Airspace Collision Risk Analysis Using AI: A Case Study of Criss-Cross Arrivals into Chicago O'Hare

Fahimeh Ghorbani, PhD Student*
Center for Air Transportation Systems Research/System Engineering & Operations Research Dept.
George Mason University, Fairfax, Virginia, and 22030
Email: fghorban@gmu.edu
ORCID: https://orcid.org/0000-0003-0974-5095

Lance Sherry, Professor,
Center for Air Transportation Systems Research/System Engineering & Operations Research Dept.
George Mason University, Fairfax, Virginia, and 22030
Email: lsherry@gmu.edu

John Shortle, Professor,
Center for Air Transportation Systems Research/System Engineering & Operations Research Dept.
George Mason University, Fairfax, Virginia, and 22030
Email: jshortle@gmu.edu

Total Number of Pages: 15

Submitted July 29, 2026

*Corresponding Author

ABSTRACT

Objectives: Criss-cross arrivals on parallel runways intentionally assign aircraft to runways from opposite or non-corresponding sides of terminal airspace, causing arrival streams to cross before final approach. Before approval, a comprehensive collision risk assessment is required to determine whether the procedure meets the target level of safety (TLS). Traditional collision risk methods use simulation models to generate flight tracks for the proposed procedure. Computational limitations on the number of generated flight tracks affect confidence intervals for rare events. Artificial intelligence/machine learning (AI/ML) algorithms for synthetic flight-track generation can produce larger samples for analysis. This paper demonstrates the application of AI/ML-generated synthetic flight tracks to Collision Risk Analysis for the "criss-cross arrivals" procedure at Chicago O'Hare International Airport (ORD).

Methods: The Collision Risk Analysis was conducted using modern AI tools and included assessment of latent collision risk. One thousand synthetic flight tracks were generated for arrivals to ORD Runway 27C and paired with 288 ADS-B historical Runway 28C trajectories. The fleet mix, sequence of flights, and timing were coordinated through a Monte Carlo process. Collision risk was assessed using current and projected proximity metrics under near mid-air collision (NMAC) and Traffic Alert and Collision Avoidance System (TCAS) thresholds.

Findings: For the [−600, 600]-second arrival-time-offset range, the estimated probabilities per simulated pair were 1.98E−04 for NMAC proximity, 2.91E−03 for TCAS proximity, 4.62E−04 for projected NMAC, and 3.06E−03 for projected TCAS. The results are conservative and assume independent arrival streams without controller coordination or intervention.

Novelty: This paper demonstrates the use of AI/ML-generated flight tracks and latent-risk metrics in airspace collision-risk analysis.

Practical Applications: The approach can support regulatory and safety agencies by expanding trajectory data for simulation and enabling efficient evaluation of complex airspace procedures.

1- INTRODUCTION

Criss-cross arrivals on parallel runways are arrival procedures in which aircraft approaching an airport are intentionally assigned to parallel runways from opposite or non-corresponding sides of the terminal airspace (Figure 1). As a result, arrival streams cross each other's paths before they are aligned with their final approach courses (Tang 2015). The crossing nature of the trajectories introduces additional complexity and can increase controller workload as well as the potential for loss-of-separation events if not appropriately managed. Such assignments can balance runway demand and accommodate traffic-flow, runway, weather, or airspace constraints when the nearest runway is not operationally preferred.

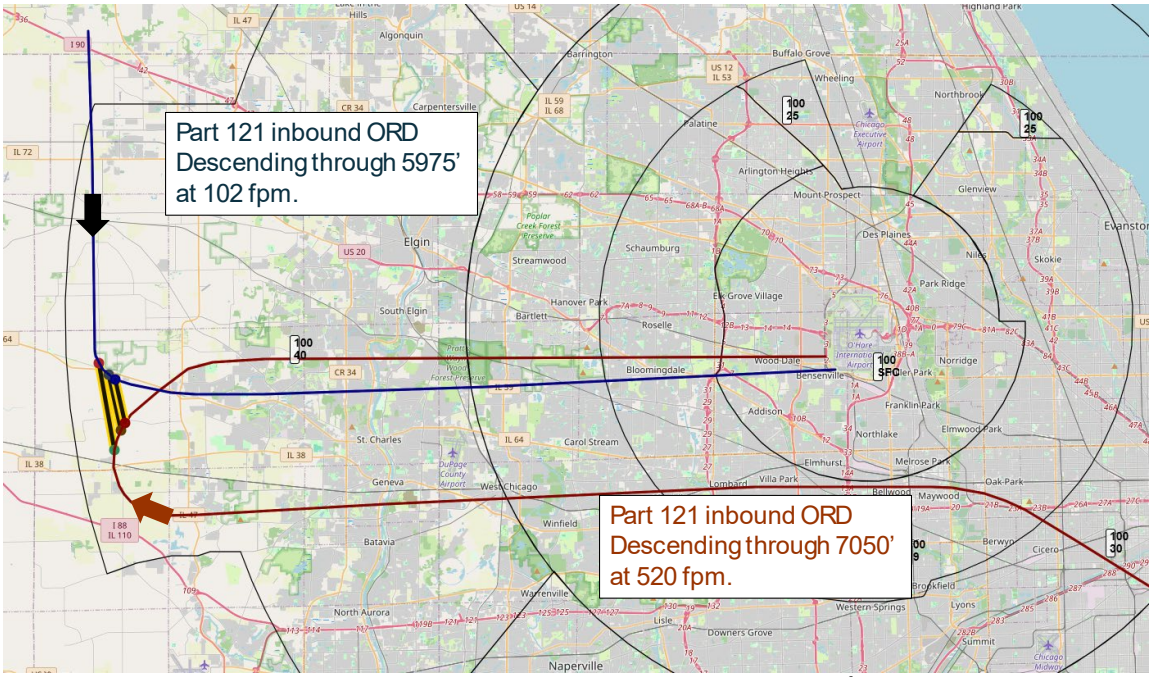


Figure 1. Example of criss-cross arrival for Chicago O'Hare, on 8th of June 2026

Before criss-cross arrival procedures can be deployed, their safety risk must be thoroughly assessed (Federal Aviation Administration 2023). In air traffic management, collision risk is a fundamental safety metric used to determine whether an airspace design, operational procedure, or surveillance system can maintain an acceptably low level of risk and satisfy an established target level of safety (TLS) (Sherry et al. 2023). Airspace collision risk generally refers to the probability that two or more aircraft occupy the same volume of airspace at the same time, potentially resulting in a mid-air collision. A loss of required separation or penetration of a defined proximity volume is commonly treated as a precursor to such an event (Thapa et al. 2023). Depending on the operation and phase of flight, collision risk may be expressed per flight hour, aircraft movement, or exposure event.

