly defined. Four flight state parameters—ground speed, climb angle, pitch angle, and engine power—were predicted using Random Forest and XGBoost models. A total of 80% of the data was used for training and 20% for testing. For the Random Forest model, the number of decision trees was set to 200, with a maximum depth of 5. The minimum number of samples required to split an internal node was 10, and the minimum number of samples at a leaf node was 5. The random seed was fixed at 42. For the XGBoost model, parameters were set as follows: 200 decision trees, a maximum depth of 5, a tree-based booster, a learning rate of 0.1, and a random seed of 42. The prediction results are shown in Figure 8. The results indicate that the model performs well in predicting ground speed, climb angle, and engine thrust, whereas the prediction accuracy for bank angle is relatively lower. Overall, as long as engine thrust can be predicted with sufficient accuracy, the bank angle is only used in the engine installation correction term within the

noise-correction calculations, where its value is typically on the order of one to two decimal places; therefore, its impact on the final noise metrics is negligible. Consequently, the model is suitable for predicting trajectory-specific flight parameters.

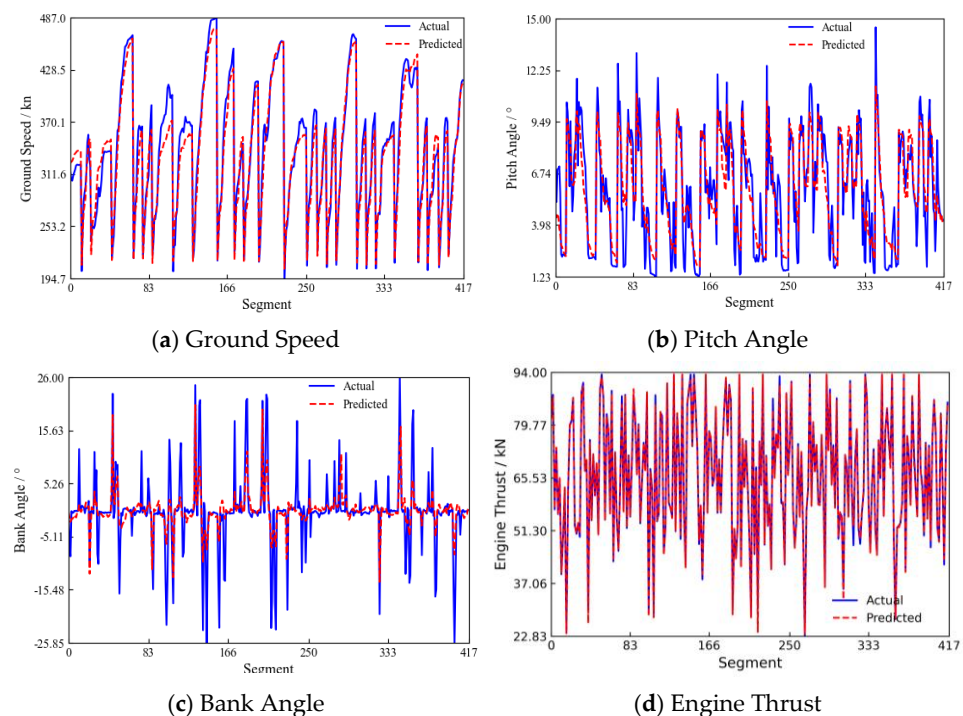
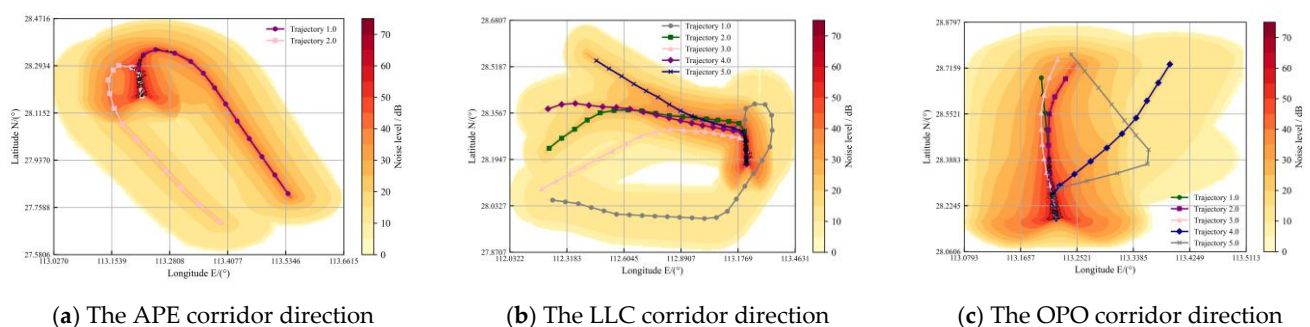


Figure 8. Comparison of real and predicted values of various characteristic parameters.

