Project 054 AEDT Evaluation and Development Support

Georgia Institute of Technology

Project Lead Investigator

Prof. Dimitri N. Mavris
Director, Aerospace Systems Design Laboratory
School of Aerospace Engineering
Georgia Institute of Technology
275 Ferst Drive NW, Atlanta, GA 30332
(404) 894-1557
dimitri.mavris@ae.gatech.edu

Dr. Michelle Kirby
Chief, Civil Aviation Division
Aerospace Systems Design Laboratory
School of Aerospace Engineering
Georgia Institute of Technology
275 Ferst Drive NW, Atlanta, GA 30332
(404) 385-2780
michelle.kirby@ae.gatech.edu

University Participants

Georgia Institute of Technology (Georgia Tech)

- P.I.s: Dr. Dimitri N. Mavris, Dr. Michelle Kirby, Dr. Ameya Behere
- FAA Award Numbers: 13-C-AJFE-GIT-098, 114, and 122
- Period of Performance: September 24, 2021, to September 20, 2023
- Tasks:
  1. Improved departure modeling
  2. Arrival profile modeling
  3. Full flight modeling
  4. System testing and evaluation of the Aviation Environmental Design Tool (AEDT)
  5. Supersonic transport (SST) aircraft modeling in the AEDT (formerly ASCENT Project 010)

Project Funding Level

Georgia Tech has received $900,000 in funding for this project. In terms of cost-share details, Georgia Tech has agreed to a total of $900,000 in matching funds. This total includes salaries for the project director, research engineers, and graduate research assistants, as well as computing, financial, and administrative support, including meeting arrangements. Georgia Tech has also agreed to provide tuition remission for the students, paid for by state funds.

Investigation Team

Prof. Dimitri N. Mavris, (P.I.), Tasks 1-5
Dr. Michelle Kirby, (co-Investigator), Tasks 1-5
Dr. Ameya Behere, (co-Investigator), Tasks 1-5
Mr. Jirat Bhanpato, (research faculty), Tasks 1-5
Dr. Raphael Gautier, (research faculty), Tasks 1-5
Ranya Almarzooqi, (graduate student), Task 1
Vishnu Sankar Manivasakan, (graduate student), Task 1
Keletso Mmlane, (graduate student), Task 1
Anushka Moharir, (graduate student), Task 2
Hansen Lian, (graduate student), Task 2

Nitya Maruthuvakudi Venkatram, (graduate student), Tasks 3 and 5
Hyungu Choi, (graduate student), Task 3
Yash Vinod, (graduate student), Task 3
Bogdan Dorca, (graduate student), Task 4
Santusht Sairam, (graduate student), Task 4
Fahraan Badruddin (graduate student), Task 4

Project Overview

This project provides data and methods to continue to improve modeling of aircraft weight, takeoff thrust, and departure and arrival procedures within the Federal Aviation Administration's (FAA) AEDT, as well as the AEDT's full flight modeling capabilities. Some of the modeling assumptions in the AEDT are considered excessively conservative and could be improved through use of industry and airport flight operation data. Funding for this project will continue to support the implementation of these methods and data in the AEDT. To facilitate these efforts, the Georgia Tech team will use real-world flight data and noise monitoring data to improve departure, full flight, and arrival modeling.

Task 1 – Improved Departure Modeling

Georgia Institute of Technology

Objective

This task aims to improve departure profile modeling in the AEDT. Over the past few years, this task provided recommendations for new noise abatement departure procedures (NADP) to be implemented in the AEDT, improving how flight trajectories are modeled and consequently, giving more accurate fuel consumption and emissions metric results.

In the previous reporting period, departure profile modeling efforts focused on identifying why existing AEDT departure procedures did not adequately represent observed operations in the 2021 Threaded Track (TT) and Flight Operations Quality Assurance (FOQA) calendar year (CY)2023 datasets. That work examined the role of aircraft energy share percentage (ES%) and takeoff weight in departures, classified key operational "anchors" for characteristics such as flap transitions, landing gear retraction, thrust cutbacks, level-offs, and constant calibrated airspeed (CAS) segments. Dynamic time warping (DTW) was used to cluster CY2023 FOQA departures using these anchors to determine distinct departure procedures for AEDT implementation.

Building on that foundation, this year's objective is to evaluate whether real-world departure operations can be represented in the AEDT using the existing AEDT procedures and previously developed NADP profiles when assessed on noise. Specifically, the task (1) clusters FOQA departures to existing AEDT/NADP profiles based on noise under flight path (NUFP) similarity, and (2) validates these cluster assignments by checking that key operational anchors for thrust cutback and acceleration segments as well as fuel burn over the corresponding segments, agree between the FOQA derived fixed point profiles (FPP) and their matched procedures within defined tolerances.

To present these findings, this report first describes the methodology used to construct NUFP from AEDT noise outputs for FPPs and existing profiles, to assign FPPs to existing procedures based on NUFP through clustering, and to extract and compare anchors and fuel burn for all flights in these clusters. It then summarizes the clustering and validation results for Airbus® A321-232 departures at several airports, highlights where existing profiles capture noise and operations well and where they do not, and concludes with planned next steps for refining the anchor detection and determining whether additional procedures are necessary to be defined in the AEDT.

Background and Previous Work

Prior work established that, for the TT-2021 dataset, none of the existing AEDT departure profiles provided a consistently accurate representation of observed operations, even when accounting for ES% and takeoff weight variations. The previous reporting period then used CY2023 FOQA data to (1) identify and classify operational anchors for departures (flaps, landing gears, thrust cutback, level-offs, constant CAS), (2) build FPPs in the AEDT to acquire performance and noise results for FOQA departures, and (3) apply DTW and related metrics to cluster flights by trajectory and thrust profiles, and to compare these clusters to existing AEDT and NADP procedures.

® Airbus is a registered trademark of Airbus Operations GmbH, Hamburg, Germany.

Initial clustering highlighted that trajectory-based similarity alone is not always aligned with noise patterns. Some flights with visibly different thrust and altitude profiles produced similar noise footprints, while others with very slight trajectory differences produced significantly different noise. This motivated the shift toward a noise-centric clustering method using NUFP as the basis for determining whether real-world operations can be represented by existing departure procedures.

Research Approach

The following section provides a detailed explanation of the methodology used in this reporting period to evaluate whether real-world departure operations can be represented in the AEDT using existing AEDT and NADP procedures. The overall process is summarized in Figure 1, which shows how FOQA data are used to build FPPs in the AEDT, obtain NUFP, and classify key operational anchors. NUFP for each FPP is then compared with the existing procedure library to determine the closest match, after which the assigned cluster is validated using anchor alignment and cumulative fuel burn. Flights that satisfy these validation checks are considered to be appropriately represented by their assigned procedures, while flights that do not indicate potentially the need for additional or modified departure procedures. The subsections that follow describe each step of this process: NUFP construction and clustering as well as anchor and fuel-based validation.

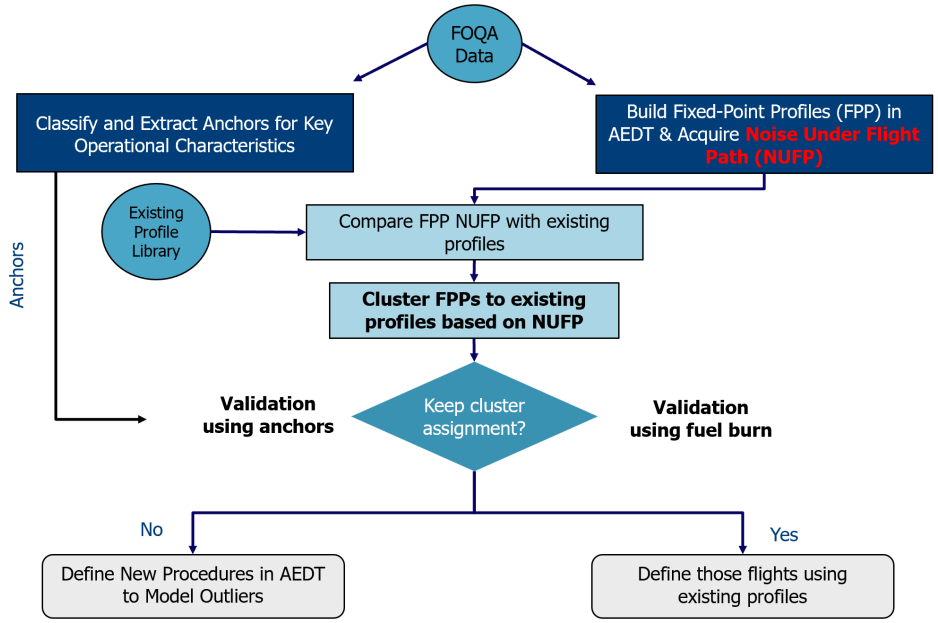


Figure 1. Overall clustering and validation process. AEDT: Aviation Environmental Design Tool, FOQA: Flight Operations Quality Assurance, NUFP: noise under flight path.

Methodology

Data and Scope

This year's analysis uses FOQA CY2023 departures, focusing on A321-232 operations from multiple airports. The NUFP-based clustering and validation were applied to A321-232 departures from Denver International Airport (KDEN), Hartsfield-Jackson Atlanta International Airport (KATL), and Los Angeles International Airport (KLAX). There were 1,987 A321-232 departures considered for KATL, 441 A321-232 departures considered for KDEN, and 532 A321-232 departures considered for KLAX.

Each of these departures is modeled in the AEDT through creating FPP that follow the observed trajectory. The AEDT is then used to compute noise on its standard receptor grid. These FPPs are compared to a library of existing AEDT departure procedures (STANDARD, International Civil Aviation Organization [ICAO]A, ICAOB, etc.) and NADP profiles (NADP 1 and 2 profiles with multiple thrust derate settings) to determine which procedure best matches each FOQA departure in terms of NUFP.

Construction of NUFP

NUFP is a one-dimensional representation of sound exposure level (SEL) directly underneath the aircraft's ground track. For each FPP, the AEDT is first run to obtain full SEL grids on the receptor layout for that airport. The aircraft ground track distance versus time, as well as the latitude and longitude over time, is taken from AEDT performance reports. A KDTree search is then used to find receptors close to the ground track; receptors within roughly 0.08 nmi of the track (about 1.5 times the receptor spacing) are kept ensuring that at least one receptor lies near each point along the trajectory.

To focus on the part of the noise footprint most relevant for NUFP clustering, the receptor set is filtered by SEL threshold. Several thresholds were tested (75, 85, and 95 dB); however, these tests showed that the 95 dB captures only the loudest part of the footprint and misses smaller differences that matter for clustering, while 75 dB retains enough of the footprint shape to separate different procedures without being dominated by far-field noise. As a result, the 75 dB NUFP was used as the primary basis for clustering.

Receptors at or above 75 dB SEL are ordered along cumulative ground track distance from the start of takeoff roll. This produces SEL as a function of cumulative distance from liftoff for each flight. For comparison with existing AEDT/NADP procedures, the NUFP curves for the FPP and the procedures are truncated at the point where either profile reaches 15,000 ft above field elevation (AFE), so that both curves cover a common downrange segment.

NUFP Similarity and Assignment to Existing Procedures

Initial work compared NUFP using distribution-based metrics such as Wasserstein distance and Hausdorff distance applied to the cumulative SEL distributions. These methods capture some differences between procedures but also produced cases where flights with clearly different thrust cutback heights or climb gradients looked “similar” in terms of NUFP because the distributions overlapped. This showed that NUFP needed to be compared as a curve along distance, not just a set of values.

The current approach, therefore, compares the full SEL versus cumulative distance curves. For a given FPP and AEDT/NADP procedure, the two NUFP curves are interpolated to a common set of cumulative distance points over their shared range up to 15,000 ft AFE. The mean absolute error (MAE) and root mean square error (RMSE) between the two curves are then computed. MAE captures average differences in SEL along the path, while RMSE emphasizes larger local mismatches. These two error measures are normalized and combined into a single similarity score. For each FPP, this score is computed against every AEDT/NADP procedure, and the procedure with the lowest score is taken as the best NUFP match for the FOQA departure. All flights whose NUFP best matches a given procedure are grouped into a cluster labeled by that procedure.

This assignment is done separately for each airport and airframe. Comparing results across KDEN, KATL, and KLAX allows for analyzing how local runway layout, elevation, and typical operating practices affect which procedures appear in the cluster library and how many flights each procedure represents.

Anchor Classification

NUFP similarity alone does not guarantee that two departures follow the same procedure in terms of thrust and speed changes, the NUFP-based cluster assignments are checked using operational anchors for the departures. All anchors are defined along cumulative ground track distance but checks also use altitude at the anchor to ensure that events occur at reasonable heights.

Thrust cutback begin (CUT) is defined as the earliest point after an altitude gate where the smoothed thrust profile shows a clear drop and then levels off. Specifically, the algorithm looks for the earliest point between 800 ft and 4,000 ft AFE where thrust drops by at least 8% over a 0.5 nmi window, followed by a flat segment with less than about 2% thrust decrease over 0.3 nmi. This identifies both the cutback and the start of the reduced thrust climb segment.

Initial acceleration begin (AB) and end (AE) are derived from true airspeed after cutback. The method computes the change in speed over a window a roughly 0.5 nmi. The begin point is classified as the first sustained increase in speed of at least 10 kt over 0.4 nmi, and the end point is where the speed gain drops back to about 2 kt over a similar distance. This captures the main acceleration phase where flap settings are changed and the aircraft transitions toward a climb schedule with higher speed. Final acceleration begin (FAB) and end are defined similarly but occurs after AE, with lower speed thresholds (about 5 kt over 0.3 nmi for FAB begin and 2 kt over 0.3 for final acceleration end) to reflect the smaller accelerations at higher altitude.

These anchors are extracted both for each FPP and for each AEDT/NADP procedure to provide a consistent set of reference points for comparison. For anchor validation, altitude checks are used to ensure that anchors occur at comparable heights: CUT is compared using a 1,500 ft altitude tolerance, and AB/FAB and AE use tolerances of about 1,600 ft and 2,200 ft, respectively.

Validation Using Anchors and Fuel Burn

Once flights have been assigned to procedures based on NUFP and the anchors have been classified, each flight-procedure pair is checked against several validation rules. The aim is to determine whether the NUFP similarity corresponds to a genuinely similar procedure in terms of its thrust, speed, and trajectory.

For each anchor (CUT, AB, AE, FAB), the difference in location between the FPP and the matched procedure is computed, along with the altitude difference. An anchor is considered aligned if it occurs within altitude tolerance and within a profile-specific distance or timing tolerance. In addition to individual anchors, an anchor order rule is applied: AE must occur after both CUT and AB, and FAB must occur after or around AE, which allows for both discrete final acceleration classification and continuous final acceleration segments.

Fuel burn is checked using cumulative fuel burn from the performance reports for the FPPs and matched procedure. Cumulative fuel is evaluated at cutback, and the fuel used over the acceleration windows (AB to AE and AE to AB) is summed. The FPP and procedure values are compared using a symmetric relative error method. A flight is considered to pass the fuel check if the combined error across cutback and the acceleration windows is within 25%. Previous work showed that, for the majority of flights, cumulative fuel burn over these windows is within this bound when the operations are consistent.

FAB is treated somewhat differently from the other anchors. As final acceleration can appear as either a discrete event or continuous increase in speed, FAB alignment is a soft check. FAB must respect the anchor order rule, but mismatches in exact FAB location are not counted as hard failures.

When these validation rules were first applied with fixed thresholds, some profiles had high outlier rates even when visual inspection suggested that the operations were reasonably similar. To address this, thresholds for anchor alignment were tuned per profile. For each profile, the distribution of anchor errors among the flights mapped to that profile is computed, and the profile-specific threshold for that anchor is set to the 80th percentile of the empirical error distribution, with an upper cap of 1.6 times the original base threshold. This keeps thresholds tied to observed variability while preventing them from becoming so large that they no longer enforce a meaningful match.

Results and Discussions

NUFP Clustering at KATL

The NUFP clustering analysis for 1,987 A321-232 departures from KATL reveals a structured and highly non-uniform distribution of operational departure profiles. A total of 38 distinct clusters were identified as shown in Figure 2. Although the absolute number of clusters appears large, the cluster size distribution is strongly skewed: only a small subset of clusters contains the majority of the 1,987 flights, while the remainder form a long tail of clusters with very few flights.

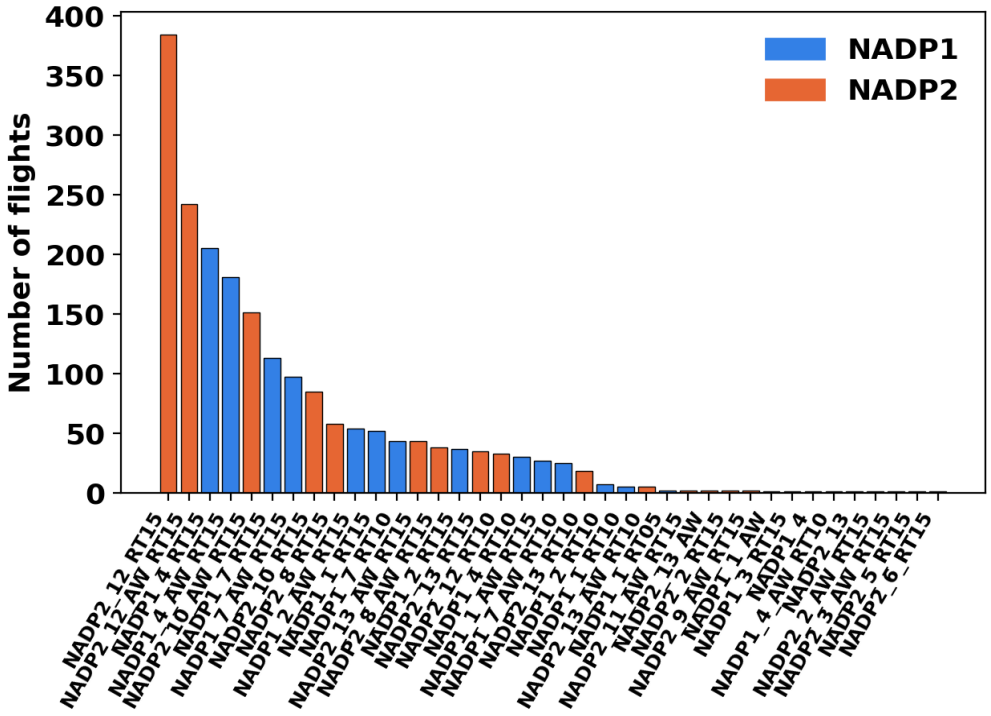


Figure 2. KATL flight distribution among clusters. NADP: noise abatement departure procedures.