Collision-risk assessments consider factors such as traffic density, navigation and surveillance performance, communication reliability, aircraft dimensions, trajectory uncertainty, and operational errors. Some models estimate risk by combining the probability of a separation-related or proximity event with the probability that the remaining protective barriers fail and the event progresses to an actual collision. These estimates provide a quantitative basis for evaluating proposed operations and determining whether separation minima can be safely maintained or modified.

Traditional collision-risk assessments use analytical and probabilistic modeling techniques. Analytical models, including those developed by Reich and Fujita, estimate collision probability by combining traffic-flow characteristics with statistical descriptions of navigation errors and aircraft dimensions (Reich 1966; Fujita 2013). Event-tree and fault-tree methods decompose an accident scenario into sequences of contributing failures and protective-barrier outcomes (Zhang et al. 2015). Monte Carlo simulation methods generate large numbers of simulated traffic encounters to estimate rare-event probabilities, while operation-specific models have been used to quantify airborne collision risk and compare the estimated risk with an established TLS (Zhang et al. 2015; Pilko, Ferraro, and

Scanlan 2023). Importance sampling and sequential Monte Carlo methods have also been developed to improve the computational efficiency of rare-event estimation (de Oliveira et al. 2007; Kharchenko et al. 2006).

A major limitation of simulation-based collision-risk analysis is its dependence on sample size. Under the Rule of Three, when no event is observed in N independent trials, the approximate one-sided 95% upper confidence bound on the event probability is $3/N$ (Jovanovic and Levy 1997). Consequently, demonstrating a collision probability below a TLS of 1E−9 would require approximately 3E9 independent simulated encounters with no observed collisions. Because each encounter consists of two flight trajectories, this would require processing approximately 6E9 trajectories. Even with high performance computing, generating and evaluating this number of terminal area trajectories can be prohibitively time and resource intensive. Rare event methods such as importance sampling and sequential Monte Carlo can reduce this computational burden, although the required sample size remains dependent on the desired estimation accuracy and confidence level (de Oliveira et al. 2007; Kharchenko et al. 2006).

Artificial intelligence and machine learning methods are increasingly being applied to diverse aviation safety and atmospheric monitoring problems. For example, Ashouri and Sherry (2026) developed a lightweight deep-segmentation framework for detecting aircraft induced clouds in ground-based sky imagery to support high altitude icing-condition identification. Recent advances in artificial intelligence and machine learning have enabled the generation of synthetic flight trajectories that preserve important statistical characteristics of observed operations while addressing limitations associated with scarce historical data and computationally expensive trajectory generation (Aref et al. 2024, 2025; Krauth et al. 2023; Murad and Ruocco 2025). Generative methods, including variational autoencoders (VAEs), generative adversarial networks, and related deep learning architectures, learn representations of historical trajectories and generate new flight tracks that were not present in the original dataset. Krauth et al. developed deep generative models for terminal area aircraft trajectories and subsequently combined a disentangled VAE with subset simulation for collision risk estimation (Krauth et al. 2023, 2024). Aref et al. demonstrated single-seed and multi-seed VAE methods for generating synthetic trajectories for collision-risk safety analysis (Aref et al. 2024, 2025). Murad and Ruocco developed a time-based vector quantized VAE and evaluated the statistical fidelity and operational feasibility of its generated trajectories (Murad and Ruocco 2025). These studies indicate that synthetic trajectories can reproduce important spatial, temporal, kinematic, and distributional characteristics of observed operations. Once trained, these models can generate large synthetic samples more efficiently than collecting an equivalent number of historical observations, although validation of trajectory realism and operational feasibility remains essential.

Conventional collision-risk assessments generally emphasize direct proximity or collision risk and may not explicitly represent the latent risk introduced by airspace geometry, planned course changes, or traffic flow procedures (Ghorbani et al. 2026a). Ghorbani et al. also applied statistical process control methods to rare proximity and projected-proximity events to monitor the stability of collision risk indicators and distinguish normal variation in risk measures, due to inherit variability in the system, versus underlying systemic changes (Ghorbani et al. 2026b). These methods complement conventional collision risk analysis by identifying hazardous trajectory interactions that may not be evident from observed collision events alone. Here, latent risk means a projected closest point of approach that enters a specified proximity volume within the look ahead time even when the aircraft are not currently inside that volume.

This paper applies AI/ML-generated synthetic flight trajectories and Monte Carlo simulation to estimate proximity and projected-proximity risks for the criss-cross arrival procedure involving Runways 27C and 28C at Chicago O'Hare International Airport. The analysis evaluates both current and latent collision risk indicators and examines how the relative arrival time offset between the two aircraft streams affects the frequency of close encounters.

The remainder of the paper is organized as follows. Section 2 provides an overview of criss-cross arrivals and relevant collision risk statistics at Chicago O'Hare International Airport. Section 3 describes the trajectory generation, pairing, and collision risk analysis methodology. Section 4 presents the results for arrivals involving Runways 27C and 28C. Section 5 discusses the implications and limitations of the approach and identifies directions for future research.

2- OVERVIEW OF CRISS-CROSS ARRIVALS

Criss-cross arrivals on parallel runways are arrival procedures in which aircraft approaching an airport are assigned to parallel runways from opposite or non-corresponding sides of the terminal airspace. As a result, the arrival streams cross before the aircraft are aligned with their respective final approach courses (Tang 2015). These procedures differ from conventional straight-in parallel approaches because aircraft must transition across the terminal airspace to reach their assigned runways, creating trajectory intersections and additional traffic interactions.