3.3. Trajectory Noise Estimation Results

Taking the A320 aircraft as an example, the corresponding Noise–Power–Distance (NPD) interpolation table is imported, and the relevant acoustic parameters are selected accordingly. Each input variable required for single-event noise estimation is implemented as a dedicated function, which are then integrated into a comprehensive noise computation function based on the established model.

Subsequently, statistical analyses are conducted on the total number of flights and their distribution across daytime and nighttime periods to compute cumulative noise exposure. For each terminal corridor direction, the flight probabilities of clustered trajectories are introduced, and the associated noise contributions are weighted and averaged accordingly. This process yields directional noise contour maps based on probabilistic trajectory distributions, as illustrated in Figure 9.



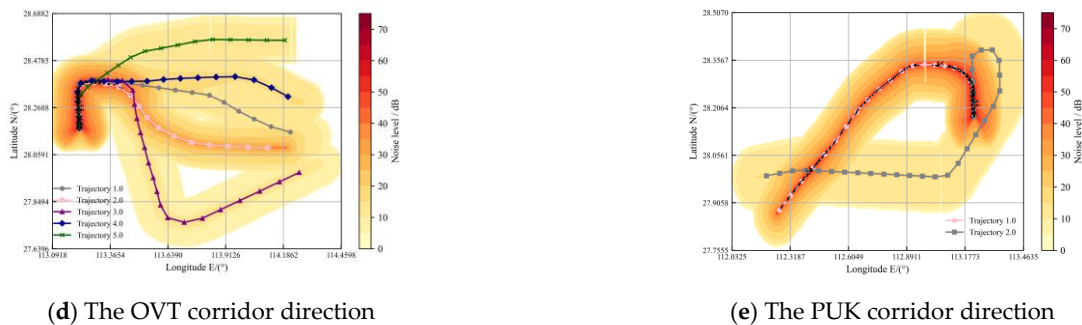


Figure 9. Contour map of clustering track noise in all directions.

From a macroscopic perspective, taking the APE and PUK direction as examples, a comparison was made between the noise impact of clustered trajectories and that of actual flight trajectories, as shown in Figure 10. Since actual flight trajectories have not undergone clustering or parameter prediction, the resulting noise calculations most accurately reflect the noise impact under real operational conditions. Therefore, by comparing the noise computed based on trajectory operational knowledge with that derived from actual flight trajectories, the effectiveness of the proposed noise prediction method is verified.