All clusters mapped to NADPs**,** specifically NADP1 and NADP2 variants with RT10 and RT15 thrust-reduction settings. These same procedures repeatedly emerged as the closest NUFP matches because their SEL-distance curves reflect the general shape and noise levels seen in the NUFP of real-world operations.

Despite the relatively large number of total clusters (38), the overall distribution reveals a small number of dominant profiles that account for the majority of departures. These dominant clusters correspond to NADP 1_7 and NADP 2_12. Overall, a small set of NADP1 and NADP2 RT variants effectively capture the dominant NUFP behavior at KATL, while the clusters with very few flights reflect infrequent or non-routine noise patterns,

NUFP Clustering at KLAX

For KLAX, NUFP clustering was applied to 532 A321-232 departures, resulting in 66 clusters as shown in Figure 3. Similar to KATL, the cluster size distribution is dominated by a few high-membership clusters, with the remainder forming a long tail of small clusters.

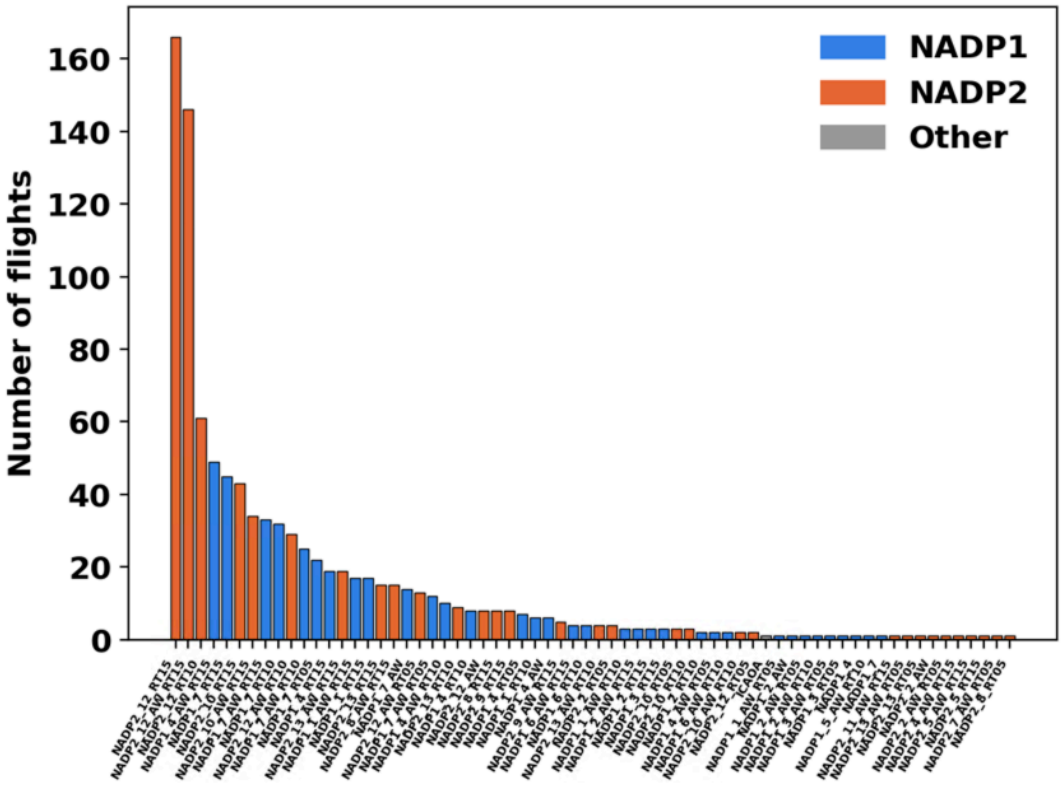


Figure 3. KLAX flight distribution among clusters. NADP: noise abatement departure procedures.

Most flights mapped to NADP2 RT15 and RT10 procedures. NADP 1 variants appear but represent a smaller share of flights compared to KATL. This is likely due to differences in airport elevation and location, impacting noise-abatement constraints and air traffic control procedures. A single flight matched to an ICAOA procedure; however, this does not seem to be a common operation for A321-232s taking off from KLAX.

A321-232s taking off from KLAX produced a larger tail of small clusters than KATL. Non-RT profiles appear within this long tail of clusters, but account for relatively few flights. However, the key observation remains the same: the most common NUFP distributions correspond to a narrow group of NADP procedures.

NUFP Clustering at KDEN

The NUFP clustering for 441 A321-232 departures from KDEN resulted in 43 distinct clusters Figure 4. As at the other airports in this study, the cluster size distribution is highly skewed (i.e., a small number of profiles account for most flights, while most profiles contain only a few flights). When profiles are sorted by membership, the six largest clusters each contain more than 20 flights and together represent about 61% of all departures. An intermediate set of 11 clusters (6–20 flights each) accounts for roughly 26% of the data, while the remaining 26 clusters each contain five or fewer flights, together representing about 12% of departures. This long tail indicates that KDEN A321-232 operations are dominated by a handful of NADP procedures, namely the NADP2_12 and NADP1_7 variants with RT10 and RT15.

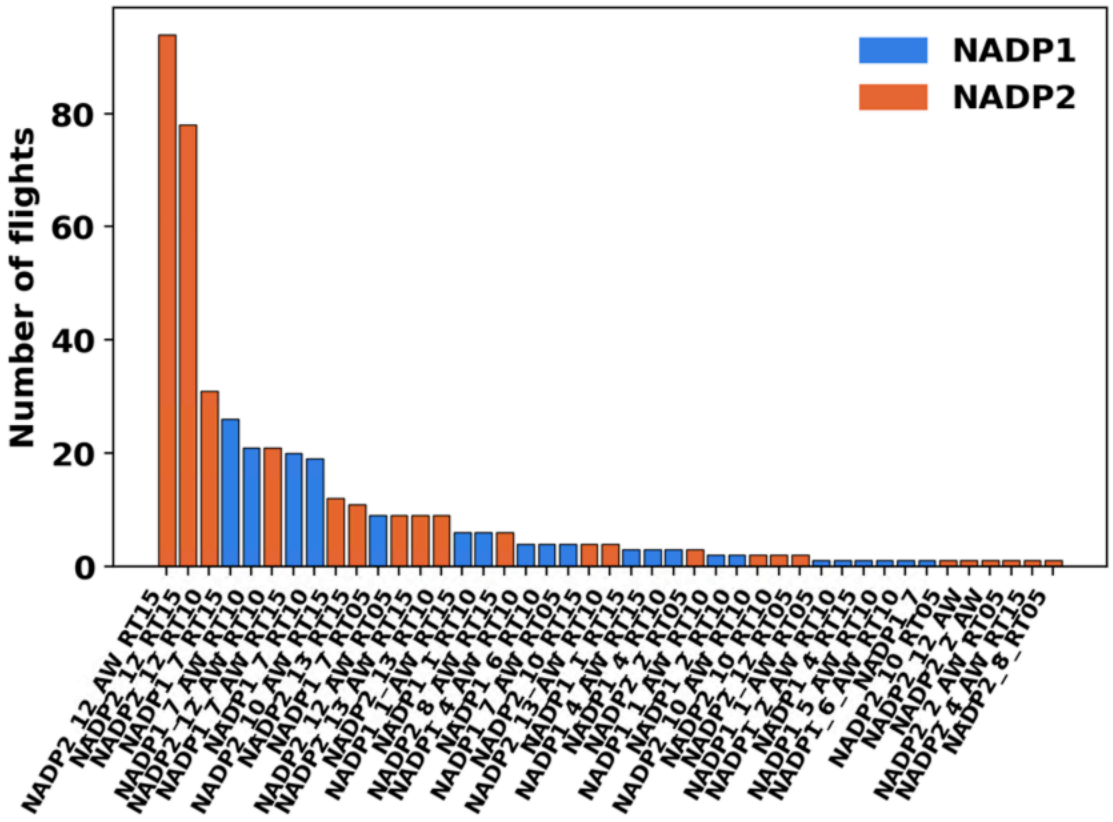


Figure 4. KDEN flight distribution among clusters. NADP: noise abatement departure procedures.

All KDEN profiles were NADP1 or NADP2 variants, aggregating across profiles, approximately 69% of flights are assigned to NADP2 procedures and 31% to NADP1, indicating a clear predominance of NADP2 operations in the NUFP matching results.

The NADP2_12_AW_RT15, NADP2_12_RT15, and NADP2_12_RT10 clusters alone account for just over 45% of all A321-232 departures at KDEN. Smaller but still substantial contributions come from several NADP1_7 variants (e.g., NADP1_7_RT15, NADP1_7_AW_RT10, NADP1_7_AW_RT15, NADP1_7_RT10), each representing 15–30 flights. Overall, the KDEN clustering results indicate that A321-232 departures are strongly organized around NADP2 profiles, with NADP1 profiles providing a secondary set of procedures.

NUFO Clustering Results

Across all three airports, the NUFP clustering results show that A321-232 departures are overwhelmingly organized around a compact set of NADP1 and NADP2 procedures with different RT settings, while many additional clusters contain only a few flights each. At KATL, 1,987 departures form 38 clusters that all match NADP-style profiles; at KLAX, 532 departures form 66 clusters, with almost all clusters mapped to NADP1/NADP2 and a single flight matching an ICAO A procedure; and at KDEN, 441 departures form 43 clusters, with every flight assigned to a NADP variant and about 69% of flights mapped to NADP2. Thus, existing AEDT procedures appear only marginally in the NUFP matching (one KLAX case), whereas NADP profiles systematically capture NUFP for real-world operations more accurately. This indicates that a relatively small library of NADP-based profiles, including key NADP1/NADP2 RT variants, can effectively represent observed A321-232 departure behavior across the three airports and provides a focused basis for enhancing AEDT departure modeling.

Validation Results for A321-232 Departures for KATL

Following NUFP-based clustering, each flight underwent anchor-level validation to assess how closely individual FOQA departures aligned with their cluster's procedural characteristics. Out of 1,987 flights, 378 were flagged as outliers,

corresponding to a validation rate of approximately 81%. This indicates that most flights followed consistent anchor sequencing and climb-profile shapes.

Anchor-Level Validation:

- Cutback (CUT) has the highest pass rate (97.6%), indicating that the thrust-reduction event occurs at consistently similar locations across flights.
- Acceleration anchors show more variability, with AB having the lowest pass rate (79.7%), while AE is somewhat more consistent (84.9%).
- Anchor-order (Order) checks pass for 87.3% of flights, indicating that the expected sequence (CUT → AB → AE) is followed in most cases.
- Fuel-based validation (Fuel) shows that 90% of flights remain within ±25% of the procedure's cumulative fuel burn over the acceleration windows, suggesting good overall agreement between FOQA FPPs and their assigned NADP procedures.

Validation Results for A321-232 Departures for KLAX
524 A321-232 departures were validated, 99 were classified as outliers (≈ 18.9%).

Anchor-Level Validation:

- CUT has a pass rate of 97.9%.
- Followed by AB: 93.5%
- 88.4% have no AE failure
- Final Acceleration Begin (FAB): 100% (i.e., no flight has FAB listed as a failure)
- The order anchor has a pass rate of 98.9%.
- The fuel anchor has a 96.9% pass rate.

Validation Results for A321-232 Departures for KDEN
Following cluster assignment, each of the 441 KDEN departures underwent anchor-level and fuel-based validation against its matched NUFP profile. Of these flights, 79 were flagged as outliers, corresponding to an outlier rate of approximately 17.9%. The remaining 362 flights (82.1%) passed all validation checks and are considered well represented by their assigned NUFP procedures.

Anchor-level pass rates are summarized below, computed over all 441 flights:

- CUT a pass rate of 98.4%.
- AB a pass rate of 92.3%.
- AE a pass rate of 88.9%.
- Order a pass rate of 99.1%.
- Fuel a pass rate of 99.1%.

The validation results suggest the following patterns for KDEN A321-232 departures:

- Cutback and anchor order are highly consistent. Almost all flights satisfy the cutback and order criteria, indicating that the thrust-reduction event and the overall temporal sequence of cutback → acceleration are very stable features across the dataset.
- Acceleration anchors are the primary source of deviation. Most validation failures involve AB and/or AE, which is consistent with those anchors being more sensitive to intra-procedure variability.
- Fuel-based checks show strong agreement. Only four flights fail the fuel check, indicating that cumulative fuel burn over the cutback and acceleration windows generally agrees well with the NUFP-assigned procedures.

In summary, the KDEN validation results show that most A321-232 departures are well captured by their NUFP-assigned NADP1 or NADP2 profiles, with a moderate outlier fraction driven primarily by variability in acceleration-phase anchors rather than by cutback or fuel behavior. These findings complement the clustering results by demonstrating that the dominant NUFP procedures at KDEN are not only frequently used but also flown with generally high internal consistency.

Conclusion
For all three airports, the NUFP validation results indicate that the assigned procedures capture A321-232 departure behavior with a high degree of internal consistency, with roughly 80% of flights passing all checks at each airport (81% at

KATL, 81.1% at KLAX, 82.1% at KDEN). Cutback and anchor-order tests are the most robust, with pass rates above about 97% everywhere, showing that the thrust-reduction event and the overall CUT → AB → AE sequence are implemented very consistently in practice. Acceleration anchors are the main source of deviations: AB and AE have the lowest pass rates at each airport, especially AB at KATL, confirming that the acceleration phase is where most intra-procedure variability appears. Fuel-based validation is strong at all three airports, with at least 90% of flights within the specified fuel-burn bounds and pass rates near 97–99% at KLAX and KDEN. Taken together, these results show that the NUFP-assigned NADP1/NADP2 profiles not only dominate procedurally but are also flown in a repeatable way, with a relatively small outlier fraction driven primarily by variations in the acceleration segment rather than by gross differences in cutback location, event ordering, or fuel usage.

Outlier Behavior and Directional Sensitivity at KATL

This section summarizes the validation outlier behavior observed for A321-232 departures. The goal is to distinguish true procedural deviations from artifacts of the current anchor-detection and validation logic.

Of the 1,987 flights at KATL, 378 grouped into 38 NUFP-based clusters (approximately 19%) failed at least one validation check and were classified as outliers. To understand where validation is most sensitive, the 378 outlier flights were decomposed by the specific anchor checks that failed, shown in Figure 5. The anchor-combination histogram shows that approximately 87% of all outlier flags are tied to the acceleration phase anchors. AE-only failures account for about 42% of the outliers, AB-only failures contribute another 41%, and combined AB+AE failures add roughly 3–4%. This concentration in AB and AE confirms that most outliers arise during the transition between cutback and climb acceleration rather than at thrust reduction itself.

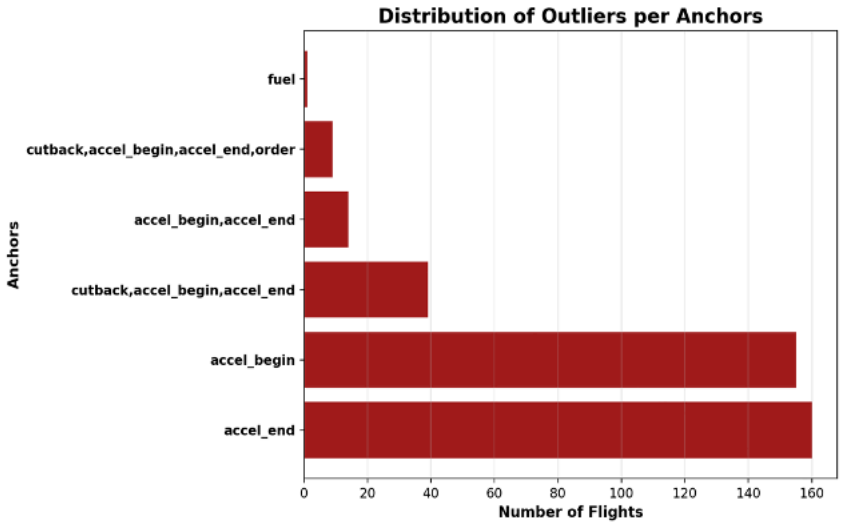


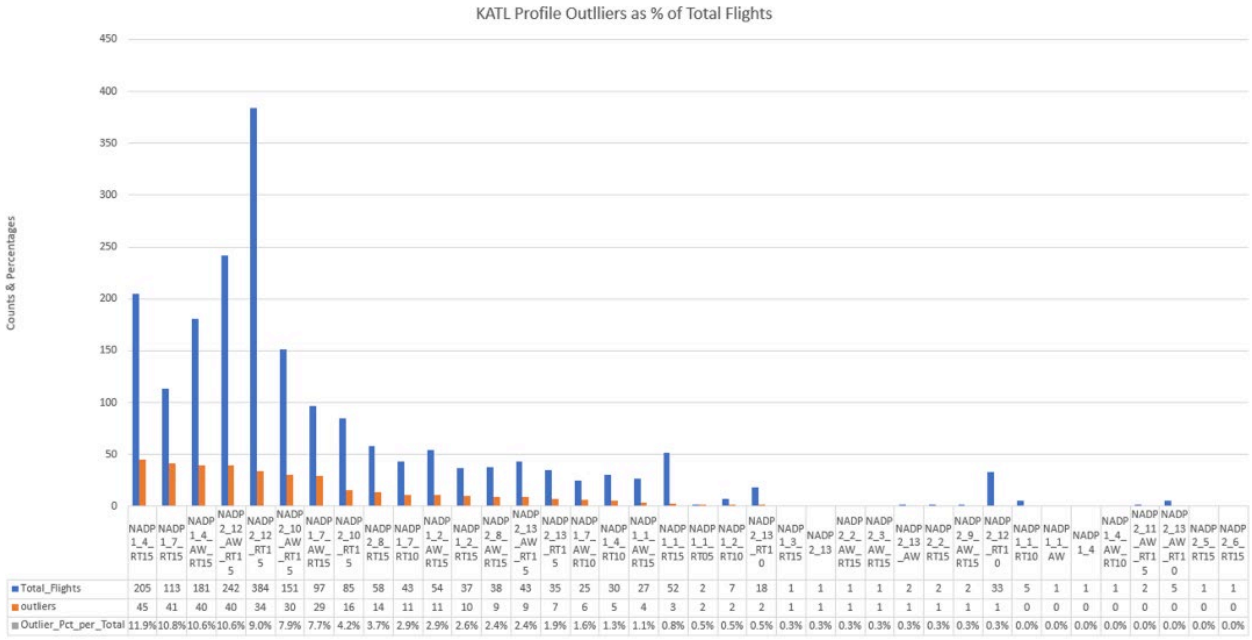
Figure 5. Match failures specific to anchor checks.