Criss-cross arrivals may be used at high-capacity airports with multiple parallel runways when traffic-flow patterns, runway-utilization constraints, or airspace design make assignment to the nearest runway operationally inefficient. Such arrival assignments can support more balanced runway utilization, improved airport throughput, and increased operational flexibility during periods of high traffic demand. By distributing aircraft among multiple runways, air traffic controllers can manage arrival demand, reduce delays, and accommodate operational constraints. These procedures may also be influenced by weather conditions, noise-abatement requirements, and the integration of arrival and departure flows.

Criss-cross arrivals are typically managed through a combination of arrival sequencing, speed control, altitude restrictions, radar surveillance, and lateral vectoring. Aircraft are strategically spaced and vectored so that crossing points occur with adequate temporal and spatial separation before they become established on their respective final approach paths. These operations may also be supported by performance-based navigation, advanced surveillance systems, and trajectory-based procedures. Consequently, the crossing geometry introduces additional traffic interactions and can increase the potential for separation conflicts if the arrival streams are not appropriately managed.

Analysis of 10 weeks of arrivals to ORD (May 4th to July 12th, 2026) identified a total of 7 interactions between pairs of flights on criss-cross procedures. Figure 2 shows one example pair.

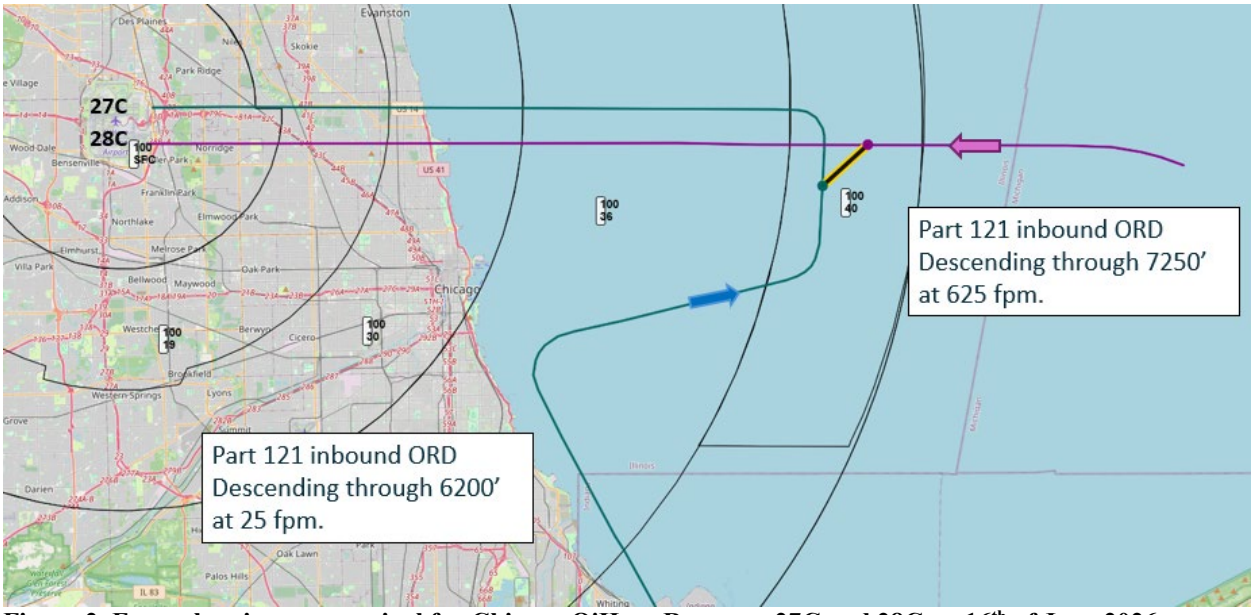


Figure 2. Example criss-cross arrival for Chicago O'Hare Runways 27C and 28C on 16th of June 2026

The observed rate of criss-cross conflicts was 0.1 per day (Figure 3). The total number of events in NMAC airspace volume proximity violations was zero (Table I). The total number of events in TCAS airspace volume proximity violations was 56. The total number of events in projected NMAC airspace volume was zero. The total number of events in projected TCAS airspace volume proximity violations was 318 in 10 weeks.

Proximity NMAC and TCAS cylinders use lateral/vertical thresholds of 500 ft/100 ft and 0.35 NM/850 ft, respectively. A proximity three-dimensional event (sphere) uses Euclidean separation of 500 ft or less. Projected metrics apply the corresponding thresholds to the closest point of approach within 30 seconds.

Table I. Statistics of criss-cross and conflicts for 10 weeks

Event Type	Count in 10 weeks
Criss-cross Count in all flights	7
Proximity NMAC Count in all flights	0
Proximity TCAS Count in all flights	56
Projection NMAC Count in all flights	0
Projection TCAS Count in all flights	318

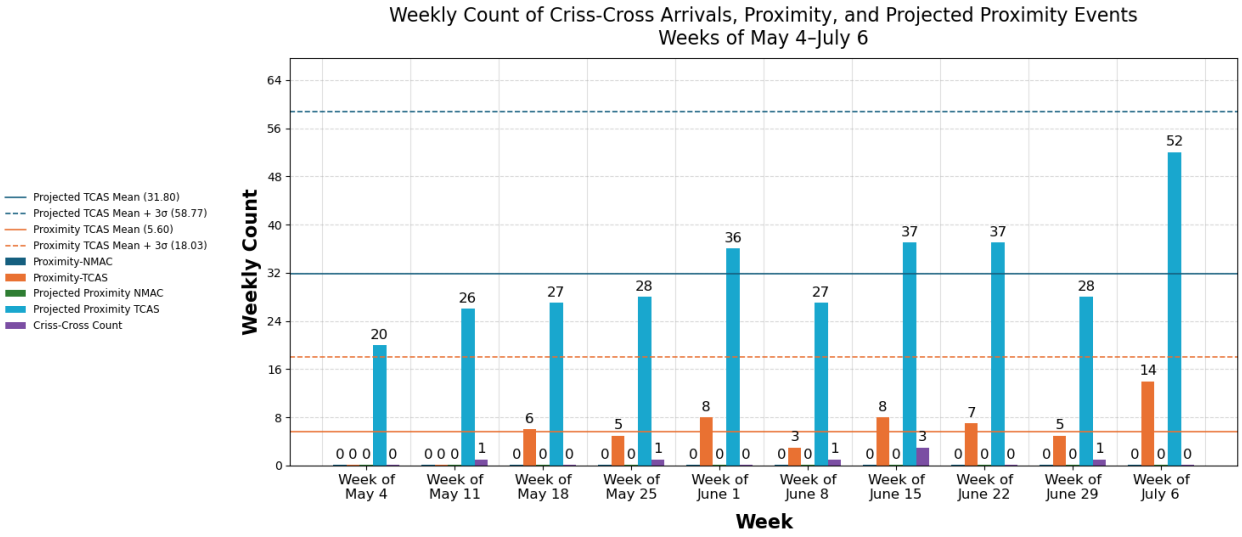


Figure 3. Weekly Counts of Criss-Cross Arrivals, Conflicts, and Proximity and Projected-Proximity Events During the 10-Week Analysis Period