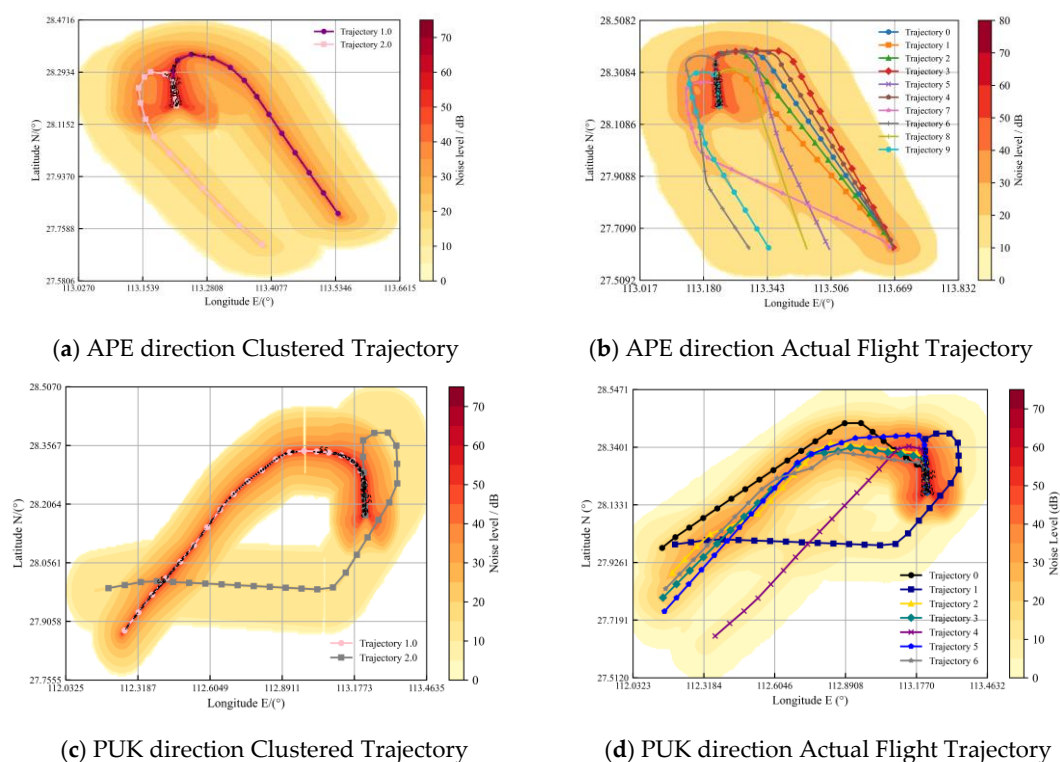


Figure 10. Comparison of clustering flight paths and actual flight path noise contour maps.

According to the guidelines for aircraft noise calculation and prediction around civil airports [38], noise analysis of day-night equivalent sound levels typically focuses on the range of 55–75 dB. Accordingly, this study compared the noise impact areas of the two types of trajectories across thresholds of ≥ 55 dB, 60 dB, 65 dB, 70 dB, and 75 dB. A comparison of the noise impact areas of the two trajectory types for the two corridor entry directions was conducted, and the results are shown in Tables 2 and 3. Although minor differences exist in the overall noise impact range, the two approaches exhibit generally similar patterns. Specifically, for regions exceeding 55 dB, the difference in impact area does not exceed 1 km².

The clustered trajectories were also compared with the nominal trajectories. As shown clearly in Figure 11, there are notable differences between the noise impact areas of the clustered and actual flight trajectories, with the clustered trajectories generally producing a broader noise impact area. This further demonstrates the necessity and rationality of the proposed noise prediction method.

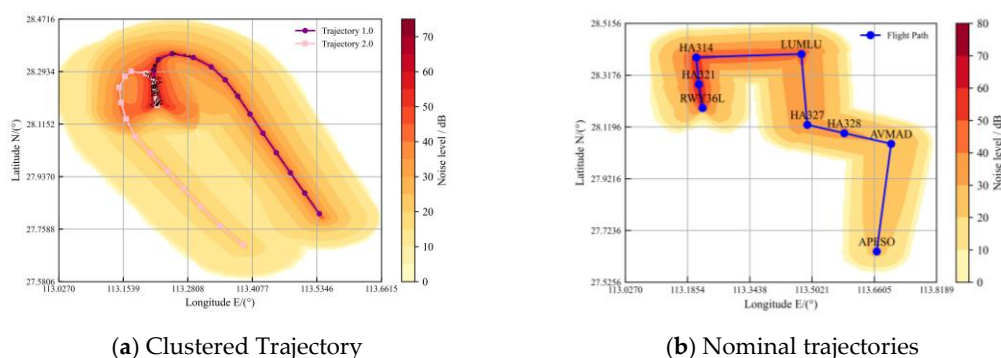


Figure 11. Comparison of APE direction clustering track noise and nominal track sound contour maps.

Table 2. Comparison of noise impact area of APE direction.

Noise Level/dB	Actual Flight Path Impact Area/km ²	Clustered Flight Path Impact Area/km ²
≥55	27.10	27.85
≥60	11.29	9.28
≥65	4.02	3.86
≥70	1.83	1.51
≥75	0.46	0.43

Table 3. Comparison of noise impact area of PUK direction.

Noise Level/dB	Actual Flight Path Impact Area/km ²	Clustered Flight Path Impact Area/km ²
≥55	61.23	63.99
≥60	26.85	28.86
≥65	11.54	12.55
≥70	5.02	4.82
≥75	0.51	0.55

From a microscopic perspective, noise levels at noise-sensitive locations were calculated for each flight segment during the aircraft departure process, and noise variation curves were plotted. The noise impact curves at the same observation point were compared between the probabilistic (clustered) trajectories and actual flight trajectories. If both exhibit similar trends, it indicates that the probabilistic trajectories can effectively substitute actual trajectories in representing the noise impact on observation points.