Secondary combinations of anchors account for the remaining outliers. Roughly 10.3% of outliers involve simultaneous mismatches in CUT, AB, and AE, while about 2.4% show additional order-related discrepancies (CUT + AB + AE + Order). These cases typically reflect compound effects where slight misplacement of cutback interacts with acceleration detection and ordering, rather than fundamentally different departure procedures. Fuel-driven mismatches are negligible: only about 0.26% of outliers are caused by the fuel check, consistent with the broader finding that 90% of all flights are within ± 25% of expected fuel burn. This supports the interpretation that the performance-based validation remains robust and that the main tuning need lies in acceleration-phase anchor thresholds.

Aggregating outliers by matched NUFP profile shows in Figure 6 that approximately 70% of all outliers are associated with just seven profiles. All these profiles are flown with RT15 thrust-reduction settings and are largely NADP1 variants, with a smaller but still significant contribution from NADP2 profiles. The key high-outlier profiles are:

- NADP1_4_RT15

- NADP1_7_RT15
- NADP1_4_AW_RT15
- NADP2_12_RT15
- NADP2_12_AW_RT15
- NADP2_10_AW_RT15
- NADP1_10_AW_RT15



	NADP 1_4_ RT15	NADP 1_7_ RT15	NADP 1_4_ AW_ RT15	NADP 2_12 _AW _RT1 5	NADP 2_12 _RT1 5	NADP 2_10 _AW _RT1 5	NADP 1_7_ AW_ RT15	NADP 2_10 _RT1 5	NADP 2_8_ RT15	NADP 1_7_ RT10	NADP 1_2_ AW_ RT15	NADP 1_2_ RT15	NADP 2_8_ AW_ RT15	NADP 2_13 _AW _RT1 5	NADP 2_13 _RT1 5	NADP 1_7_ AW_ RT10	NADP 1_4_ RT10	NADP 1_1_ AW_ RT15	NADP 1_1_ RT15	NADP 1_1_ RT05	NADP 1_2_ RT10	NADP 2_13 _RT1 0	NADP 1_3_ RT15	NADP 2_13	NADP 2_2_ AW_ RT15	NADP 2_3_ AW_ RT15	NADP 2_13 _AW	NADP 2_2_ RT15	NADP 2_9_ AW_ RT15	NADP 2_12 _RT1 0	NADP 1_1_ RT10	NADP 1_1_ AW	NADP 1_4	NADP 1_4_ AW_ RT10	NADP 2_11 _AW _RT1 5	NADP 2_13 _AW _RT1 0	NADP 2_5_ RT15	NADP 2_6_ RT15
Total_Flights	205	113	181	242	384	151	97	85	58	43	54	37	38	43	35	25	30	27	52	2	7	18	1	1	1	1	2	2	2	33	5	1	1	1	2	5	1	1
outliers	45	41	40	40	34	30	29	16	14	11	11	10	9	9	7	6	5	4	3	2	2	2	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Outlier_Pct_per_Total	11.9%	10.8%	10.6%	10.6%	9.0%	7.9%	7.7%	4.2%	3.7%	2.9%	2.9%	2.6%	2.4%	2.4%	1.9%	1.6%	1.3%	1.1%	0.8%	0.5%	0.5%	0.5%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 6. Outlier flights matched to NUFP profiles and shown as a percentage of total flights at Hartsfield-Jackson Atlanta International Airport (KATL). NADP: noise abatement departure procedures.

Taken together, these profiles capture most outlier flights. At this stage, the analysis treats the association between high RT settings, NADP1 procedures, and outliers as a correlation rather than a proven causal link; some of these profiles also carry substantial traffic, which naturally increases their outlier counts. Further normalization and profiling are planned before drawing stronger conclusions about causality.

A second analysis examined the percentage of outliers within each profile Figure 7. Profiles with very small sample sizes (1–2 flights) sometimes show outlier rates above 50% or even 100%, but these represent only nine outliers in total (≈2.4% of all outliers) and are treated as case-by-case anomalies rather than evidence of systemic issues. For statistical robustness, the evaluation was restricted to profiles with at least seven flights. Within this subset, the highest outlier rates are again associated with NADP1 procedures:

- NADP1_7_RT15 (~36%)
- NADP1_7AW_RT15 (~30%)
- NADP1_2_RT10 (~29%)
- NADP1_2_RT15 (~27%)
- NADP1_7_RT10 (~26%)
- NADP1_4_RT15 (~22%)

NADP1 combinations with RT10 or RT15 continue to dominate the high-outlier segment.

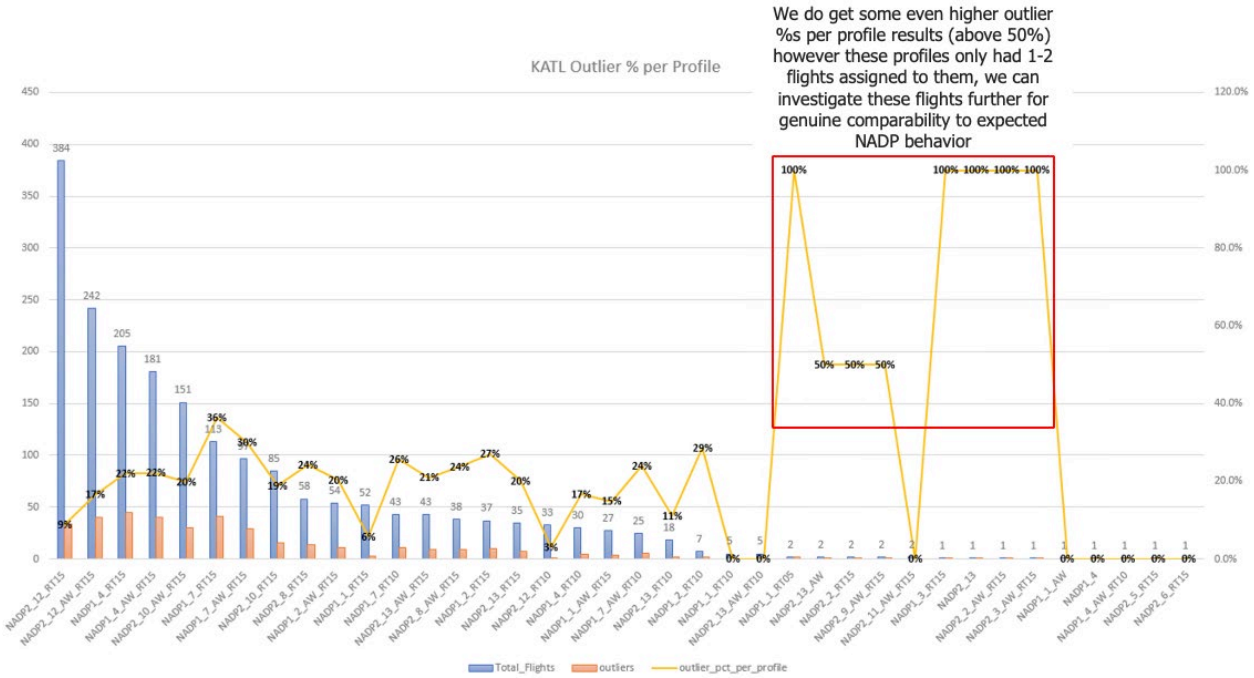


Figure 7. Hartsfield-Jackson Atlanta International Airport (KATL) percentage of outliers matched to each profile. NADP: noise abatement departure procedures.

Detailed cluster plots were generated for representative high-outlier profiles, including NADP1_4_RT15, NADP1_7_AW_RT15 (shown in Figure 8 and Figure 9), NADP1_7_RT10, and NADP1_7_RT15 (shown in Figure 10 and Figure 11). In each case, speed- and altitude-versus-distance traces for outlier flights are overlaid against the corresponding reference NADP profile. Visual inspection shows that the overall climb shapes of the outlier flights remain closely aligned with the nominal procedure: cutback occurs in the expected region, the acceleration segment transitions into clean climb, and the general trajectory family is preserved.

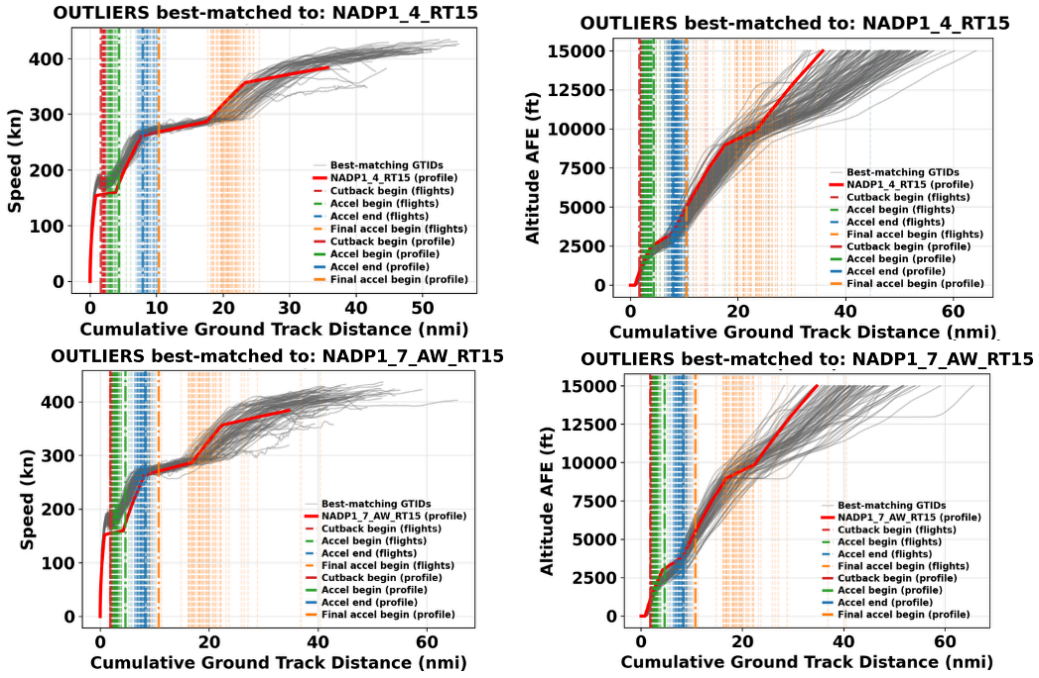


Figure 8. Cluster examples - NADP1_4_RT15 and NADP1_7_AW_RT15 (speed and altitude vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

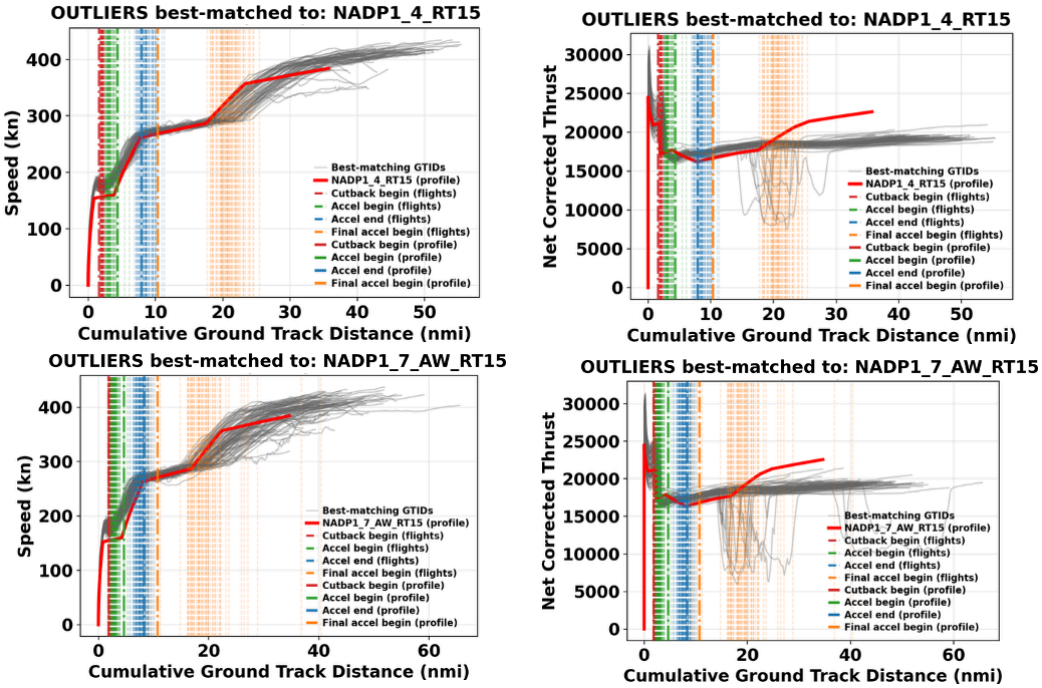


Figure 9. Cluster examples - NADP1_4_RT15 and NADP1_7_AW_RT15 (speed and net corrected thrust vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

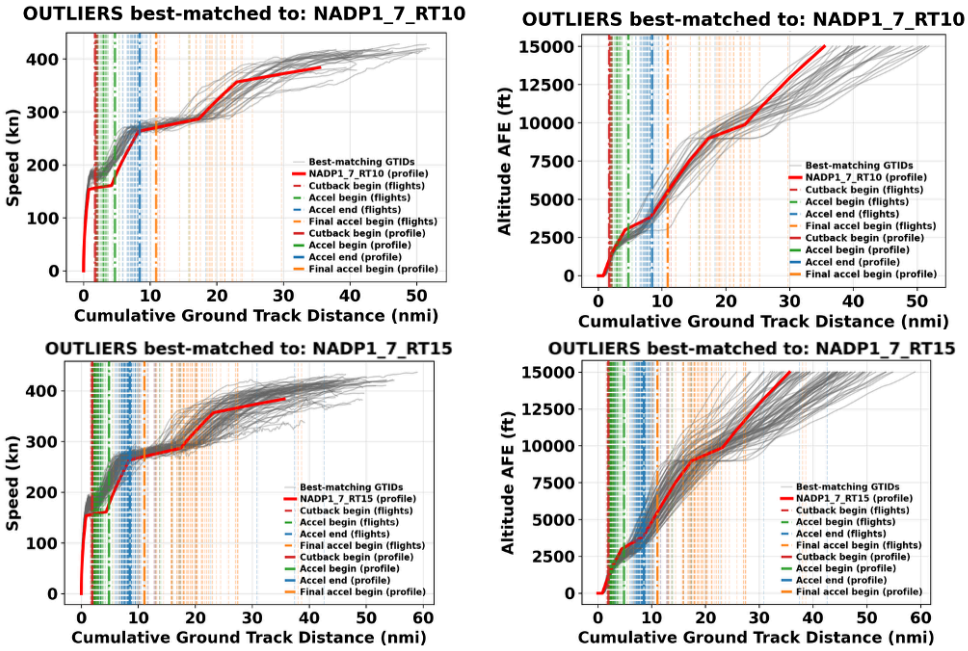


Figure 10. Cluster examples - NADP1_7_RT10 and NADP1_7_RT15 (speed and altitude vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

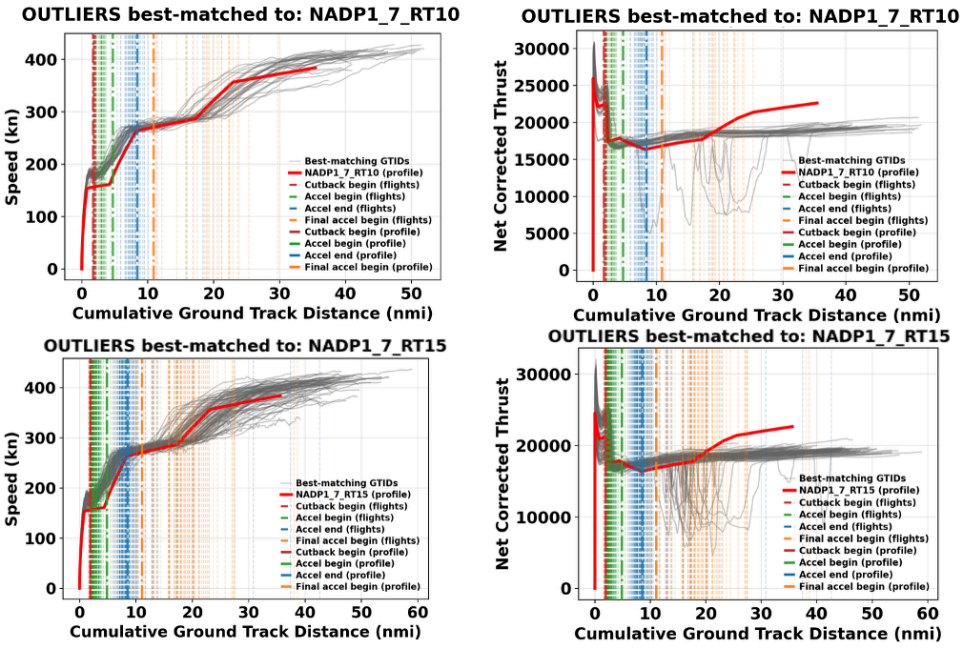


Figure 11. Cluster examples - NADP1_7_RT10 and NADP1_7_RT15 (speed and net corrected thrust vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

The primary differences lie in the exact altitude and distance at which AB and AE are detected. Outlier flights tend to show modest shifts in these anchors relative to the reference profile, even though the broader pattern is consistent. FAB markers, included for visualization, further highlight that much of the late-phase variation is driven by FOQA data noise, sampling resolution, and small differences in thrust derate or weight, rather than fundamentally different procedures.

Net corrected thrust plots for the same clusters reinforce this interpretation. The thrust traces follow the same qualitative pattern as the reference NADP curves, but changes in derate level (e.g., RT10 versus RT15) modify climb performance enough to shift the altitude of the acceleration phase. Under the current validation tolerances, these shifts are sufficient to trigger AB and AE failures, even when the operational intent and overall NADP structure are preserved.

For clusters with the very highest outlier rates (100% failures, such as NADP1_1_RT05 and NADP1_3_RT15, as shown in Figure 12 and Figure 13), all primary anchors fail simultaneously: cutback begin, AB, and AE. Yet the speed and altitude traces again closely follow the NADP1 reference shapes. In these cases, the thrust decrease at “cutback” is often smaller than the 8% threshold used in the current logic, so the algorithm never registers a formal cutback event. Combined with derate-driven changes in acceleration timing and the presence of level-offs, this leads to systematic anchor misdetections and multi-anchor failures, even for flights that are procedurally consistent.