3- METHOD FOR COLLISION RISK ANALYSIS USING SYNTHETIC FLIGHT TRACKS AND LATENT RISK

The method for calculating collision risk and latent collision risk is described through the four-phase process below.

- Phase 1 generates the synthetic ORD Runway 27C flight tracks used in the analysis using the toolset described by Aref et al. (2024, 2025, 2026).
- Phase 2 prepares the historical ORD Runway 28C flight tracks for pairing with the synthetic 27C flight tracks.
- Phase 3 analyzes the pairs of synthetic flight tracks selected and synchronized via a Monte Carlo simulation as described in Shortle and Sherry (2026). The calculation of collision risk and latent collision risk follows the methods described by Ghorbani et al. (2026b).
- Phase 4 calculates event probabilities and confidence intervals for the proximity and projected-proximity metrics, as described by Ghorbani et al. (2026b) and Shortle and Sherry (2026).

1- Phase 1: Generate synthetic ORD 27C flight tracks for analysis.
   a. Obtain historic flight tracks for ORD 27C (e.g., from ADS-B data).
   b. Pre-process historic flight tracks for VAE training, including conversion to a Cartesian coordinate system, standardization, and normalization.

c. Train the VAE on historic flight tracks.
d. Source a representative seed track for ORD Runway 27C, such as a simulator-generated flight track.
e. Preprocess the seed flight track, including standardization and normalization.
f. Generate synthetic flight tracks (e.g. 1,000).
g. Validate characteristics of synthetic flight tracks (e.g. course changes, bank angles, deceleration rates, lateral/vertical/groundspeed dispersion at specific along-track distances from the runway).
h. Postprocess the synthetic flight tracks by converting them back to a geographic coordinate system and assigning a consistent time reference.
i. Interpolate the synthetic flight tracks to a fixed one-second update rate.

2- Phase 2: Prepare historic ORD 28C flight tracks for analysis.
a. Source historical flight tracks for ORD Runway 28C (e.g., from ADS-B data).
b. Filter the historical flight tracks to retain valid Runway 28C arrival trajectories.
c. Convert the historical flight tracks to the same Cartesian coordinate system used for the synthetic Runway 27C trajectories.
d. Standardize the trajectory structure, variables, units, and time format.
e. Validate the characteristics and completeness of the historical flight tracks.
f. Interpolate the historical flight tracks to a fixed one-second update rate.

3- Phase 3: Analyze Pairs of Flight Tracks
a. Time synchronization of flight tracks
i. Establish the ORD Runway 27C threshold as the common arrival-time reference.
ii. Define two symmetric uniform distributions for the Runway 28C arrival-time offset: [−300,300] and [−600,600] seconds. A positive offset indicates that the Runway 28C aircraft arrives later than the Runway 27C aircraft.
iii. Use the [−600,600]-second range as the primary timing envelope because the principal trajectory-interaction region occurs approximately 600 seconds before the Runway 27C threshold on the reference trajectory. This range allows the Runway 28C aircraft to be shifted from substantially earlier to substantially later than the Runway 27C aircraft.
iv. Use the narrower [−300,300]-second range as a sensitivity case to evaluate how the assumed width of the arrival-time-offset distribution affects the estimated event probabilities. Both distributions are centered at zero to avoid imposing an earlier- or later-arrival bias.
b. Conduct a separate Monte Carlo experiment for each time-offset range. For each time-offset range, perform 500 replications, with 1,000 paired-flight samples in each replication.
i. Randomly select one synthetic 27C trajectory.
ii. Randomly select one historical 28C trajectory.
iii. Randomly sample one arrival-time offset from the uniform time offset distribution.
iv. Synchronize both trajectories by their arrival/reference time.
v. Shift 28C relative to 27C using time_28C = time_27C − offset
vi. Calculate the proximity Euclidian distance, lateral distance, and vertical distance at each aircraft present position at each time stamp for the pair of flight track data points.
vii. Apply airspace volume proximity thresholds and count the events (The NMAC proximity volume is defined by a lateral separation of 500 ft and a vertical separation of 100 ft. The TCAS proximity volume is defined by a lateral separation of 0.35 NM and a vertical separation of 850 ft.)
viii. Calculate the projected proximity Closest Point of Approach (CPA) for Euclidian distance, lateral distance, and vertical distance at each aircraft present position given heading and groundspeed at each time stamp for the pair of flight track data points.

ix. Apply airspace volume projected proximity thresholds (e.g. NMAC airspace volume, TCAS airspace volume) for the projected closest point of approach and count the events (The NMAC projected proximity volume is defined by a lateral separation at CPA of 500 ft and a vertical separation at CPA of 100 ft. The TCAS projected proximity volume is defined by a lateral separation at CPA of 0.35 NM and a vertical separation at CPA of 850 ft. Time to CPA for both threshold is defined 30 seconds.)

4- Phase 4: Conduct statistical analysis on count data for each type (i.e. proximity, projected proximity) for each threshold (e.g. NMAC, TCAS).
   a. Calculate each event probabilitiy per pair.
   b. Calculate a confidence interval (using the Poisson distribution for rare events).
   c. Generate tables, histograms, and plots.

4- RESULTS OF COLLISION RISK ANALYSIS OF CRISS-CROSS PROCEDURE

The analysis was conducted according to the methods described above. Five hundred repetitions of one thousand iterations using 1000 flight tracks for arrival ORD 27C and 288 historic flight tracks for arrival ORD 28C (Figures 4).