The results show that the probabilistic trajectories exhibit a broadly similar noise impact trend to that of the actual flight trajectories at the same noise-sensitive point, as illustrated in Figure 12. In summary, compared with actual flight trajectories, the use of clustered trajectories significantly reduces computational complexity and demonstrates superior performance in predicting high-noise regions. Therefore, the proposed method can serve as an effective approach for terminal area noise analysis.

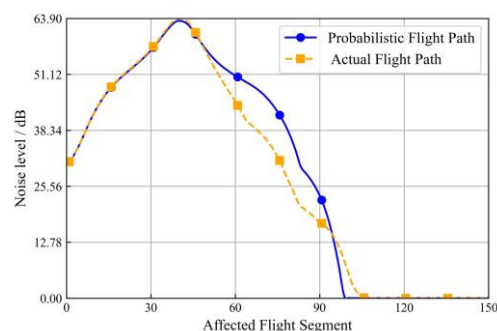


Figure 12. Comparison of the trend in noise variation at a single observation point affected by different trajectories.

4. Conclusions

The operational-knowledge-driven framework proposed in this study establishes a coherent mapping from “traffic flow–trajectory patterns–operational states–acoustic contributions”. Compared with the event-by-event baseline using actual trajectories, it reproduces the spatial morphology of noise contours and the spatiotemporal variation trends at typical receptors with good consistency, while substantially reducing the computational burden associated with large-scale trajectory sets.

Specifically, (a) since terminal-area noise exposure is systematically influenced by air traffic flow, the probabilistic representative trajectories preserve dominant operational modes and their usage frequencies while attenuating the disproportionate impact of a small number of atypical tracks, thereby offering a practical trade-off between physical realism and computational tractability for routine prediction and rapid scenario screening; (b) trajectory-occurrence probabilities provide a principled weighting mechanism for cumulative acoustic impact, ensuring that frequent patterns dominate aggregate exposure whereas infrequent patterns contribute proportionally, and segment-wise parameter completion supplies the operational states required by the standardized noise engine, reducing reliance on fixed-profile assumptions and improving robustness across varying demand levels and operational regimes; and (c) from an engineering standpoint, the framework achieves a favorable balance between efficiency and accuracy, supporting comparative assessments of alternative route structures and operational strategies.

Although the framework is grounded in observed operational trajectories and standardized noise computation and is therefore inherently transferable, its performance may still vary with runway configurations, route-network complexity, terrain constraints, and procedure-design differences (e.g., multi-runway mode switches, strong coupling between arrival and departure routes, and constrained turning geometry). In such settings, the clustering granularity, corridor identification, and probability-estimation strategies may require adaptive tuning to better capture topology-driven operational modes. Moreover, extreme meteorological and operational conditions—such as strong crosswinds, convective weather, low-visibility operations, and frequent runway-configuration changes—can induce non-stationary trajectory dispersion and atypical thrust/flight-state patterns, posing higher requirements on the robustness of both representative-trajectory construction and parameter prediction. Future work may therefore explore weather-regime-conditioned clustering and probability modelling, domain-adaptive state-prediction models, and uncertainty quantification to express predictive confidence under rare or rapidly changing conditions, thereby improving robustness for cross-airport deployment and extreme-weather applications.

Author Contributions: Z.T.: Conceptualization, Methodology. Y.F.: Writing—original draft. X.C.: Data curation. X.S.: Investigation. Z.N.: Writing—review & editing. Y.Z.: Validation. M.J.: Supervision. X.T.: resources. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2024YFB2605201, Open Project of Sichuan Provincial Engineering Technology Research Center for Civil Aviation Flight Technology and Safety, grant number GY2024-03B, and the Fundamental Research Funds for the Central University, grant number NJ2024023.