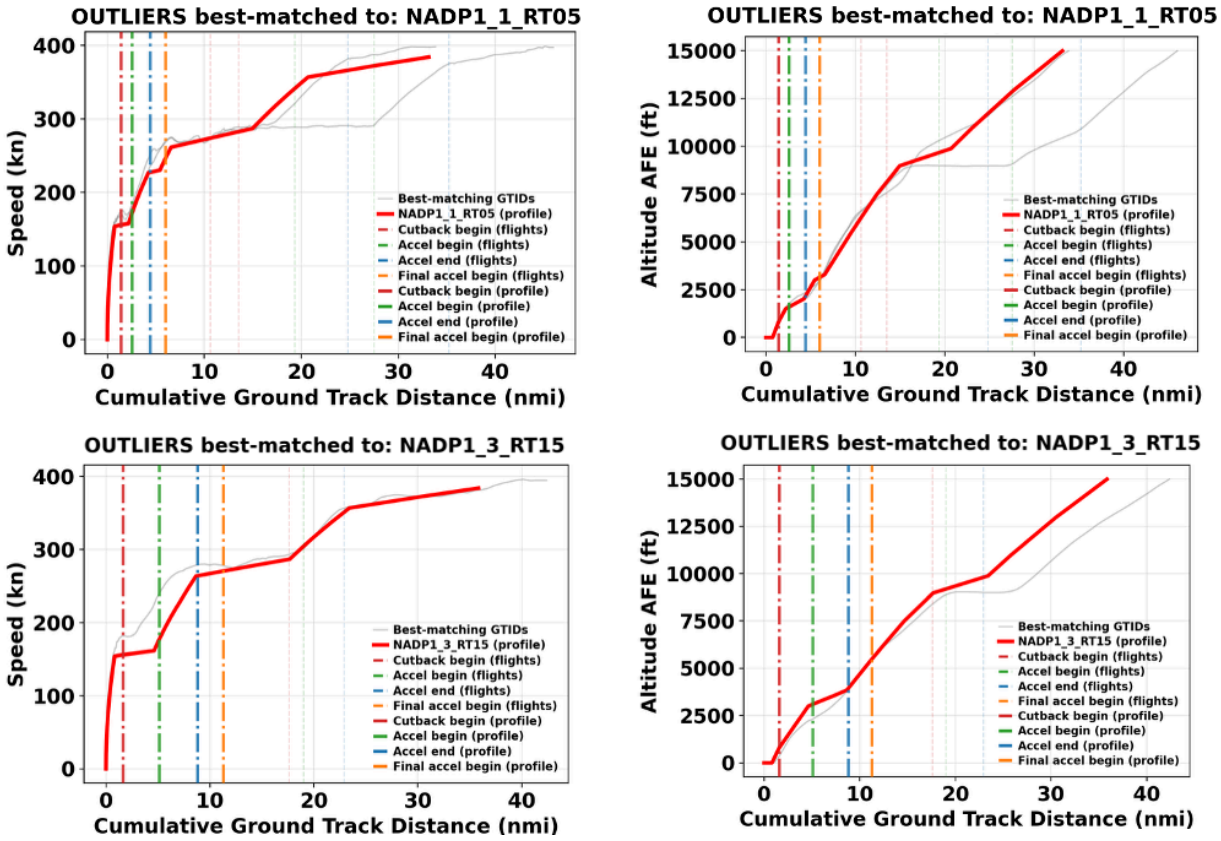


Figure 12. Cluster examples - NADP1_1_RT05 and NADP1_1_RT05 (speed and altitude vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

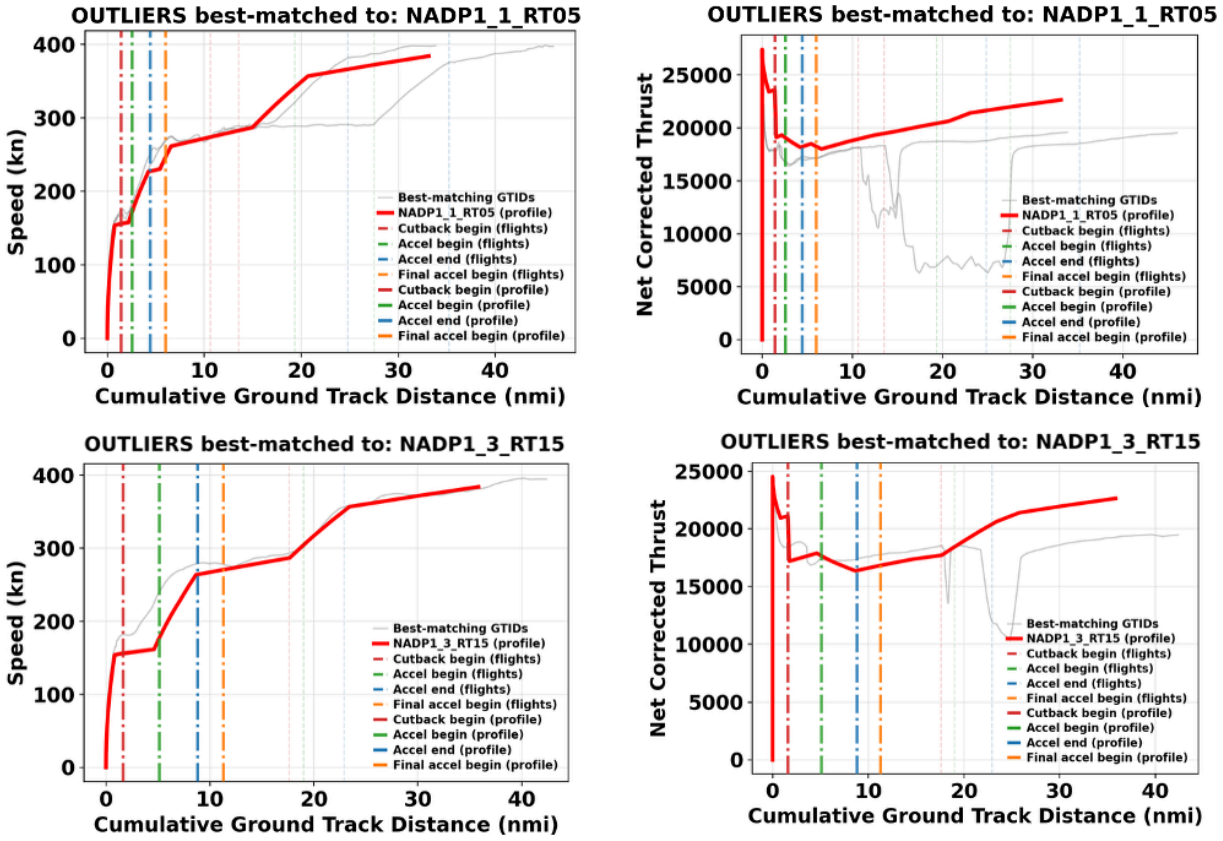


Figure 13. Cluster examples - NADP1_3_RT15 and NADP1_3_RT15 (speed and net corrected thrust vs. cumulative ground tract distance [CGTD] plots). NADP: noise abatement departure procedures.

To place the observed KATL level-off effects in context, a broader analysis was conducted across the full CY 2023 FOQA dataset (approximately 49,647 departures and 596 city pairs). Only about 5% of flights contain at least one level-off segment in the departure phase, indicating that level-offs are relatively rare in the overall population. The top 10 origins account for roughly 74% of these level-offs, driven largely by high-traffic hubs such as KATL, KDTW, and KSEA, with additional contributions from airports like Fort Lauderdale-Hollywood International Airport (KFLL), KDEN, Cancun International Airport (MMUN), KSLC, KLAX, and KJFK. These concentration patterns are interpreted primarily as volume effects, and there is no evidence of a standardized, fleet-wide “level-off procedure,” as summarized in Table 1.

Table 1. Departure level-off analysis summary. ICAO: International Civil Aviation Organization. KATL: Hartsfield-Jackson Atlanta International Airport, KDEN: Denver international Airport, KDTW: Detroit Metropolitan Wayne County Airport, KFLL: Fort Lauderdale-Hollywood International Airport, KJFK: John F. Kennedy International Airport, KLAX: Los Angeles International Airport, KMSP: Minneapolis-Saint Paul International Airport, KSEA: Seattle-Tacoma International Airport, KSLC: Salt Lake City International Airport, MMUN: Cancun International Airport.

Takeoff Airport ICAO Code	Flights with Level-Offs	Total Flights from Takeoff Airport	Rates
KATL	838	10,324	9.7%
KDTW	363	2,559	14.2%
KSEA	144	2,353	6.1%
KFLL	105	497	21.1%
KDEN	70	803	8.7%
KMSP	66	3,998	1.7%
MMUN	65	246	26.4%
KSLC	63	2,465	2.6%
KLAX	58	1,880	3.1%
KJFK	55	1,716	3.2%

In the context of validation, however, even these infrequent level-offs can have a disproportionate impact. Level-offs shift the altitude and distance at which acceleration begin and end are detected, causing anchors to fall outside the prescribed tolerance windows and generating outlier flags. This is especially relevant for profiles where level-offs occur later in the departure leg, compounding the effects of thrust derate and data noise on acceleration detection. Thus, there is a need to remove flights with level-offs from the analysis before clustering and validation is done.

Milestones

None.

Major Accomplishments

- Completed NUFP-based clustering of all FOQA A321-232 departures to the existing AEDT procedure/NADP library, producing cluster distributions for each airport and identifying the dominant procedures used in real-world operations.
- Extracted and classified anchors for thrust cutback, initial AB/AE, and final acceleration for all flights to support validation of the noise-based cluster assignments.
- Performed anchor and fuel-burn validation for each defined cluster, comparing the altitudes at which the anchors occur as well as the cumulative fuel burn at anchor locations between FPPs and their matched AEDT/NADP procedures to determine whether the NUFP-based assignments reflect operationally consistent behavior.
- Identified departures that cannot be represented by any existing AEDT/NADP procedure, based on anchor and fuel-burn mismatches, providing the basis for determining where new procedures may be required.
- Documented level-off behavior across the three airports, quantifying how often level-offs occur and assessing whether there is a need to define new procedures to account for them.

Publications

None.

Outreach Efforts

Participated in biweekly calls and biannual ASCENT meetings.

Awards

None.

Student Involvement

Graduate research assistants Ranya Almarzooqi, Vishnu Sankar Manivasakan, and Keletso Mmalane from the School of Aerospace Engineering participated in this task. They were involved with the generation of noise results, evaluation of clustering methods, and analysis of results.

Plans for Next Period

- Further analyze outliers for KDEN and KLAX to see if the same patterns that were seen for KATL occur.
- Study whether the outliers are occurring due to aircraft weights not being accurately represented in procedures (may need more AW variants) or if more thrust-derate settings are necessary as well (e.g., RT20, RT25, etc.).
- Revise validation thresholds to ensure only flights that require a new procedure definition are classified as outliers.
- Finalize analyses and develop recommendations for how additional departure profile could be modeled to enhance aviation environmental analysis with the AEDT.

Task 2 – Arrival Profile Modeling

Georgia Institute of Technology

Objectives

Currently, the AEDT implements arrival profiles using either specified fixed-point trajectories or manufacturer-provided procedures. The primary objective of Task 2 is to compare data from real-world flights to AEDT models to make recommendations that will allow the AEDT to better capture aircraft behaviors during arrival.

This year, the specific focuses of this task are (1) to develop and validate quantitative metrics to compare AEDT noise outputs with FOQA-derived fixed-point profiles, and (2) to investigate how variations in thrust, level-off height, and flap deployment affect the resulting noise contours. The end goal of this project is to provide recommendations regarding a subset of standardized AEDT arrival profiles which should be integrated into the system and what information those profiles should include.

Methodology

In the previous year, an evaluation of arrival trajectories was conducted to identify key level-off characteristics and quantify flap variations along descent paths. This analysis laid the foundation for the next phase of the study, which focused on understanding how configuration-driven variations such as flap deployment timing, thrust settings, and level-off altitudes translate into acoustic impacts in the AEDT. This year's work extends the investigation from trajectory and configuration modeling to noise outcome validation. Specifically, the study quantifies how well AEDT's standard arrival profiles capture real-world noise behavior observed in FOQA data and develops metrics to objectively measure these deviations.

For generating the FPPs, a large-scale analysis was performed using FOQA arrival data for A321-200 operations into KATL, Runway 26R, covering CY2023. A total of 3,030 FOQA arrival flights were processed through the AEDT version 3g Batch Report Run Tool to generate SEL noise contours. To handle this dataset efficiently, the jobs were divided into 13 chunks (12×234 flights and 1×222 flights), achieving roughly one-minute runtime per flight on the Batch Report Tool. This configuration improved throughput compared to prior single-job runs (~20 minutes/flight), enabling a comprehensive set of FOQA-based AEDT outputs for comparing metrics with the standard profile (see Table 2).

Table 2. Batch Tool Execution Summary.

No. of Jobs	~Time per Job (mins)	~Total Time
3,030 (Actual size)	18-20	38 days
1,000	7	5 days
500	1	8 hours
250	0.5 -0.75	2.5 hours

Quantitative Metrics

To assess the agreement between modeled and observed arrivals, several spatial similarity metrics were implemented:

- **Average Nearest Neighbor (ANN)**: Mean distance between the FOQA flight's noise contour and the standard (STD) AEDT contour along the ground track; a smaller score correlates to a closer match.
- **Hausdorff Distance**: Maximum contour separation; this metric emphasizes worst-case deviation. Because maximum noise deviation generally occurs over the ground track, reducing the dimensionality of the noise contours to the 1D ground track does not lose information for this metric.
- **Intersection over Union (IoU)**: Ratio of shared footprint area to total combined area (0–1 scale); as opposed to the first two, a higher score indicates a better overlap between two noise contours.
- **Percent Area Difference**: Relative footprint expansion or contraction versus the STD contour, with smaller values in magnitude corresponding to a closer match.
- **ΔSEL (FPP – STD)**: Pointwise difference in SEL values across the AEDT receptor grid, revealing localized regions of increased or decreased noise.

When selecting noise contours, the 75 dB SEL threshold was chosen as the lowest noise level of interest. A distribution of the Hausdorff-1D values used in this analysis is shown in Figure 14. Similarly, the distribution of ANN-1D values across all FOQA arrivals is shown in Figure 15. Together, the ANN and Hausdorff metrics form the backbone of this year's analysis. ANN capturing average conformity, and Hausdorff identifying localized anomalies such as isolated "thrust/noise islands." The ANN thresholds used for binning (0–5, 5–10, 10–20, and 20+ ft) were selected based on clear inflection points observed in the nonlinear relationship between ANN and Hausdorff distance. These breakpoints delineate natural groupings of nominal, moderate, and high-deviation arrivals and allow for consistent comparison across all noise-based performance metrics.

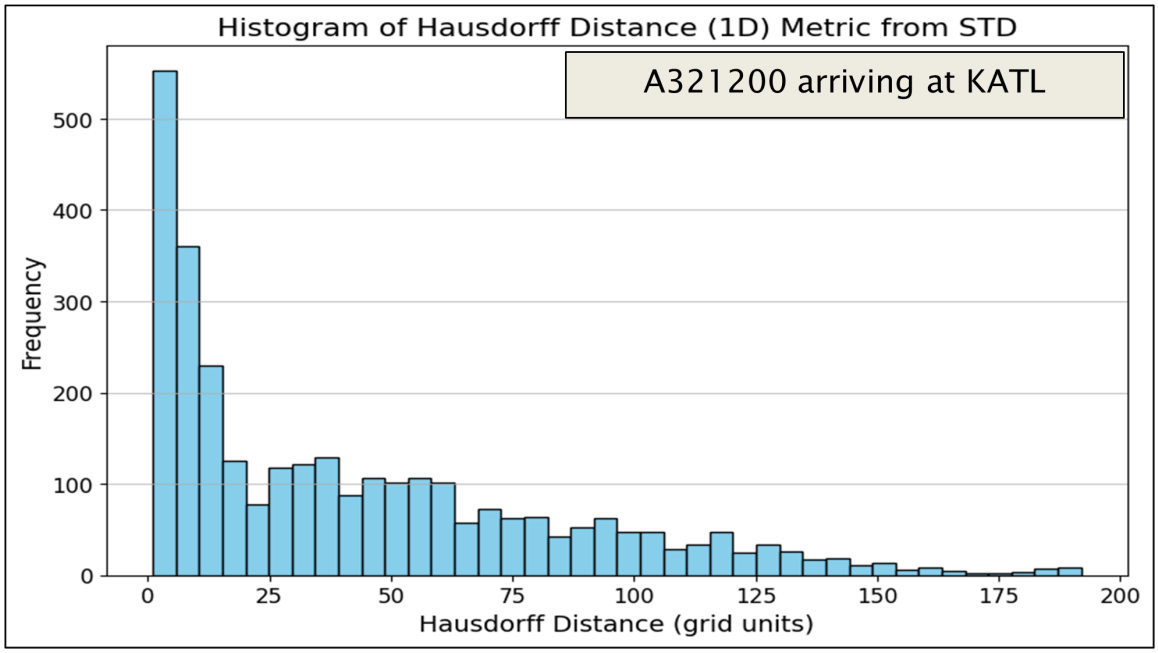


Figure 14. Distribution of Hausdorff distance (1D) from standard profile (STD) at Hartsfield-Jackson Atlanta International Airport (KATL).