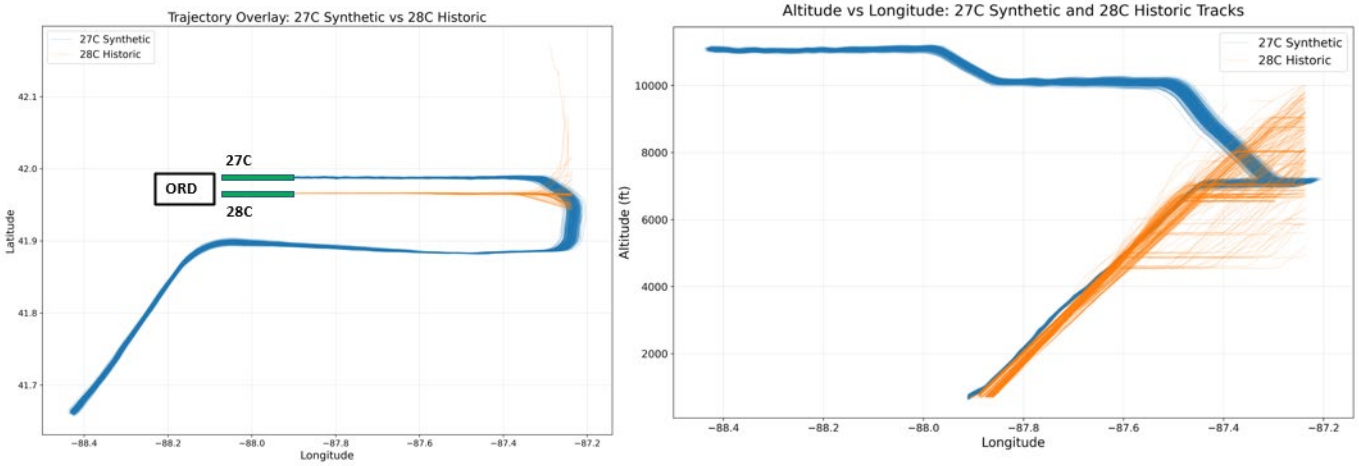


Figure 4. A Top-down View Trajectory Overlay and Altitude vs Longitude Plot, 27 C Synthetic vs 28C Historic Flight Tracks

4.1. Characteristics of the Synthetic Flight Tracks

The characteristics of the 1,000 synthetic Runway 27C flight tracks were evaluated at a common location of 3 NM from the runway reference point, obtained through linear interpolation between the two trajectory points bracketing the evaluation location. At 3 NM, the synthetic trajectories had a mean altitude of 1,778.8 ft MSL with a standard deviation of 19.4 ft. The mean groundspeed was 142.8 kt with a standard deviation of 0.8 kt, while the mean vertical rate was −1,615 ft/min with a standard deviation of 78.7 ft/min. The mean cross-track position was −0.047 NM, with a standard deviation of 0.029 NM. Negative cross-track values indicate positions south of the Runway 27C extended centerline, or the left side when viewed in the landing direction. The trajectories also exhibited a mean heading deviation of −1.47 degrees and an estimated mean bank angle of 1.98 degrees. The small dispersions indicate a tightly clustered synthetic set at 3 NM with measurable spatial and kinematic variation.

Table II. Characteristics of the Synthetic Flight Tracks

Measure	Mean	Standard deviation	5th percentile	Median	95th percentile
Cross-track position	−0.047 NM	0.029 NM	−0.096 NM	−0.047 NM	−0.002 NM
Altitude	1,778.8 ft MSL	19.4 ft	1,745.6 ft	1,779.1 ft	1,811.2 ft
Groundspeed	142.79 kt	0.76 kt	141.63 kt	142.78 kt	144.01 kt
Vertical rate	−1,615 ft/min	78.7 ft/min	−1,744.5 ft/min	−1,616.8 ft/min	−1,485.9 ft/min
Longitudinal acceleration	−0.116 kt/s	0.024 kt/s	−0.155 kt/s	−0.117 kt/s	−0.075 kt/s
Heading deviation	−1.47°	1.00°	−3.13°	−1.46°	0.09°
Estimated bank angle	1.98°	1.06°	0.26°	1.96°	3.70°

4.2. Characteristics of the Historic Flight Tracks

Similarly, the characteristics of the 288 historical Runway 28C flight tracks were evaluated at a common location of 3 NM from the runway reference point. At 3 NM, the historical trajectories had a mean altitude of 2,106.7 ft MSL with a standard deviation of 135.4 ft. The mean groundspeed was 157.3 kt with a standard deviation of 13.1 kt, while the mean vertical rate was −892.5 ft/min with a standard deviation of 130.2 ft/min. The mean cross-track position was 0.005 NM, with a standard deviation of 0.003 NM. The trajectories also exhibited a mean heading deviation of −0.02 degrees and an estimated mean bank angle of 0.21 degrees. Compared with the synthetic Runway 27C tracks, the historical Runway 28C tracks show greater variability in altitude, groundspeed, vertical rate, and acceleration, but less cross-track dispersion.

Table III. Characteristics of the Historic Flight Tracks

Measure	Mean	Standard deviation	5th percentile	Median	95th percentile
Cross-track position	0.005 NM	0.003 NM	0.000 NM	0.004 NM	0.009 NM
Altitude	2,106.7 ft MSL	135.4 ft	1,920.7 ft	2,082.2 ft	2,313.3 ft
Groundspeed	157.25 kt	13.05 kt	135.82 kt	156.96 kt	175.82 kt
Vertical rate	−892.5 ft/min	130.2 ft/min	−1,125.1 ft/min	−876.2 ft/min	−673.8 ft/min
Longitudinal acceleration	−0.624 kt/s	0.332 kt/s	−1.114 kt/s	−0.640 kt/s	0.000 kt/s
Heading deviation	−0.02°	0.33°	−0.57°	0.00°	0.48°
Estimated bank angle	0.21°	0.21°	0.00°	0.19°	0.58°

4.3. Proximity and Projected-Proximity Results

After pairing the two sets of flight trajectories using the proposed simulation framework, the resulting safety-risk measures are presented below.

All probabilities reported in this section assume that the Runway 27C and Runway 28C arrivals are operated independently, without controller coordination or intervention to maintain separation.