Data Availability Statement: The dataset presented in this article are not readily available because the authors have no permission to share them.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QAR	Quick Access Recorder
QNH	Altimeter setting to mean sea level pressure
SI	Similarity Index
SID	Standard Instrument Departure
STAR	Standard Terminal Arrival Route
ATC	Air Traffic Control
DTW	Dynamic Time Warping
Fast-DTW	Fast Dynamic Time Warping
SAT	Static Air Temperature
ISA	International Standard Atmosphere
N1%	low-pressure rotor speed expressed as a percentage
RF	Random Forest
XGB	XGBoost
NPD	Noise–Power–Distance
SEL	Sound Exposure Level
ICAO	International Civil Aviation Organization
FAA	Federal Aviation Administration
AEDT	Aviation Environmental Design Tool
ECAC	European Civil Aviation Conference
EASA	European Union Aviation Safety Agency
EEA	European Environment Agency

References

1. Basner, M.; Clark, C.; Hansell, A.; Hileman, J.I.; Janssen, S.; Shepherd, K.; Sparrow, V. Aviation Noise Impacts: State of the Science. *Noise Health* **2017**, *19*, 41–50.
2. World Health Organization. *Environmental Noise Guidelines for the European Region*; WHO Regional Office for Europe: Copenhagen, Denmark, 2018; ISBN 978-92-890-5356-3.
3. Münzel, T.; Gori, T.; Babisch, W.; Basner, M. Cardiovascular effects of environmental noise exposure. *Eur. Heart J.* **2014**, *35*, 829–836.
4. Janssen, S.A.; Centen, M.R.; Vos, H.; van Kamp, I. The effect of the number of aircraft noise events on sleep quality. *Appl. Acoust.* **2014**, *84*, 9–16.
5. Saucy, A.; Schäffer, B.; Tangemann, L.; Vienneau, D.; Wunderli, J.-M.; Röösli, M. Does night-time aircraft noise trigger mortality? A case-crossover study on 24 886 cardiovascular deaths. *Eur. Heart J.* **2021**, *42*, 835–843.
6. European Parliament and Council of the European Union. *Directive 2002/49/EC of 25 June 2002 relating to the assessment and management of environmental noise*. *Off. J. Eur. Communities* **2002**, *L189*, 12–25.
7. International Civil Aviation Organization (ICAO). *Guidance on the Balanced Approach to Aircraft Noise Management (Doc 9829)*, 2nd ed.; ICAO: Montreal, QC, Canada, 2008.