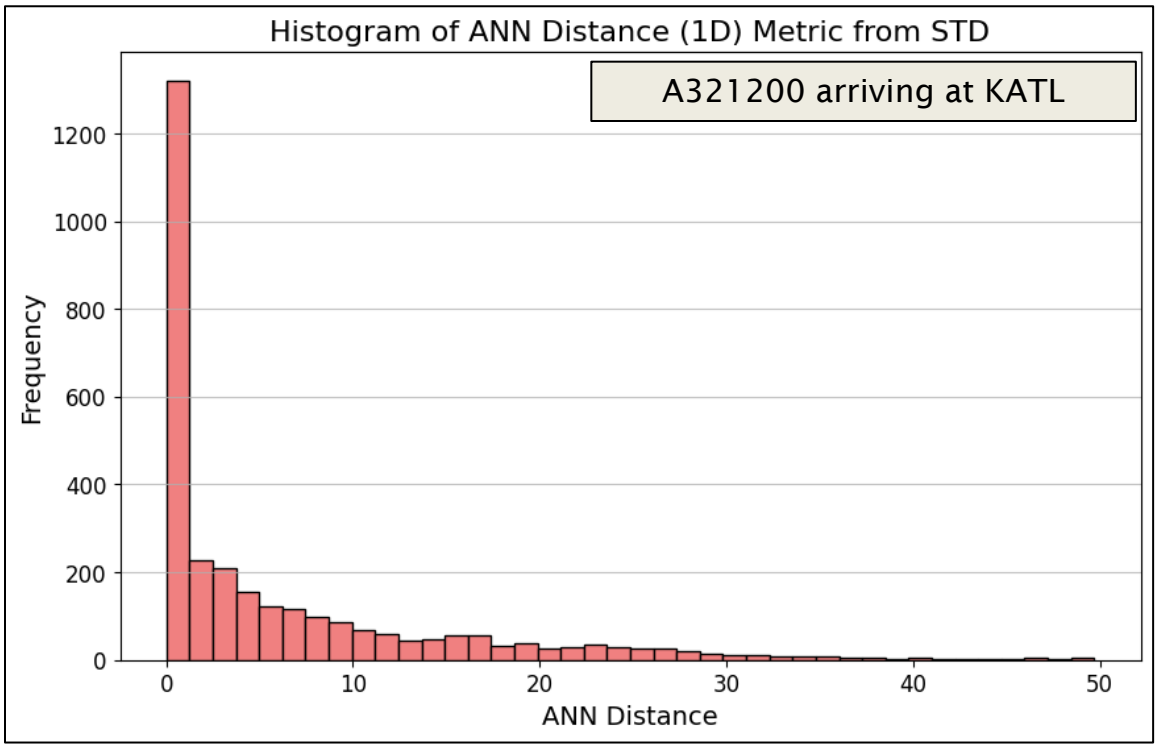


Figure 15. Distribution of ANN distance (1D) from standard profile (STD) at Hartsfield-Jackson Atlanta International Airport (KATL).

Analytical Workflow

The workflow for this year was designed to connect trajectory deviations directly to their noise footprints:

1. **Baseline Evaluation** – Compute AEDT noise contours for the standard arrival profile and compare them with FOQA-derived FPPs.
2. **Metric Binning** – Categorize flights by ANN distance into bins (0–5, 5–10, 10–20, 20+) to separate nominal versus anomalous operations.
3. **Noise Contour Analysis** – Evaluate footprint area, overlap, and ΔSEL distribution for each ANN bin across 75–95 dB SEL thresholds.
4. **Profile Recreation Experiments** – Modify key procedural parameters (level-off altitude, thrust mode, flap timing) in the AEDT to reproduce noise behavior seen in outlier FOQA flights.

The baseline comparison indicated that most FOQA flights align closely with the AEDT standard profile, with ANN scores <5 and Hausdorff <50, confirming strong conformity. The long right-tail of higher ANN values pinpointed a subset of non-standard arrivals requiring further evaluation.

Results & Discussion

Metric Correlation and Threshold Validation

The relationship between ANN-1D and Hausdorff-1D values for all FOQA arrivals is shown in Figure 16. A nonlinear correlation was observed between ANN and Hausdorff metrics: flights with high ANN distances generally exhibited high Hausdorff distances as well, though the sensitivity varied. Inflection points near ANN = 5, 10, and 20 validated the chosen binning scheme, distinguishing nominal, moderate, and highly deviating arrivals. For this analysis, both ANN-1D and Hausdorff-1D were computed using the 75 dB SEL contour, as this threshold corresponds to footprint extents of approximately 10–15 nmi that are most relevant for community noise exposure and reveal the clearest separation between standard and non-standard arrivals. The "1D" versions of these metrics quantify contour differences strictly along the centerline ground track, isolating along-track deviations without accounting for lateral dispersion.

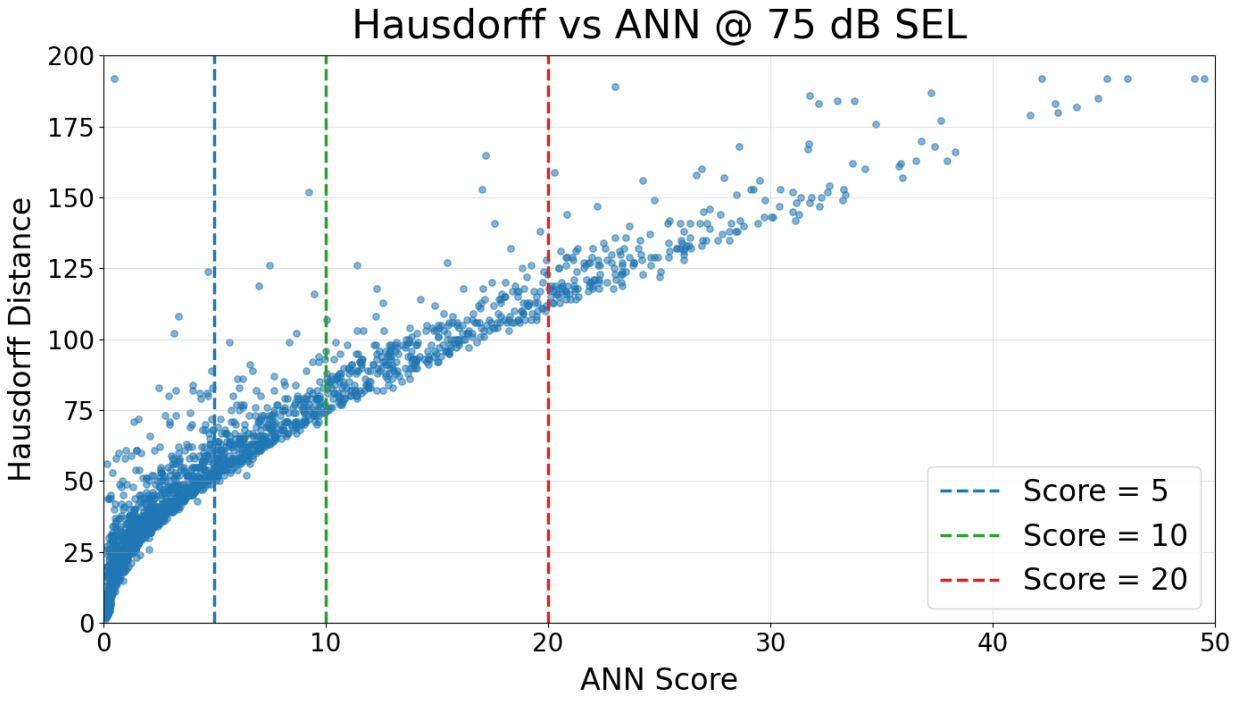


Figure 16. Hausdorff vs. Average Nearest Neighbor (ANN). SEL: sound exposure level.

Noise Footprint Trends

As the ANN score increased, flights exhibited systematically larger noise footprints and reduced spatial overlap with the standard contour. These trends are illustrated in Figure 17 and Figure 18, which show how IoU and percent area difference vary with SEL threshold for the 10–20 and 20+ ANN bins. At the 75 dB SEL threshold, flights in the 10–20 ANN bin showed roughly +80% footprint inflation, while those in the 20+ bin exceeded +130%. The IoU values correspondingly dropped from 0.75–0.8 for low-ANN cases to below 0.4 for the high-ANN group. At higher thresholds (85–90 dB), these differences narrowed, indicating that the most significant deviations occur in the outer, lower-noise regions of the footprint.

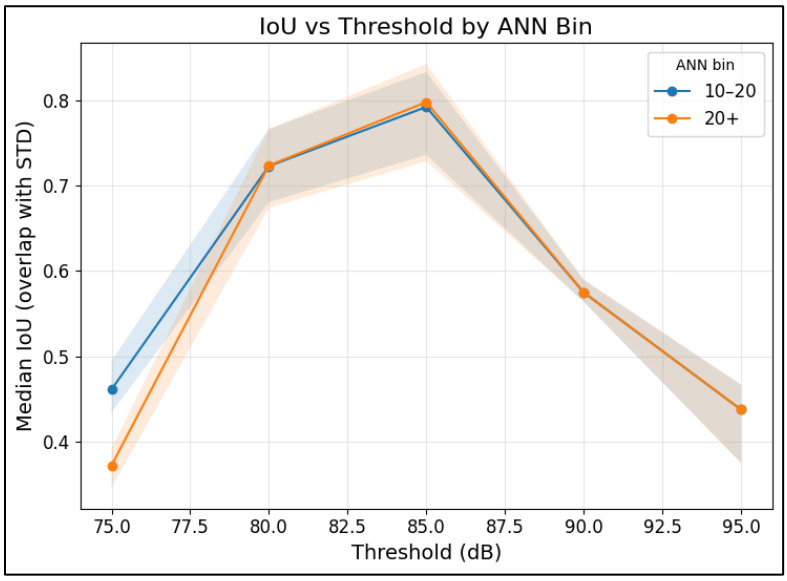


Figure 17. Intersection over Union (IoU) vs. threshold for A321-200 arrivals. ANN: Average Nearest Neighbor, STD: standard profile.

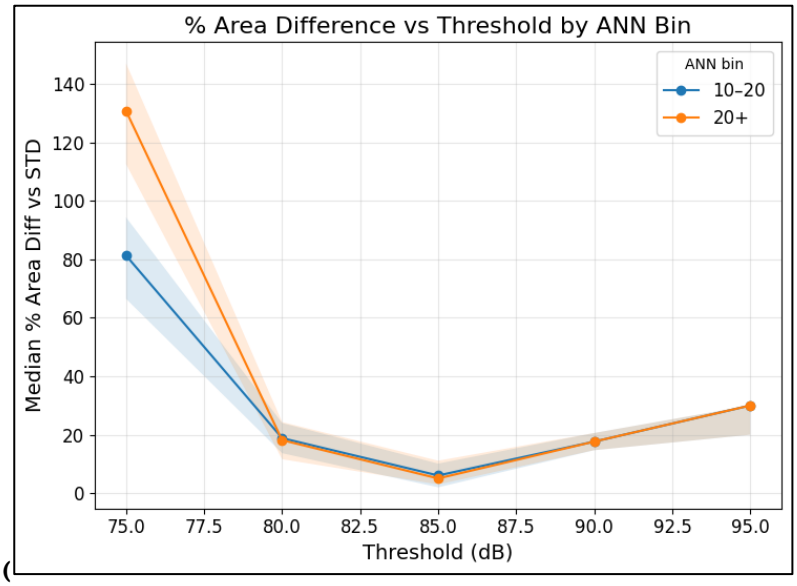


Figure 18. Percent area difference vs threshold for A321-200 arrivals. ANN: Average Nearest Neighbor, STD: standard profile.

Spatial ΔSEL Hotspots

To visualize how noise deviations manifest spatially, ΔSEL maps were averaged by ANN bin. These spatial noise deviation patterns are shown in Figure 19, Figure 20, Figure 21, and Figure 22 which illustrate the mean ΔSEL contours for the 0-5, 5-10, 10-20 and 20+ ANN bins respectively.

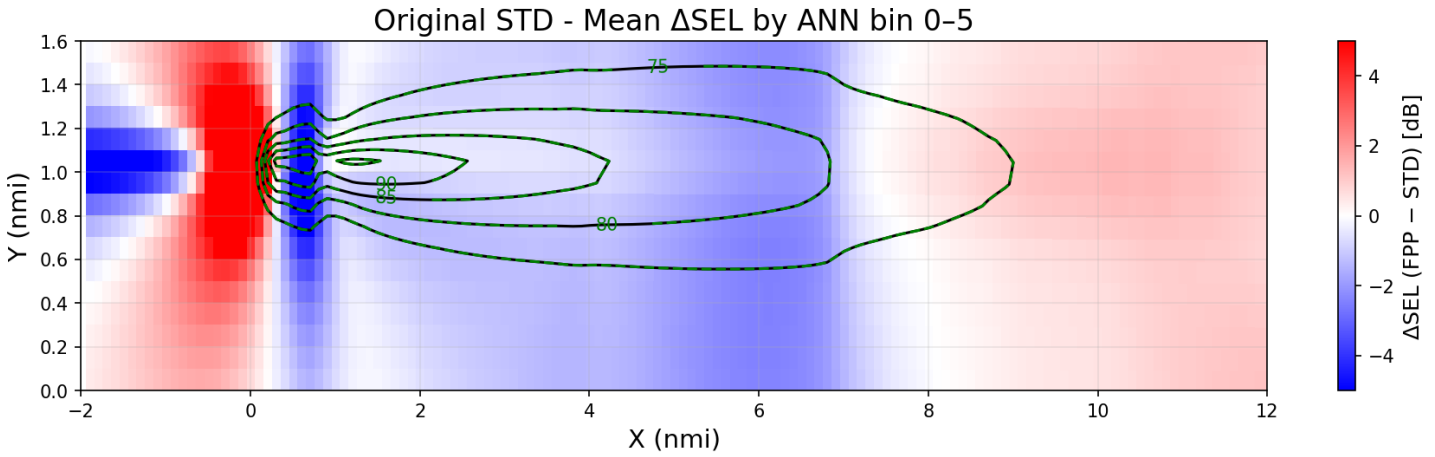


Figure 19. Mean ΔSEL for ANN 0–5 flights (FPP – STD). ANN: Average Nearest Neighbor, FPP: fixed point profile, SEL: sound exposure level, STD: standard profile.

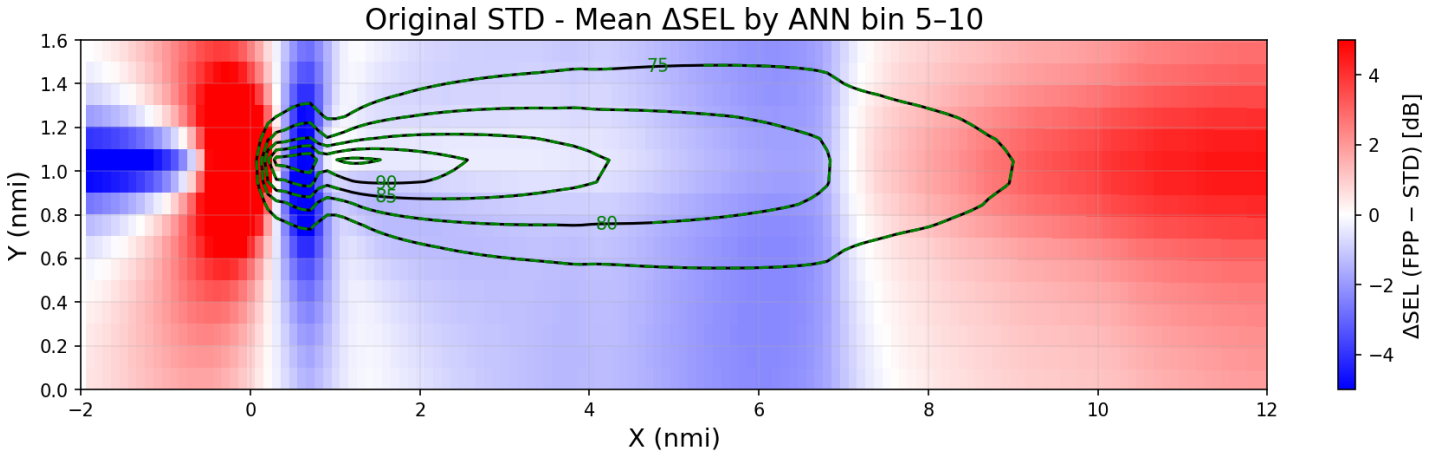


Figure 20. Mean ΔSEL for ANN 5–10 flights (FPP – STD). ANN: Average Nearest Neighbor, FPP: fixed point profile, SEL: sound exposure level, STD: standard profile.

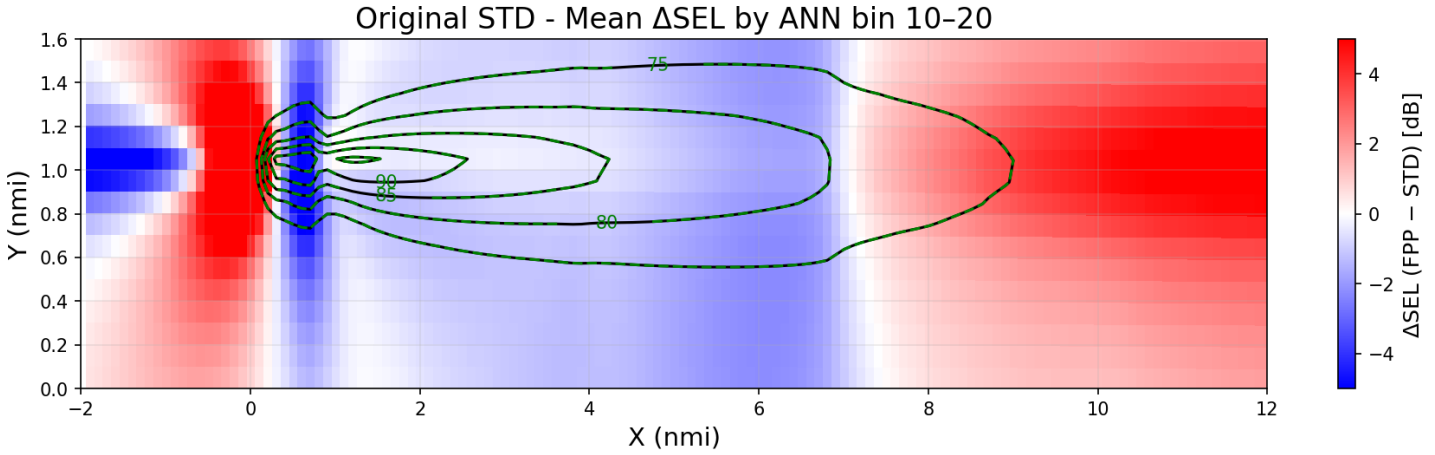


Figure 21. Mean ΔSEL for ANN 10–20 Flights (FPP – STD). ANN: Average Nearest Neighbor, FPP: fixed point profile, SEL: sound exposure level, STD: standard profile.