- For the [−600, 600]-second offset range, the current NMAC- and TCAS-cylinder probabilities per simulated pair are 1.98E-04 and 2.91E-03, respectively (Table IV).
- The projected NMAC- and TCAS-cylinder probabilities are 4.62E-04 and 3.06E-03. These quantify latent risk: the projected closest point of approach enters the volume within 30 seconds even when current positions do not.
- Current TCAS-cylinder events were 14.68 times as frequent as current NMAC-cylinder events, indicating that most TCAS-volume encounters did not reach the NMAC threshold. This is not a direct collision margin.
- Projected TCAS-cylinder events were 6.61 times as frequent as projected NMAC-cylinder events, indicating that most projected TCAS encounters did not reach the projected NMAC threshold.
- The projected NMAC-cylinder probability was 2.33 times the current NMAC-cylinder probability, indicating more predicted than current NMAC-volume entries.
- The projected TCAS-cylinder probability was 1.05 times the current TCAS-cylinder probability; the two estimates were therefore similar.

Table IV. Summary of Counts and Probabilities for Collision Risk Metrics

Event Type	Counts			Probabilities per Pair		
	Count	Lower CI	Upper CI	Probability	Lower CI	Upper CI
Proximity NMAC Sphere	202	175.10	231.85	4.04E-04	3.50E-04	4.64E-04
Proximity NMAC Cylinder	99	80.461	120.52	1.98E-04	1.61E-04	2.41E-04
Proximity TCAS Cylinder	1,454	1,379.37	1,528.62	2.91E-03	2.76E-03	3.06E-03
Projected NMAC Cylinder	231	202.16	262.78	4.62E-04	4.04E-04	5.26E-04
Projected TCAS Cylinder	1,528	1,451.50	1,604.49	3.06E-03	2.90E-03	3.21E-03

The events were concentrated almost entirely between approximately 100 and 350 seconds, indicating that the critical encounters occurred when the 28C aircraft arrived later than the 27C aircraft (Figure 5). The current and projected TCAS-cylinder events were most frequent in the 160–170-second bin, with 167 and 153 events, respectively. In contrast, the current and projected NMAC-cylinder events reached their highest counts in the 240–250-second bin, with 19 and 35 events, respectively. The current three-dimensional separation events of 500 ft or less peaked in the 120–130-second bin with 26 events. Overall, the results demonstrate that the frequency and severity of close encounters depend strongly on the relative arrival-time offset between the two aircraft streams.

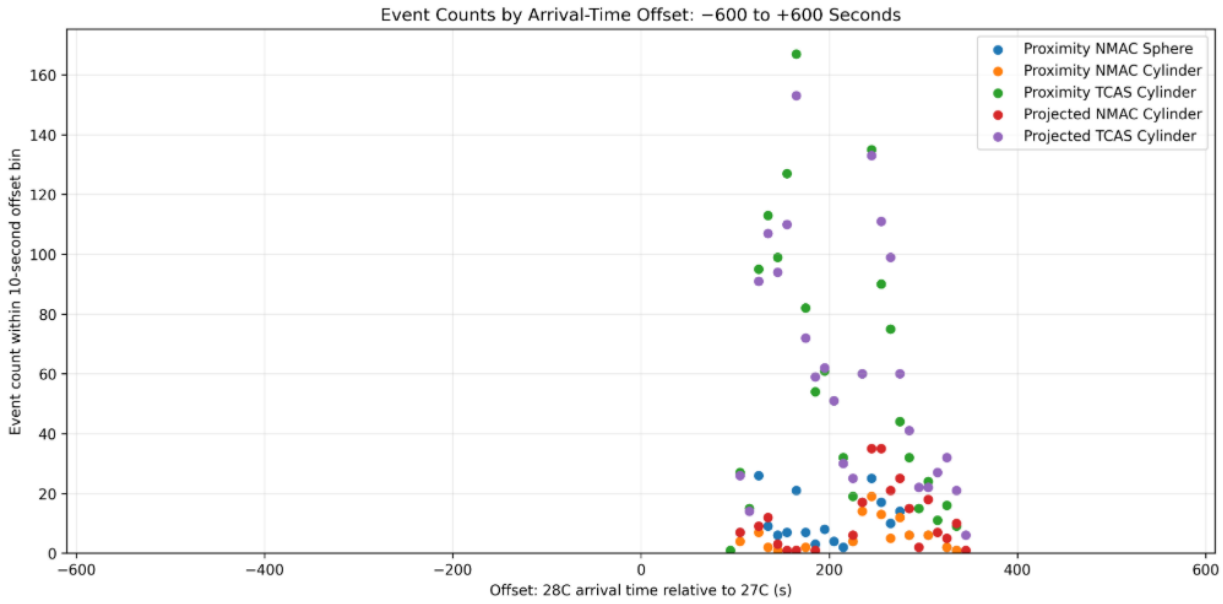


Figure 5. Event count per 10-second offset bin.

Two arrival-time-offset ranges were evaluated: [−300,300] and [−600,600] seconds. The estimated proximity NMAC-cylinder probabilities per simulated pair were 3.06E−4 and 1.98E−4, respectively (Figure 6). The higher probability for the [−300,300]-second scenario occurred because this distribution assigned a larger proportion of the simulated pairs to the critical positive-offset region in which NMAC events were concentrated, particularly around 240–250 seconds. The wider [−600,600]-second scenario included the same critical timing region but assigned more samples to offsets farther from it, where few or no NMAC events occurred. Consequently, the probability averaged over the entire [−600,600]-second distribution was lower. These findings demonstrate that the estimated NMAC probability depends strongly on the assumed arrival-time-offset distribution. The results should therefore not be interpreted as indicating that widening the offset range inherently improves safety; rather, they show that close encounters are concentrated within a limited positive-offset window, approximately 100–350 seconds, while offsets outside this window produced substantially fewer events.