8. International Civil Aviation Organization (ICAO). *Recommended Method for Computing Noise Contours Around Airports (Doc 9911)*, 2nd ed.; ICAO: Montreal, QC, Canada, 2018.
9. European Union Aviation Safety Agency (EASA); European Environment Agency (EEA); EUROCONTROL. *European Aviation Environmental Report*; Publications Office of the European Union: Luxembourg, 2022.
10. Air Traffic Management Bureau of China Civil Aviation Administration. *Construction of Visual and Instrument Flight Procedures: MH/T 4023-2007*; China Civil Aviation Publishing House: Beijing, China, 2014.
11. Federal Aviation Administration (FAA). *Aviation Environmental Design Tool (AEDT) Version 2b Service Pack 2 Technical Manual*; DOT-VNTSC-FAA-16-09; U.S. Department of Transportation: Washington, DC, USA, 2016.
12. European Civil Aviation Conference (ECAC). *Report on Standard Method of Computing Noise Contours Around Civil Airports (ECAC Doc 29)*, 4th ed.; ECAC: Paris, France, 2016.
13. Clemente, J.; Gaja, E.; Clemente, G.; Reig, A. Sensitivity of the FAA Integrated Noise Model to input parameters. *Appl. Acoust.* **2005**, *66*, 263–276.
14. Zaporozhets, O.; Levchenko, L. Accuracy of Noise–Power–Distance definition on results of single aircraft noise event calculation. *Aerospace* **2021**, *8*, 121.
15. van der Grift, R.C.; Snellen, M.; Amiri-Simkooei, A. The effect of using N1% as input for aircraft noise modelling. *Transp. Res. Part D Transp. Environ.* **2025**, *142*, 104710.
16. Brooks, C.; Hansman, R.J. *Modeling the Effects of Aircraft Flight Track Variability on Community Noise Exposure*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2017.
17. Pretto, M.; Dorbolò, L.; Giannattasio, P.; Zanon, A. Aircraft operation reconstruction and airport noise prediction from high-resolution flight tracking data. *Transp. Res. Part D Transp. Environ.* **2024**, *135*, 104397.
18. Rindfleisch, T.C.; Alonso, J.J.; Jackson, D.C.; Munguía, B.C.; Bowman, N.W. A large-scale validation study of aircraft noise modeling for airport arrivals. *J. Acoust. Soc. Am.* **2024**, *155*, 1928–1949.
19. Gariel, M.; Srivastava, A.N.; Feron, E. Trajectory clustering and an application to airspace monitoring. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 1511–1524.
20. Basora, L.; Morio, J.; Mailhot, C. A Trajectory Clustering Framework to Analyse Air Traffic Flows. In Proceedings of the 7th SESAR Innovation Days (SIDs 2017), Belgrade, Serbia, 28–30 November 2017.
21. Eerland, W.J.; Box, S. Trajectory clustering, modelling and selection with the focus on airspace protection. In Proceedings of AIAA Infotech@Aerospace (SciTech 2016), San Diego, CA, USA, 4–8 January 2016; p. 1411.
22. Bhanpato, J.; Puranik, T.G.; Mavris, D.N. Data-driven analysis of departure procedures for aviation noise mitigation. *Eng. Proc.* **2021**, *13*, 2.
23. Chen, X.R. Research on Flight Procedure Optimization for Reducing Airport Noise Impact. Master's Thesis, Civil Aviation University of China, Tianjin, China, 2016.
24. Hang, S.Z. Research on the Evaluation of Flight Procedure Design Quality Based on QAR Data. Master's Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2024.
25. Zhao, Y.; Zhang, K.Y. A rapid prediction method for aircraft noise under laterally dispersed flight tracks. *J. Saf. Environ.* **2023**, *23*, 4609–4616.
26. Zhuang, N.J.; Yu, J.X.; Gu, R.P. Four-dimensional track simulation and noise influence analysis of departure flight procedure. *J. Saf. Environ.* **2025**, in press.
27. Wang, L.; Wu, C.; Sun, R. An analysis of flight Quick Access Recorder (QAR) data and its applications in preventing landing incidents. *Reliab. Eng. Syst. Saf.* **2014**, *127*, 86–96.
28. Wang, X.; Liu, M.; Zhao, X. Research on risk analysis of an airport based on QAR data. In *Proceedings of the 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA 2023)*, Chongqing, China, 24–26 February 2023; IEEE: New York, NY, USA, 2023; pp. 1470–1474.
29. Berndt, D.J.; Clifford, J. Using Dynamic Time Warping to Find Patterns in Time Series. In Proceedings of the AAAI-94 Workshop on Knowledge Discovery in Databases (KDD-94), Seattle, WA, USA, 31 July 1994; pp. 359–370.
30. Salvador, S.; Chan, P. Toward accurate dynamic time warping in linear time and space. *Intell. Data Anal.* **2007**, *11*, 561–580.
31. Wu, D.; Lü, R.; Yang, Y.; Deng, J.J.; Zhang, B.S. Research on trajectory identification based on density clustering with dynamic time warping. *Fire Control Command. Control* **2021**, *46*, 73–78.
32. Meister, J. Power setting estimation of departing civil jet aircraft based on machine learning. *J. Journal of Aircraft.* **2024**, *61*, 1195–1204.

33. Datacamp. Winsorized Mean: A Robust Approach to Handling Outliers[EB/OL]. 2024. Available online: <https://www.data-camp.com/tutorial/winsorized-mean> (accessed on 11 February 2026).
34. Boudt, K.; Todorov, V.; Wang, W. Robust Distribution-Based Winsorization in Composite Indicators Construction. *Soc. Indic. Res.* **2020**, *149*, 375–397.
35. Wu, Z.X.; Zhang, N.; Gao, K.Y.; Reng, R. Flight fuel prediction based on a risk-preference adjusted random forest algorithm. *Acta Aeronaut. Astronaut. Sin.* **2022**, *43*, 224933.
36. Yang, J.L.; Li, L.K.; Qian, Y.; Tang, S.X.; Yang, P.D. Fuel flow prediction for domestic civil aircraft based on LSTM-XGBoost. *Adv. Aeronaut. Eng.* **2025**, *16*, 207–214.
37. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009.
38. Civil Aviation Administration of China. *Calculations and Predictions of Airplane Noise Around Airports: MH/T 5105-2007*; Science and Technology of China Press: Beijing, China, 2007.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.