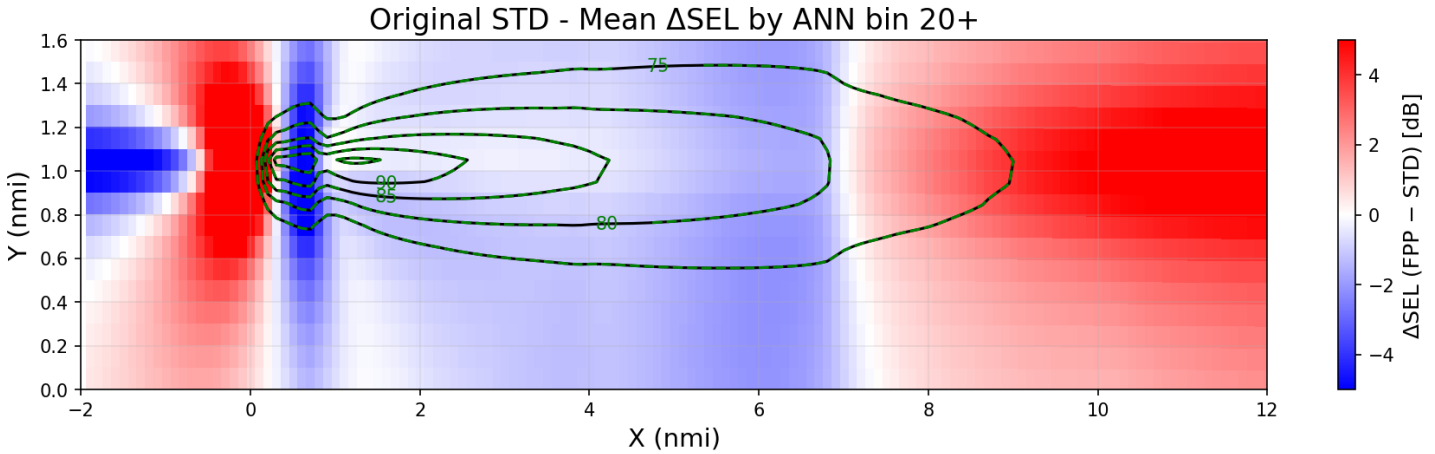


Figure 22. Mean ΔSEL for ANN 20+ Flights (FPP – STD). ANN: Average Nearest Neighbor, FPP: fixed point profile, SEL: sound exposure level, STD: standard profile.

Low-ANN flights (0–5) showed near-neutral behavior, with minor positive deviations near touchdown. Moderate-ANN bins (5–10 and 10–20) displayed progressively stronger upstream "hotspots," approximately 8–11 nautical miles from touchdown along the centerline, corresponding to level-off and deceleration regions. The highest-ANN group (20+) showed the most extensive and persistent positive ΔSEL band, confirming that these flights generate broader noise contours during the terminal descent.

Effect of Level-Off Altitude

Follow-up experiments adjusted the level-off altitude in the standard AEDT profile from 3,000 ft AFE to 2,000 ft AFE. This modification reduced the number of high-ANN outliers, particularly in the 20+ range, and yielded slightly lower Hausdorff extremes. The 2,000-ft STD thus better matched FOQA arrivals exhibiting low-altitude level-offs and stronger thrust applications. Conversely, 4000-ft level-offs showed near-perfect conformity (ANN < 5), consistent with smoother Continuous Descent Approaches (CDAs). These effects are illustrated in Figure 23, which compare altitude–distance descent trajectories for the original 3,000-ft standard profile and the modified 2,000-ft profile.

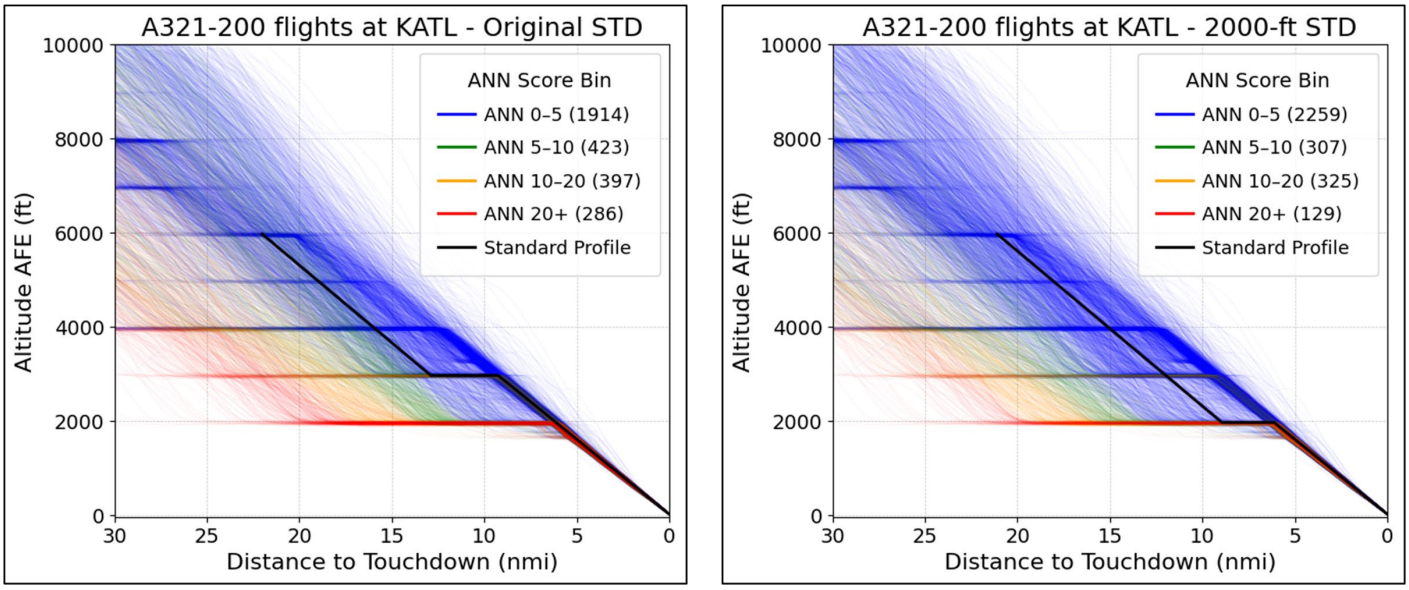


Figure 23. Trajectory plot by Average Nearest Neighbor (ANN) bin for original (left) and 2,000-ft modified (right) standard profile (STD) at Hartsfield-Jackson Atlanta International Airport (KATL).

Profile Perturbation and Noise Reproduction
To replicate noise "islands" observed in high-ANN flights, several perturbed profiles were simulated in the AEDT. Lowering airspeed during level-off (from ~250 kt to ~200 kt) and deploying flaps earlier reproduced the characteristic elongated 75 dB contours. These results suggest that noisier profiles are driven by early flap transitions and prolonged low-altitude level-offs, where sustained thrust produces elevated acoustic energy.
Across multiple case studies, the simulations confirmed that:

- **Low-speed, early-flap profiles** yield larger 75 dB footprints and higher ANN scores.
- **Higher-altitude, short level-offs** (≥4,000 ft) are acoustically benign and closely match AEDT's default model.
- **Intermediate cases (2,000–3,000 ft)** exhibit the most variability and drive the need for refined procedural modeling.

Overall, the analysis demonstrated that ANN and Hausdorff metrics can effectively characterize deviations between modeled and real-world noise footprints. Most arrivals are well captured by AEDT's standard procedures; however, specific subsets of flights–particularly those with low-altitude or extended level-offs and early flap deployment—produce significantly different and generally noisier signatures. Introducing modified profiles that account for these behaviors substantially improves contour alignment and reduces ANN deviations.

Milestones
None.

Major Accomplishments
None.

Publications
None.

Outreach Efforts
None.

Awards
None.

Student Involvement
Graduate research assistants, Anushka Moharir, and Hansen Lian from the School of Aerospace Engineering participated in this research. Anushka and Hansen were involved with the generation of noise results, exploration of various distance metrics, and analysis of results.

Plans for Next Period
- Refine metric thresholds and expand the perturbed profile library to include additional thrust and configuration variations.
- Quantify how each modified profile impacts overall contour statistics across multiple aircraft types. These results will support recommendations for new AEDT arrival procedures that better reflect real-world descent and configuration behavior.

Task 3 – Full-flight Modeling
Georgia Institute of Technology

Objectives
The objective of this study is to enhance the usability and analytical capabilities of full-flight modeling in the AEDT by developing an effective alternative to the existing sensor path method. While AEDT's sensor path remains the standard method, it is a high-fidelity and complex model, requiring heavy computation, and depends on detailed trajectory data that might not be accessible for large-scale analyses. This work aims to develop a simplified, statistically robust modeling framework that reduces dependence on high-resolution trajectory data while still delivering accurate gate-to-gate fuel burn and emissions estimates. The overall goal is to improve AEDT's capacity to support nationwide, data-driven assessments of flight performance within the National Airspace System (NAS).

In the previous year, we developed regression-based and machine learning models to estimate fuel consumption for both enroute and gate-to-gate operations across 16 representative city pairs. This research identified critical factors impacting fuel consumption, including aircraft attributes, flight operation variables, and relevant weather conditions. The results demonstrate that a data-driven modeling framework implemented in standalone Python® workflows can accurately estimate fuel burn, without requiring direct integration into the AEDT. The proposed methods serve as a complementary alternative to AEDT's sensor-path based approach and enable scalable fuel burn analyses outside the AEDT environment while using operationally relevant inputs.

Building on previous work that examined fuel burn modeling for enroute flight segments above 10,000 ft mean sea level (MSL), the current study refines and extends the research scope to focus on gate-to-gate fuel-burn prediction. This shift reflects the need to capture fuel consumption across all phases of flight, including surface operations and terminal-area maneuvers, which are not fully represented in enroute-only analyses. Using the CY2023 FOQA dataset, we investigated various approaches to improve the accuracy of fuel burn and emissions forecasts. To organize the analysis and display our findings clearly, the study is divided into three separate investigations.

1. Weather Integration Strategies

This investigation examines various methods of integrating Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) weather data, such as a baseline without weather, real-time atmospheric conditions, and monthly or seasonal climatological fields. It aims to identify which approach best enhances fuel-burn prediction for the initial 16 city pairs.

2. Operational and Directional Feature Enhancement

This study investigates the impact of additional operational variables, including takeoff weight and flight direction, on model accuracy and physical consistency. These variables help better represent wind effects and changes in fuel consumption caused by aircraft weight.

® Python is a registered trademark of Voxx International Corporation, Hauppauge, New York.

3. Large-Scale Application and Scalability Assessment

The final investigation applies the refined modeling approach to an expanded set of 86 city pairs with varied geographic and operational characteristics. Multiple algorithms, including Random Forest and Multilayer Perceptron (MLP), are employed to evaluate the scalability, prediction stability, and generalizability of the modeling framework.

Together, these three investigations establish a solid foundation for improving the AEDT's fuel-burn modeling capabilities and guiding future updates to its performance modeling framework.

The other side of this task builds upon and completes the initial clustering work by the team in 2024. The objective of the task is to analyze the performance of AEDT's sensor path-based modeling process, through a comparative study between representative 'core paths' for city pairs modeled in the AEDT, versus real-world flight data sourced from the CY2023 FOQA dataset. Secondary objectives for the study include developing and streamlining the workflow for modeling sensor path information in the AEDT, investigating bugs and issues with the modeling process, and providing recommendations for automating the process and solving any issues encountered. The overarching motivation behind this study is to set up an efficient workflow that can leverage the modeling of large-scale flight databases such as the NAS database in order to provide recommendations for future tool features and improvements.

Research Approach

Investigation 1. Weather Integration Strategies

Methodology

The initial study explored how various weather-representation strategies affect the accuracy of gate-to-gate fuel-burn predictions. Although the investigation focused on gate-to-gate fuel burn prediction, the weather integration methodology was applied only to the enroute phase of flight above 10,000 ft MSL. The objective was to identify a practical approach that improves model performance without adding extra data processing or computational burden, consistent with typical operational modeling limitations. The research analyzed CY23 FOQA flight data from 16 city pairs, encompassing 2,302 flights. Table 3 provides the FOQA flight counts per city pair, and Figure 24 illustrates the geographical distribution of these routes across the continental United States (U.S.), highlighting the range of operating conditions encompassed in the dataset.

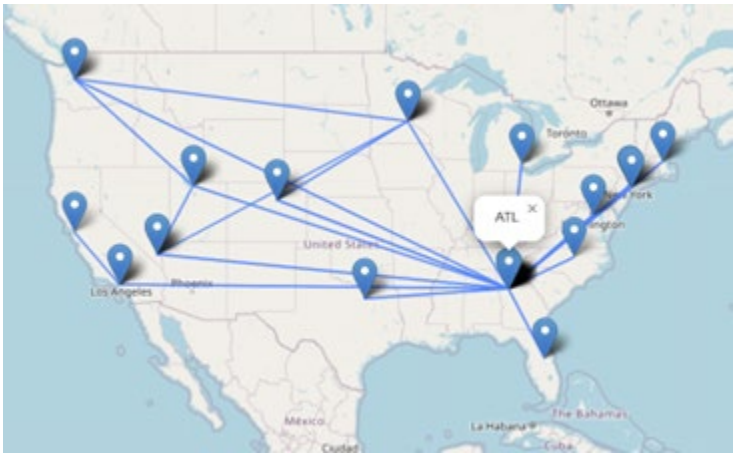


Figure 24. Visualization of 16 Origin-Destination (OD) pairs.

Table 3. Total flights selected from CY23 FOQA data for 16 city pairs. KATL: Hartsfield-Jackson Atlanta International Airport, KDCA: Ronald Reagan Washington National Airport, KJFK: John F. Kennedy International Airport, KLAS: Harry Ried International Airport, KLGA: LaGuardia Airport, KMSP: Minneapolis-Saint Paul International Airport, KSEA: Seattle-Tacoma International Airport, KSLC: Salt Lake City International Airport.

City Pair	Number of Flights	City Pair	Number of Flights
KATL-KDCA	163	KDCA-KATL	213
KATL-KJFK	72	KJFKKATL	59
KATL-KLAS	174	KLAS-KATL	160
KATL-KLGA	172	KLGA-KATL	279
KATL-KMSP	173	KMSP-KATL	147
KATL-KSEA	157	KSEA-KATL	127
KATL-KSLC	151	KSLC-KATL	129
KLAS-KSLC	68	KSLC-KLAS	58

Alternative Method 1 – Baseline (Aircraft and airport features only)
The baseline configuration relied exclusively on aircraft, airport, and basic flight-operation variables. No weather information was incorporated. This method required minimal preprocessing and served as the foundational reference for evaluating the added value of weather-based inputs in subsequent configurations. The features included in the baseline model are summarized in Table 4 below.

Table 4. Feature categories included in the baseline configuration. ICAO: International Civil Aviation Organization.

Category	Parameters
Airport	• Takeoff airport ICAO code • Takeoff runway • Landing airport ICAO code • Landing runway
Aircraft	• Aircraft category • Airframe • Engine type
Flight Operation	• Departure time • Season • Month • Azimuth • Flight direction • Flight time • Flight distance

Alternative Method 2 – Instantaneous MERRA-2 En-route Weather
The second method incorporated meteorological variables from instantaneous MERRA-2 data. Figure 25 illustrates the processing steps involved in obtaining weather data at each trajectory point. The process included three steps: (1) altitude interpolation, (2) spatial interpolation, and (3) temporal interpolation.

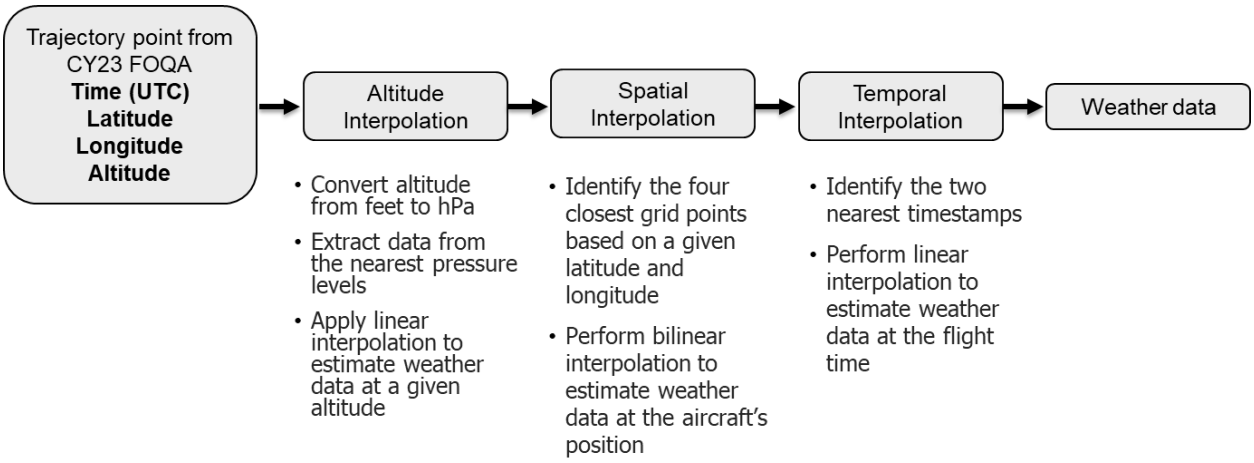


Figure 25. Workflow for obtaining instantaneous MERRA-2 weather at Flight Operations Quality Assurance (FOQA) trajectory points. CY: calendar year.

For each FOQA trajectory point, the aircraft's altitude in feet was converted to the corresponding pressure level in hectopascals. Weather variables from the nearest pressure surfaces were extracted, and linear interpolation was used to estimate the atmospheric conditions at the aircraft's exact altitude. Spatial interpolation was then performed by identifying the four closest MERRA-2 grid points based on latitude and longitude, applying bilinear interpolation to estimate weather conditions at the aircraft's horizontal position. Finally, temporal interpolation employed the two nearest MERRA-2 timestamps to calculate weather values at the precise time of flight.

Wind components and temperature were obtained for trajectory points above 10,000 ft MSL to accurately capture en-route atmospheric conditions. The interpolated data along each flight path were averaged to create a composite weather profile for each flight. This approach provided a clear, flight-specific perspective of the large-scale atmosphere, perfectly complementing the real-time weather method.

Alternative 3 - Monthly or Seasonal MERRA-2 Profiles
This method used monthly or seasonal averages from MERRA-2 to create a simplified model of enroute atmospheric conditions. For each origin–destination pair, a representative flight path was identified using a trajectory-clustering procedure applied to the FOQA data. The centroid trajectory from the most common cluster served as the standard route for that city pair.

Weather information was gathered only from the part of the representative trajectory above 10,000 ft MSL. Five sampling points were chosen along this enroute segment. At each point, monthly or seasonal averages of wind components and temperature were obtained from the MERRA-2 climatological fields. The five values were combined into a single climatological weather profile for each city pair.

This approach reduces data volume and preprocessing requirements compared with instantaneous weather retrieval. However, it does not capture day-specific atmospheric variability. The method therefore supports an assessment of the trade-off between simplified climatology-based inputs and the resulting changes in model performance.