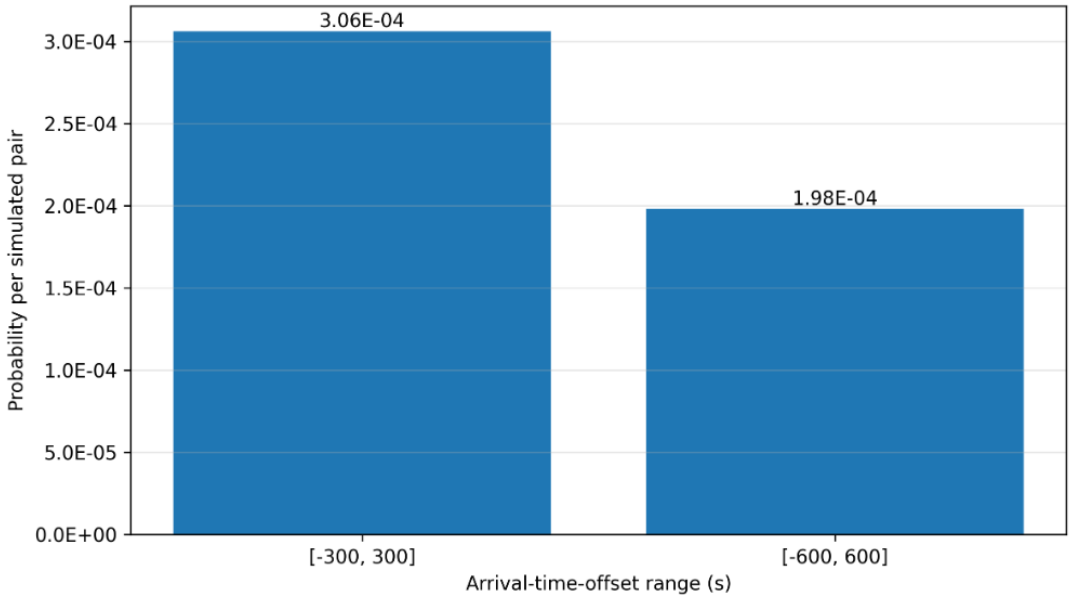


Figure 6. NMAC Cylinder Probability per Pair by Time-Offset Range

5- DISCUSSION

Near mid-air collision events can be used as surrogate indicators of collision risk. Collision probability may be expressed as the product of the probability of an NMAC and the conditional probability that an NMAC results in a collision. Some collision-risk studies assume that approximately one out of every 100 NMAC events result in a collision, corresponding to P (collision | NMAC) = 0.01. Under this assumption, a collision target level of safety of 1E−09 per operation corresponds to a derived NMAC benchmark of 1E−07 per operation. This value is a model-derived benchmark and depends on the assumed NMAC-to-collision conversion factor, encounter definition, aircraft characteristics, and exposure unit (Kochenderfer et al. 2010).

The results reported in this analysis for the NMAC and TCAS airspace volumes should be interpreted as conservative risk estimates due to the assumption of independent pairing and arrival-time-offset distributions. The estimated proximity risk is conservative, based on the assumption that the two arrival streams operate independently. The result does not account for potential coordination between the controllers responsible for the two approaches. Per ICAO guidelines for CRA, the model also does not account for operational interventions intended to maintain safe temporal spacing.

The arrival-time offset strongly affects the NMAC probability. The offset time determines the aircraft pairs to reach the trajectory-interaction region at the same time. For this specific criss-cross procedure, trajectory interaction events were concentrated at positive offsets of approximately 100–350 seconds and peaked at around 240–250 seconds. The positive critical offset reflects the combined effects of trajectory geometry, remaining path length, and the complete speed and descent profiles of both streams. For example, at 3 nm from the runway thresholds, the historical 28C trajectories had a higher mean groundspeed (157 knots) than the synthetic 27C trajectories (142 knots).

The impact of the time offset (Figure 6) shows that [−300,300]-second scenario produced a higher NMAC probability than the [−600,600]-second scenario because a larger proportion of its samples fell within the critical positive-offset region. The wider range included more samples outside this region, which lowered the average probability.

6- CONCLUSIONS

This paper demonstrates the use of AI/ML-generated synthetic flight tracks for collision-risk analysis of airspace navigation procedures. AI/ML algorithms for synthetic flight track generation can now generate a sufficient sample size to reach the required confidence level. This helps to overcome the computational limitations on the number of generated flight tracks.

One thousand synthetic Runway 27C trajectories were paired with historical Runway 28C trajectories through Monte Carlo simulation to evaluate criss-cross arrivals at Chicago O'Hare. The paired trajectories were assessed using current and projected proximity metrics under NMAC and TCAS thresholds.

For the [−600,600]-second arrival time offset range, the estimated probabilities per simulated pair were 1.98E−04 for NMAC proximity, 2.91E−03 for TCAS proximity, 4.62E−04 for projected NMAC, and 3.06E−03 for projected TCAS. The events were concentrated within a limited positive-offset interval, demonstrating the importance of relative timing between the two arrival streams.

The results assume that the Runway 27C and Runway 28C arrival streams operate independently, without controller coordination or intervention to maintain separation. The comparison of the [−300,300]- and [−600,600]-second ranges also showed that the estimated probabilities depend on the assumed arrival time offset distribution.

Overall, the study shows that validated AI/ML-generated trajectories can expand the flight-track data available for simulation and can be integrated with Monte Carlo methods for collision-risk modeling. This approach

provides a practical framework for evaluating complex airspace procedures when historical trajectory data alone are insufficient for repeated safety analysis.

ACKNOWLEDGMENTS

The authors acknowledge the technical contributions of Somil Shah, Rany Azzi, Sherri Shearon, Jiang Hong, Cody Nichols, Lacey Thompson, Luke Wood, and Jacob Powers of the Federal Aviation Administration; and Dr. Ran Ji, Dr. Chen, Ashim Thapa, Erfan Ashouri, Shahab Aref, Andrew Kirk, and Jon West of CATSR, SEOR, and George Mason University.

A variational autoencoder (VAE) was used to generate synthetic flight trajectories analyzed in this study. OpenAI ChatGPT was used to assist with grammar, wording, and clarity. All AI-assisted content and outputs were reviewed and validated by the authors, who take full responsibility for the manuscript.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Sherry, Ghorbani, Shortle; data collection: Ghorbani; analysis and interpretation of results: Ghorbani, Sherry, Shortle; draft manuscript preparation: Ghorbani, Sherry. All authors reviewed the results and approved the final version of the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

This research was funded by the Federal Aviation Administration (FAA) under the NEXTOR IDIQ Contract No.: 693KA9-20-D00004, Delivery Order No. 693KA9-22-F00004
.