In this study, our team implemented two models — (1) decision tree and (2) random forest — to predict aircraft fuel consumption and identify the most accurate predictor. Each model was evaluated using standard performance metrics, including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2).

MAE measures the average magnitude of errors between predicted and actual values, with lower values indicating more accurate predictions. MSE calculates the average of squared differences, where a lower value also signifies better performance. RMSE, the square root of MSE, offers an error value in the same unit as the target variable, aiding

interpretation. R^2 reflects how well the model explains variability in the target, with values closer to 1 indicating a better fit.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Fuel\ Burn_{Actual} - Fuel\ Burn_{Predicted}| \quad \text{(Eq. 1)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Fuel\ Burn_{Actual} - Fuel\ Burn_{Predicted})^2 \quad \text{(Eq. 2)}$$

$$RMSE = \sqrt{MSE} \quad \text{(Eq. 3)}$$

As mentioned earlier, this study implemented two types of predictive models. The first is the decision tree, which effectively captures complex nonlinear relationships by iteratively splitting the data and making decisions at each node. This approach provides insights into the relative importance of each feature; however, decision trees are prone to overfitting, making them sensitive to variations in the training data.

The final model, the random forest, addresses this limitation by aggregating multiple decision trees, thereby reducing overfitting and enhancing generalization. By combining the outputs of many trees, random forests capture complex feature interactions, resulting in improved robustness and accuracy of predictions.

The dataset was partitioned into an 80% training set and a 20% testing set. In addition, five-fold cross-validation was applied within the training portion to assess model stability and ensure that performance estimates were not dependent on a single data split.

Results and Discussion

Investigation 1 examined how different methods of representing en-route weather affect gate-to-gate fuel-burn prediction. Four configurations were evaluated:

1. Operational features only
2. Instantaneous MERRA-2 averages
3. Monthly MERRA-2 averages
4. Seasonal MERRA-2 averages

Decision Tree and Random Forest models were applied consistently across all configurations to ensure that performance differences were due to the weather representation rather than the modeling technique.

Table 5 summarizes the Decision Tree performance across the four configurations. Incorporating instantaneous MERRA-2 weather reduced the MAE from 1.207 to 0.969 and reduced the RMSE from 1.927 to 1.832 compared with the operational-only baseline. The R^2 value increased from 0.957 to 0.961. Monthly and seasonal MERRA-2 averages resulted in MAE values of 1.029 and 1.035, respectively, and RMSE values of 1.981 and 1.863. These results show that the instantaneous weather method produced the lowest overall errors among the four configurations.

Table 5. Performance of Decision Tree fuel burn prediction models using different weather representation methods. MAE: mean absolute error, MERRA-2: Modern-Era Retrospective analysis for Research and Applications, Version 2, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	Operational Feats	Avg. En-route Wx from Inst. MERRA2	Avg. En-route Wx from Monthly MERRA-2	Avg. En-route Wx from Seasonal MERRA-2
MAE	1.207	0.969	1.029	1.035
MSE	3.878	3.435	3.982	3.519
RMSE	1,927	1.832	1.981	1.863
R^2	0.957	0.961	0.955	0.960

Figure 26 shows the actual and predicted fuel-burn values and the corresponding residuals. The red dashed line indicates the condition in which the predicted value equals the observed value. Most points in the actual-versus-predicted panel lie

close to this line for lower and moderate fuel-burn values. Points begin to show a greater vertical spread when the actual fuel burn exceeds approximately 60,000 lb In the residual panel, most values are near zero, and larger residuals occur at higher fuel-burn levels. Some flights have residuals exceeding 20,000 lb.

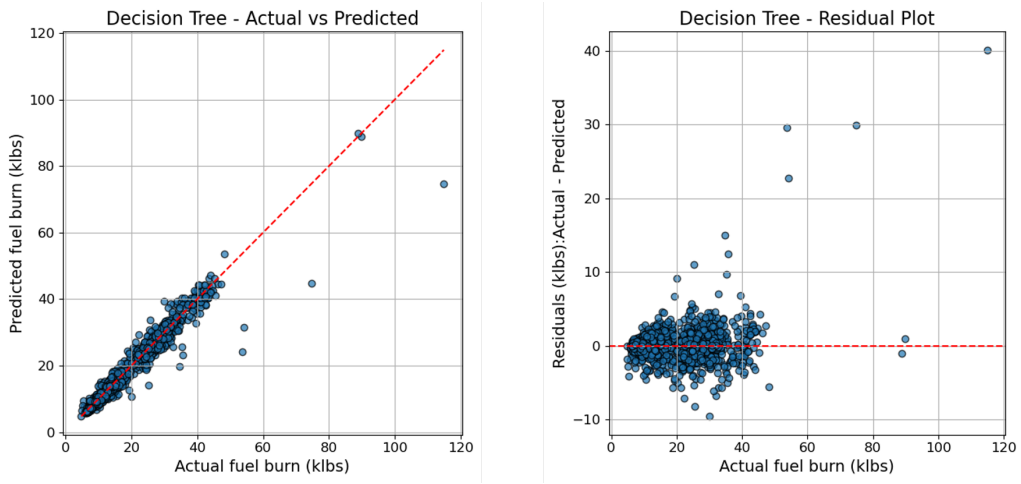


Figure 26. Decision Tree model performance for gate-to-gate fuel-burn prediction (16 city pairs).

Table 6 summarizes the Random Forest performance across the four configurations. The lowest MAE value was 0.701 and occurred when instantaneous MERRA-2 en-route weather was used. This configuration also produced the lowest RMSE value at 1.536. The operational-features-only configuration produced an MAE of 0.786 and an RMSE of 1.566. The monthly and seasonal MERRA-2 averages yielded MAE values of 0.792 and 0.808, respectively, and RMSE values of 1.582 and 1.684. The highest R^2 value was 0.974 for the operational-features-only configuration, followed closely by 0.972 for the instantaneous weather case.

Table 6. Performance of Random Forest fuel burn prediction models using different weather representation methods. MAE: mean absolute error, MERRA-2: Modern-Era Retrospective analysis for Research and Applications, Version 2, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	Operational Feats	Avg. En-route Wx from Inst. MERRA-2	Avg. En-route Wx from Monthly MERRA-2	Avg. En-route Wx from Seasonal MERRA-2
MAE	0.786	0.701	0.792	0.808
MSE	2.734	2.577	2.798	3.372
RMSE	1.566	1.536	1.582	1.684
R^2	0.974	0.972	0.970	0.964

Figure 27 shows the actual and predicted fuel-burn values and the associated residuals. The Random Forest predictions are tightly clustered along this line for most flights, with a narrower vertical spread compared with the Decision Tree results. The residual plot shows a concentrated distribution around zero for lower and moderate fuel-burn values. A smaller number of flights exhibit large positive residuals, with several exceeding 20,000 pounds.

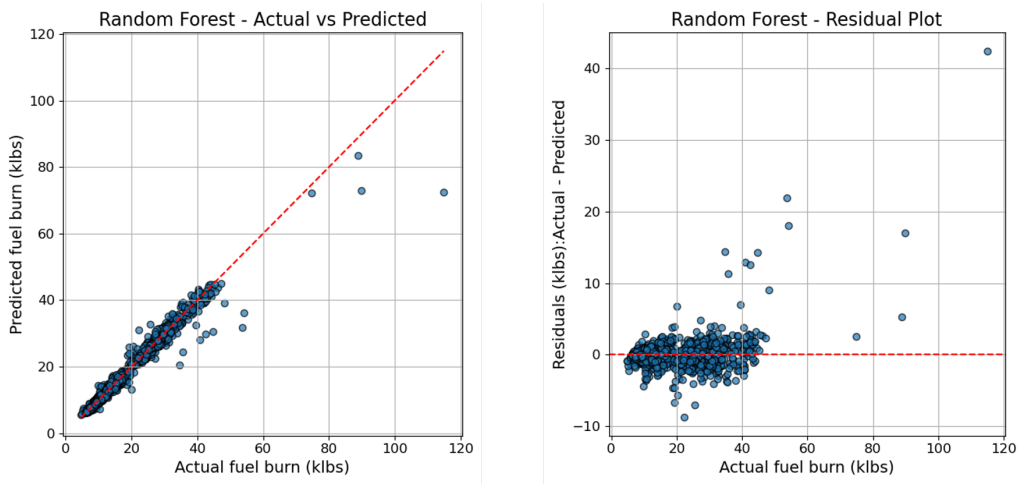


Figure 27. Random Forest model performance for gate-to-gate fuel-burn prediction (16 city pairs).

In all configurations, incorporating instantaneous MERRA-2 weather data reduced MAE and RMSE relative to the baseline that used only operational data. Monthly and seasonal averages from MERRA-2 yielded error levels similar to the baseline, offering no extra benefit. These findings suggest that only the instantaneous weather inputs significantly contributed to lowering prediction errors.

Investigation 2: Operational and Directional Feature Enhancement

Methodology

This investigation examined how two operational factors, aircraft takeoff weight and flight direction, affect the accuracy of gate-to-gate fuel-burn predictions. The analysis used the same weather-integrated framework as Method 2 in Investigation 1, which incorporated instantaneous MERRA-2 en-route weather data. The objective was to evaluate whether adding these operational variables improves prediction accuracy beyond what is obtained using aircraft and weather variables alone.

Two modifications were applied to the baseline configuration:

1. Aircraft takeoff weight: Aircraft takeoff weight was obtained from the FOQA dataset and represents the weight recorded for each flight at takeoff. Because takeoff weight affects the thrust and fuel required across all phases of flight, including this variable, allows the models to represent fuel-burn differences associated with variations in aircraft mass.
2. Flight direction classification: Flight direction was determined by comparing the longitudinal positions of the origin and destination airports for each city pair. Flights ending east of the departure airport were classified as eastbound, and flights ending to the west were classified as westbound. This classification provides a straightforward way to separate flights operating in opposite directions along the same route, allowing wind-related differences between the two directions to be evaluated.

In this investigation, the same modeling framework and evaluation metrics used in Investigation 1 were applied to assess the impact of the additional operational variables. Decision tree and random forest models were trained using the expanded feature set, and their performance was evaluated using MAE, MSE, RMSE, and R^2.

Results and Discussion

This initial analysis was conducted for 16 city pairs, which was later expanded to 86 city pairs in Investigation 3.

The inclusion of aircraft takeoff weight improved prediction accuracy in both models. For the Decision Tree, MAE decreased from 0.969 to 0.882, and RMSE decreased from 1.832 to 1.747 (Table 7). For the Random Forest model, MAE decreased from 0.701 to 0.607, and RMSE decreased from 1.536 to 1.300 (Table 8). In both models, the R^2 values increased slightly. These results indicate that adding takeoff weight consistently reduced prediction errors relative to the configuration without this variable.

Table 7. Fuel burn prediction performance of Decision Tree models with and without aircraft takeoff weight. MAE: mean absolute error, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	Without Takeoff Weight	With Takeoff Weight
MAE	0.969	0.882
MSE	3.435	3.261
RMSE	1.832	1.747
R^2	0.961	0.963

Table 8. Fuel burn prediction performance of Random Forest models with and without aircraft takeoff weight. MAE: mean absolute error, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	Without Takeoff Weight	With Takeoff Weight
MAE	0.701	0.607
MSE	2.577	1.950
RMSE	1.536	1.300
R^2	0.972	0.979

As shown in Table 9 and Table 10, flight direction resulted in measurable differences in prediction accuracy for both models. For the Decision Tree, the MAE values for eastbound and westbound flights were 0.899 and 0.929, compared with 0.882 for all directions. The corresponding RMSE values were 1.620 for eastbound flights and 1.806 for westbound flights. For the Random Forest model, the MAE values were 0.596 for eastbound flights and 0.671 for westbound flights, compared with 0.607 for the combined dataset. The RMSE values were 1.301 for eastbound flights and 1.548 for westbound flights. These results indicate that eastbound flights generally exhibited lower error values than westbound flights in both modeling approaches.

Table 9. Fuel burn prediction performance of Decision Tree models by flight direction. MAE: mean absolute error, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	All Flight Directions	Eastbound	Westbound
MAE	0.882	0.899	0.929
MSE	3.261	3.367	3.701
RMSE	1.747	1.620	1.806
R^2	0.963	0.954	0.968

Table 10. Fuel burn prediction performance of Decision Tree models by flight direction. MAE: mean absolute error, MSE: mean squared error, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	All Flight Directions	Eastbound	Westbound
MAE	0.607	0.596	0.671
MSE	1.950	2.322	3.080
RMSE	1.300	1.301	1.548
R^2	0.979	0.969	0.975

This study evaluated the influence of aircraft takeoff weight and flight direction on gate-to-gate fuel-burn prediction accuracy using the same modeling framework as Investigation 1. Adding the takeoff weight variable reduced prediction

errors for both the Decision Tree and Random Forest models, resulting in lower MAE and RMSE values compared to the configuration without this variable. Flight direction also affected error levels, producing different MAE and RMSE values for eastbound and westbound flights in both models. However, training direction-specific models did not yield an overall improvement in prediction accuracy relative to the combined-direction model. In the Random Forest case, the eastbound subset showed slightly lower errors; however, this reflects the properties of that subset rather than enhanced model performance. Overall, the results indicate that takeoff weight provides a clear contribution to accuracy, while separating models by flight direction does not provide a predictive benefit.

Investigation 3: Large-Scale Application and Scalability Assessment

Methodology

This investigation assessed how well the fuel-burn prediction framework generalizes when applied to a larger and more diverse set of operations. The dataset was expanded from 16 to 86 city pairs by adding routes in order of available flight counts, resulting in approximately 11,300 FOQA flights. The expanded network introduced additional variation in the atmospheric and geographic conditions represented in the dataset. All model configurations used the instantaneous MERRA-2 en-route weather variables defined in Method 2 of Investigation 1 and incorporated aircraft takeoff weight to maintain consistency with the enhanced feature set. Figure 28 illustrates the distribution of the 86 city pairs across major U.S. corridors, highlighting the increased range of operating conditions captured in the expanded dataset.

Two predictive modeling methods were evaluated: Random Forest and a Multilayer Perceptron (MLP). Random Forest was selected as the primary model because it demonstrated stronger performance than the other methods tested in the earlier investigations. Both models used the complete feature set developed in Investigation 2, which included aircraft type, FOQA-derived takeoff weight, aircraft construction year, flight direction, airport characteristics, and averaged en-route MERRA-2 weather variables.

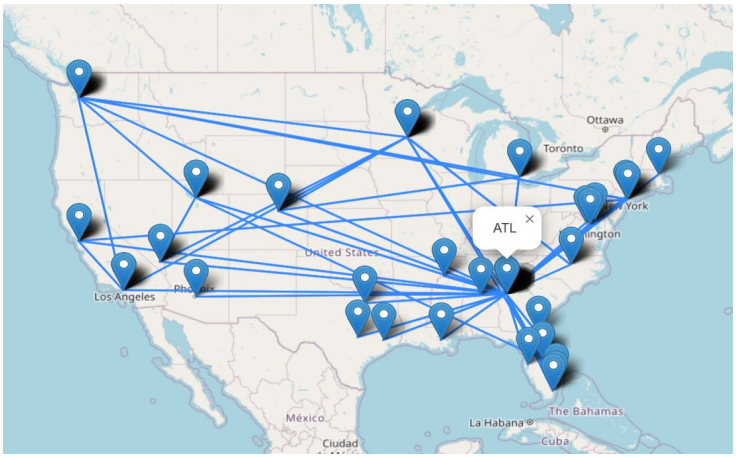


Figure 28. Visualization of 86 Origin-Destination (OD) pairs.

An MLP model was also implemented to examine whether a neural-network approach could represent nonlinear relationships among operational variables and atmospheric conditions that the Random Forest could not capture. The network consisted of two hidden layers with 64 and 32 nodes using rectified linear activation and a linear output layer. All numerical features were standardized before training. The MLP used a 70/15/15 train–validation–test split to allow independent assessment of hyperparameter tuning and final generalization performance.

Results and Discussion

The expansion from 16 to 86 city pairs improved the performance of the Random Forest model. As shown in Table 11, the MAE decreased from 0.607 to 0.474, and the RMSE decreased from 1.300 to 0.887. The R^2 value increased from 0.979 to 0.988. The actual-versus-predicted plot shows that most predictions fall close to the line of perfect agreement, with larger deviations appearing primarily at higher fuel-burn values. The corresponding residuals are centered near zero, with greater spread for flights with higher fuel-burn values.

Table 11. Fuel burn prediction performance with expanded city pairs and an added Multilayer Perceptron (MLP) model. MAE: mean absolute error, MSE: mean squared error, OD: Origin-Destination, RMSE: root mean squared error, R^2: coefficient of determination.

Metric	Random Forest 16 OD pairs	Random Forest 86 OD pairs	MLP 86 OD pairs
MAE	0.607	0.474	0.442
MSE	1.950	0.807	0.477
RMSE	1.300	0.887	0.691
R^2	0.979	0.988	0.993

The MLP model was evaluated only on the expanded 86-route dataset and produced the lowest overall error values among the models tested. The MAE and RMSE were 0.442 and 0.691, and the R^2 value was 0.993. The actual-versus-predicted plot shows close alignment with the reference line across the full fuel-burn range, and the residuals exhibit less dispersion than those from the Random Forest model (shown in Figure 29).