REFERENCES

Aref, Shahab, John F. Shortle, and Lance Sherry. 2024. "Generating synthetic flight tracks for collision risk safety analysis: Variational autoencoders with a single seed track." In *2024 Integrated Communications, Navigation and Surveillance Conference (ICNS)*, 1–9. IEEE. https://doi.org/10.1109/ICNS60906.2024.10550543.

Aref, Shahab, John F. Shortle, and Lance Sherry. 2025. "Physics-based 'realism checks' for AI-generated synthetic flight tracks." In *2025 AIAA/IEEE Digital Avionics Systems Conference (DASC)*, 1–10. IEEE. https://doi.org/10.1109/DASC66011.2025.11257394.

Aref, Shahab, John F. Shortle, and Lance Sherry. 2026. "Synthetic flight track generation for collision risk analysis: An iterative multiseed approach." Paper presented at the 23rd Annual Conference on Systems Engineering Research, Arlington, VA, April 6–9, 2026.

Ashouri, Erfan, and Lance Sherry. 2026. "High altitude icing condition identification: A lightweight deep segmentation framework for detecting aircraft-induced clouds in ground-based sky photography." In *Proceedings of the 2026 Integrated Communications, Navigation and Surveillance Conference (ICNS).*

de Oliveira, Ítalo Romani, Paulo Sérgio Cugnasca, Henk A. P. Blom, and G. J. "Bert" Bakker. 2007. "Estimation of collision risk in time-based spacing air traffic by sequential Monte Carlo simulation." Paper presented at the 11th World Conference on Transport Research, Berkeley, CA, June 24–28, 2007.

Federal Aviation Administration. 2023. *Safety Risk Management Policy*. FAA Order 8040.4C. Washington, DC: U.S. Department of Transportation.

Fujita, Masato. 2013. "Collision risk model for independently operated homogeneous air traffic flows in terminal area." *Electronic Navigation Research Institute Papers*, no. 130: 1–17.

Ghorbani, Fahimeh, Lance Sherry, and John F. Shortle. 2026a. "Insights into airspace flight-track latent proximity risk: A case study of Chicago TRACON and adjacent airspace." Paper accepted for presentation at the AIAA/IEEE Digital Avionics Systems Conference, Orlando, FL, September 15–17, 2026.

Ghorbani, Fahimeh, John F. Shortle, and Lance Sherry. 2026b. "Separating the signal from the noise: Statistical process control for steady-state monitoring of airspace collision risk." Paper presented at the 105th Annual Meeting of the Transportation Research Board, Washington, DC, 2026.

Jovanovic, B. D., and Paul S. Levy. 1997. "A look at the rule of three." *The American Statistician* 51 (2): 137–139. https://doi.org/10.1080/00031305.1997.10473947.

Kharchenko, V. P., S. V. Nagaev, A. G. Kukush, E. A. Znakovskaya, and S. I. Dotsenko. 2006. "Determination of sample size in a rare event simulation method." *Cybernetics and Systems Analysis* 42 (1): 65–74. https://doi.org/10.1007/s10559-006-0038-y.

Kochenderfer, M. J., J. D. Griffith, and J. E. Olszta. 2010. "On Estimating Mid-Air Collision Risk." AIAA Aviation Technology, Integration, and Operations Conference. DOI: 10.2514/6.2010-9333.

Krauth, Timothé, Antoine Lafage, Jérôme Morio, Xavier Olive, and Matthias Waltert. 2023. "Deep generative modelling of aircraft trajectories in terminal maneuvering areas." *Machine Learning with Applications* 11: 100446. https://doi.org/10.1016/j.mlwa.2022.100446.

Krauth, Timothé, Jérôme Morio, Xavier Olive, and Benjamin Figuet. 2024. "Advanced collision risk estimation in terminal manoeuvring areas using a disentangled variational autoencoder for uncertainty quantification." *Engineering Applications of Artificial Intelligence* 133: 108137. https://doi.org/10.1016/j.engappai.2024.108137.

Murad, Abdulmajid, and Massimiliano Ruocco. 2025. “Synthetic aircraft trajectory generation using time-based VQ-VAE.” In *2025 Integrated Communications, Navigation and Surveillance Conference (ICNS)*. Brussels, Belgium: IEEE. https://doi.org/10.1109/ICNS65417.2025.10976929.

Pilko, Aliaksei, Mario Ferraro, and James Scanlan. 2023. “Quantifying specific operation airborne collision risk through Monte Carlo simulation.” *Aerospace* 10 (7): 593. https://doi.org/10.3390/aerospace10070593.

Reich, Peter G. 1966. “Analysis of long-range air traffic systems: Separation standards—III.” *The Journal of Navigation* 19 (3): 331–347. https://doi.org/10.1017/S0373463300047445.

Sherry, Lance, John F. Shortle, Alexia Payan, Evan Harrison, Ashim Kumar Thapa, Alberto Cardenas Melgar, and Yohan Auguste. 2023. *Review of Current State of Artificial Intelligence/Machine Learning and Other Advanced Techniques Related to Air-to-Air Collision Risk Models in the Terminal Airspace*. Report DOT/FAA/TC-23/37. Washington, DC: Federal Aviation Administration. https://doi.org/10.21949/1528218.

Shortle, John F, Lance Sherry, 2026 .“Methodology for Collision Risk Assessment Combining AI-Generated and Simulated Flight Tracks”.

Tang, Huabin. 2015. “Conflict alerts for aircraft conducting visual approaches.” In *15th AIAA Aviation Technology, Integration, and Operations Conference*. Dallas, TX: American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2015-2277.

Thapa, Ashim Kumar, John F. Shortle, and Lance Sherry. 2023. “Air-to-air collision risk models in the terminal airspace.” In *2023 Integrated Communications, Navigation and Surveillance Conference (ICNS)*. IEEE. https://doi.org/10.1109/ICNS58246.2023.10124323.

Zhang, Yimin, John F. Shortle, and Lance Sherry. 2015. “Methodology for collision risk assessment of an airspace flow corridor concept.” *Reliability Engineering & System Safety* 142: 444–455. https://doi.org/10.1016/j.ress.2015.05.015.