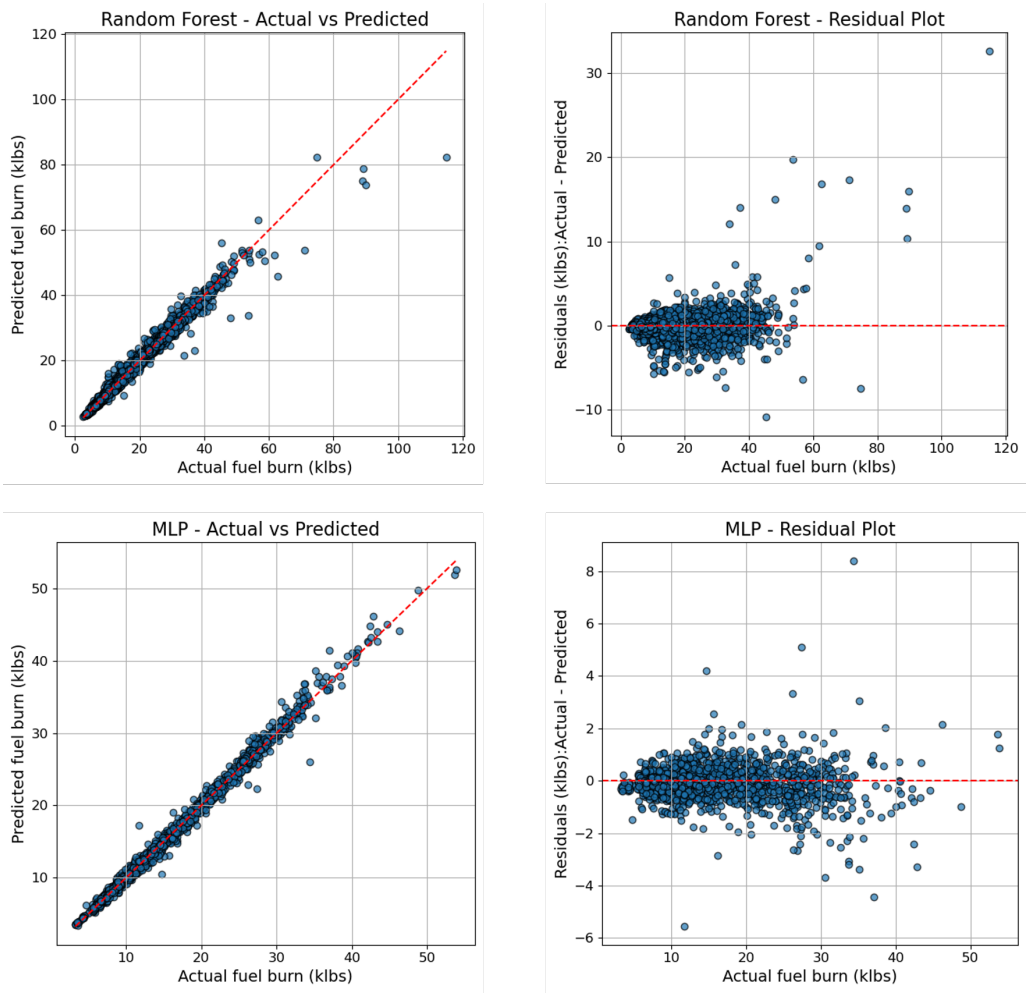


Figure 29. Multilayer Perceptron (MLP) and Random Forest models performance for gate-to-gate fuel-burn prediction (86 city pairs).

Overall, the results indicate that expanding the dataset enhanced the accuracy of the Random Forest model, and that the MLP model achieved the highest accuracy within the expanded 86-route configuration.

Investigation 4: Fuel Burn Analysis in AEDT Using Clustered Core Paths and Real-World Flight Data

Methodology

This study utilizes the AEDT 3e public release version 174.0.15710.1 and real-world full flight data provided by an airline partner, which forms part of the CY2023 FOQA dataset. As the entire FOQA dataset consists of over 50,000 individual flights, a total of 40 city pairs were chosen from the dataset for the purposes of this study based on flight frequency, industry inputs, and national importance. The 40 city pairs account for about 7,200 flights; these are described in Table 12. A flowchart summarizing the workflow followed by the team is shown in Figure 30.

Table 12. Total flights selected from FOQA data for selected city pairs. KATL: Hartsfield-Jackson Atlanta International Airport, KLAS: Harry Ried International Airport (Las Vegas, Nevada): KLGA: LaGuardia Airport (Elmhurst, New York), KMSP: Minneapolis-Saint Paul International Airport, KSEA: Seattle-Tacoma International Airport, KSLC: Salt Lake City International Airport, KJFK: John F. Kennedy International Airport (Queens, New York), KDCA: Ronald Reagan Washington International Airport (Arlington, Virginia), KRDU: Raleigh-Durham International Airport, KMCO: Orlando International Airport, KBOS: Boston Logan International Airport, KLAX: Los Angeles International Airport, KDEN: Denver International Airport, KDTW: Detroit Metropolitan Wayne County Airport, KDFW: Dallas Fort Worth International Airport, KSFO: San Francisco International Airport.

City Pair	Number of Flights	City Pair	Number of Flights	Great-Circle Distance (nmi)
KATL-KLAS	196	KLAS-KATL	169	1518
KATL-KLGA	175	KLGA-KATL	283	662
KATL-KMSP	191	KMSP-KATL	169	788
KATL-KSEA	192	KSEA-KATL	153	1896
KATL-KSLC	157	KSLC-KATL	141	1382
KLAS-KSLC	89	KSLC-KLAS	81	320
KATL-KJFK	103	KJFK-KATL	95	660
KATL-KDCA	163	KDCA-KATL	215	475
KATL-KRDU	274	KRDU-KATL	315	309
KATL-KMCO	306	KMCO-KATL	252	351
KATL-KBOS	147	KBOS-KATL	226	822
KATL-KLAX	152	KLAX-KATL	221	1691
KATL-KDEN	312	KDEN-KATL	215	1042
KATL-KDTW	165	KDTW-KATL	163	517
KATL-KDFW	170	KDFW-KATL	195	636
KMSP-KDEN	204	KDEN-KMSP	147	591
KSEA-KMSP	168	KMSP-KSEA	145	1216
KLAX-KSFO	157	KSFO-KLAX	116	293
KSEA-KSLC	161	KSLC-KSEA	135	598
KMSP-KLAS	156	KLAS-KMSP	130	1129

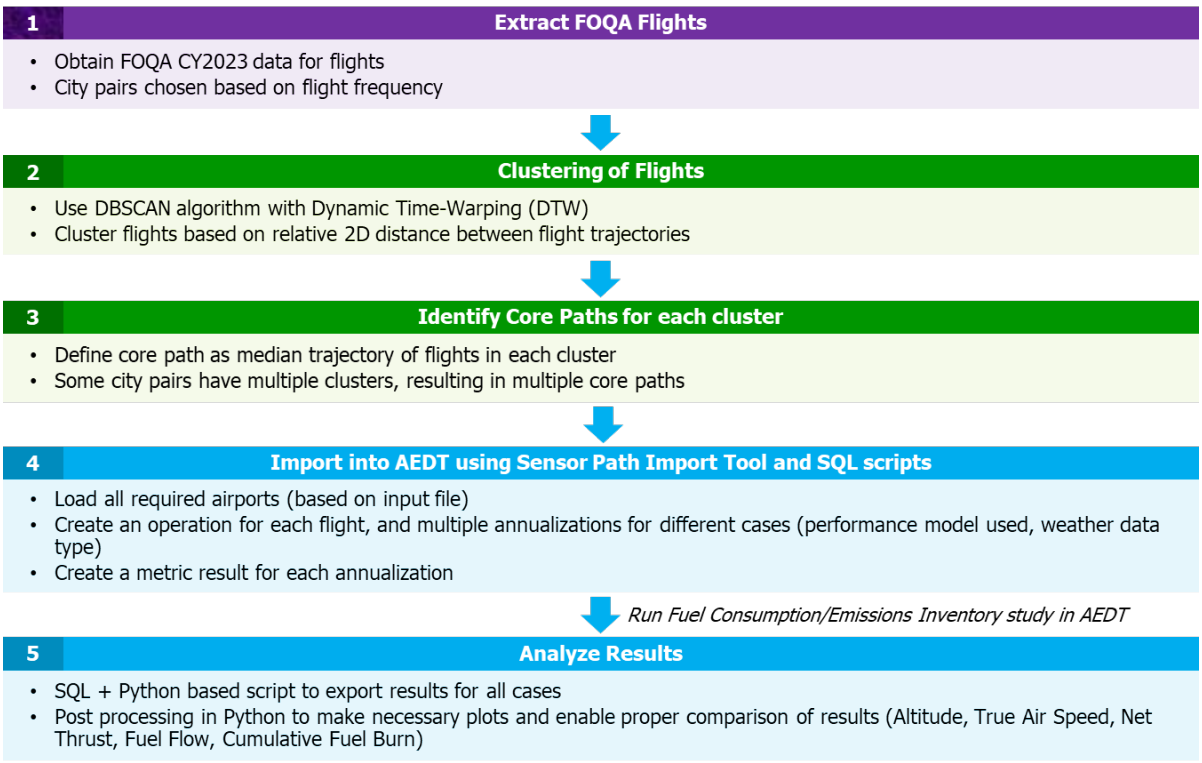


Figure 30. Flowchart summarizing the runway-to-runway fuel burn analysis process workflow. 2D: two-dimensional, CY: calendar year, FOQA: Flight Operations Quality Assurance, SQL: Structured Query Language.

All usable flight data from the chosen city pairs are extracted and properly formatted using a Python script written by the team. Any outlying flights (i.e., flights that return to their departure airport, or flights with invalid or corrupted data points) are rejected. Parameters extracted include flight trajectory information, airspeed, groundspeed, airframe and engine information, etc. Distance matrices are created out of the saved trajectory information of all flights for each city pair and saved in a high-efficiency binary format suitable for handling large amounts of numerical data. These distance matrices incorporate Dynamic Time-Warping, in order to properly and accurately calculate the relative two dimensional distance between two sets of coordinates belonging to different flights. The pre-computing of distance matrices significantly helps optimize clustering speed.

Once distance matrices are created, a clustering workflow developed by the team is run. The clustering workflow is based on the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, which is useful for clustering datasets such as trajectories. While the Hierarchical DBSCAN algorithm was used in the initial work done by the team in 2024, DBSCAN is chosen for this study as more granular control over clustering performance is possible with it. The algorithm clusters flight trajectories per city pair using the pre-computed distance matrices. Two runs of the algorithm are run. First, an iterative run that goes through combinations of compatible clustering parameters that are used to control the clustering performance (eps and min_samples). Eps is iterated from 0 to 200, and the min_samples are iterated from 2 to 20, covering all the possible combinations of both. The clusters generated by the first run are compared using two metrics, their R^2 value and their Davies-Bouldin score, to determine the optimal combination of parameters for each city pair. Once optimized parameters are decided, a second run of the clustering algorithm is performed to generate the final clusters for all city pairs. While it is noticed that many city pairs have a single general flight route dominating resulting in one major cluster, certain city pairs may have flights flying in one of multiple routes, leading to distinct clusters for those city pairs. In this case, if one route outweighs the other in terms of flights, the other is discarded. However, if the clusters are comparable in their flight frequency, they are all considered.

From the chosen cluster(s), the FOQA flight closest to the median flight trajectory within the cluster is selected as the 'core path,' which is essentially the representative flight path for a given city pair. As mentioned above, because of multiple clusters being chosen for certain city pairs, they may also have multiple core paths. Once determined, flight information for these core paths is extracted, reformatted and saved as two files, that will be required as inputs for the Safety Performance Indicator (SPI) tool in the AEDT. These two files are:

1. An 'operations' file, which contains information about the flight's wheels-off and wheels-on times, departure and arrival airports, aircraft, and AEDT-specific equipment ID.
2. A 'trajectory' file, which contains a sequential time series of the entire flight path and parameters, from departure runway to arrival runway. Parameters include latitude, longitude, pressure and Global Positioning System altitudes, true airspeed, ground track distance and timestamp.

An important aspect of these input files is the EquipID value for each flight, which controls the specific airframe and engine variant that will be used by the AEDT during the modeling process. The AEDT can contain multiple airframe and engine variants for a single aircraft type. Without any EquipID specified in the SPI input files, the AEDT may use a default airframe and engine variant, which may not exactly be the same configuration as the real-world flight, and is found to lead to inaccurate comparisons. Thus, it is important to specify an EquipID that is the closest match to the actual aircraft flown during a given flight. This is done by looking up the tail number of the FOQA flights within the FAA registry website and cross-referencing the specific airframe and engine type with the available options in the AEDT equipment database.

Once input files are created for all core path flights, a new study is created in the AEDT. All airports of interest are added into the study, after which the flights are imported using the SPI tool and the input files.

As this is a comparative study, four alternative configurations of the AEDT are considered in order to get an all-encompassing basis, including choices in both performance and weather models used by the AEDT for modeling. The performance models considered are Base of Aircraft Data (BADA) 3 and BADA 4; the weather models considered are AEDT's built-in average airport weather and high-resolution weather data imported from the National Aeronautics and Space Administration's MERRA-2 dataset. Individual operation groups, annualizations and emissions inventory cases are created for the four configurations. Once created, the cases are run in the AEDT, and results are exported as comma separated values (CSV) files for post-processing and further analysis.

Automation Efforts for Batch Report Tool

Large portions of the AEDT modeling workflow were automated to manage the scale of flights evaluated in the study. The Batch Report Tool was first considered as a potential end-to-end automation solution, since it can generate model runs and export results without manual intervention. However, during initial testing of the version available for AEDT 3e, all flight cases were automatically executed using BADA 3, regardless of the performance-model settings specified by the user. After working with Subject Matter Experts (SME) to identify the underlying cause, it became clear that this version of the tool relied on a fixed internal configuration that ignored study-specific settings and defaulted to BADA 3. Because accurate comparison across BADA 3 and BADA 4 was central to this work, relying on the Batch Report Tool was not feasible.

To address this, the team developed an alternative automation workflow (Figure 31) directly within the AEDT using Structured Query Language (SQL) scripting. The scripts automatically added airports to the study, created operation groups and annualizations, and generated four modeling cases per flight: BADA 3 with airport weather, BADA 3 with MERRA-2, BADA 4 with airport weather, and BADA 4 with MERRA-2. This allowed all performance-model and weather-model combinations to run consistently for every trajectory.

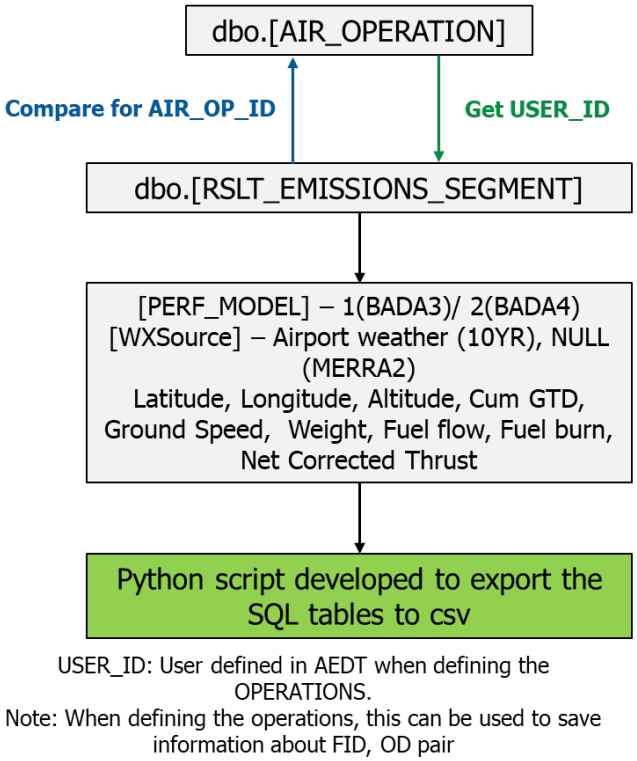


Figure 31. Workflow used to automate the bulk processing and export AEDT results. BADA: Base of Aircraft Data, CSV: comma separate values, FID: Flight ID, GTD: ground track distance, MERRA-2: OD: Origin-Destination, SQL: Structured Query Language.

A custom Python script was then written to extract modeled segment-level results from the SQL database and convert them into CSV files for post-processing, ensuring that flight ID, OD pair, weather source, and performance model were associated properly with each result. This end-to-end workflow enabled reproducible and efficient processing of all the flight cases without relying on the Batch Report Tool.

Results and Discussion

The modeled trajectories generated from the SPI tool were imported into the AEDT and evaluated using both BADA 3 and BADA 4 performance models with MERRA-2 and airport weather datasets. Because FOQA paths contain high-resolution time histories, the team applied a controlled reduction in data frequency before modeling so that the AEDT could process the flights without numerical instability. All analyses used the same filtered input paths, ensuring that the comparison across FOQA, BADA 3, and BADA 4 reflected differences in the performance models and weather assumptions rather than differences in trajectory fidelity.

The cumulative fuel burn is compared between the generated FOQA core paths and the same core paths modeled in the AEDT. The SPI tool was used to import the core paths, after which a fuel consumption study is run, using BADA 3 and BADA 4 models and both the average airport and MERRA-2 weather data. The cumulative fuel burn is computed for the modeled flights in the AEDT, subtracting the taxi-out and taxi-in phases.

As seen in Figure 32, the clustering algorithm does a satisfactory job in identifying grouped flight trajectories and generates clusters based on them. Proceeding on to inspecting the cumulative fuel burn for each city pair, Table 13 shows a comparison of the cumulative fuel burn computed by the AEDT and the fuel burn from the FOQA core path flight data for each city pair.

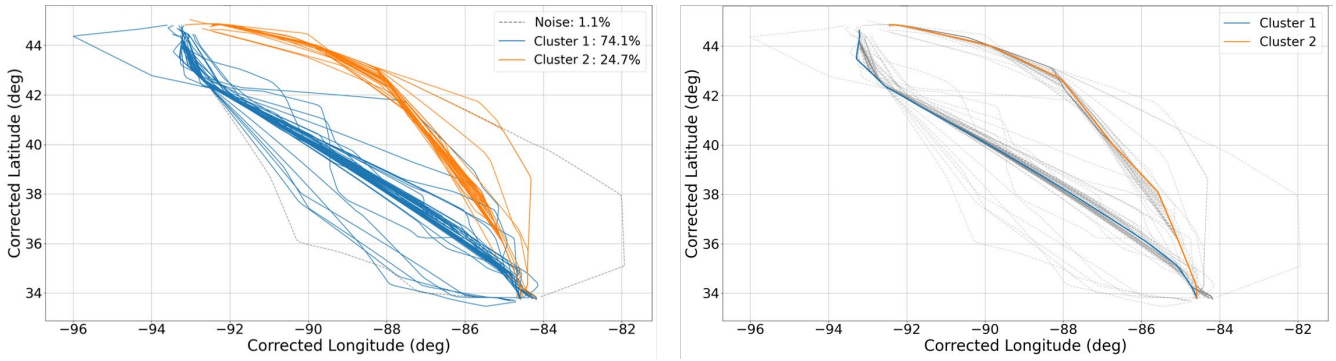


Figure 32. Clusters generated for Hartsfield-Jackson Atlanta International Airport (KATL)- Minneapolis-Saint Paul International Airport (KMSP) (left); chosen core path for the city pair (right).

Table 13. Comparison of cumulative fuel burn between FOQA data and AEDT output for selected city pairs. MERRA refers to the fuel burn calculations done with the MERRA-2 weather data, whereas APT is an identifier indicating the fuel burn calculations using average airport weather data in the AEDT. BADA: Base of Aircraft Data, FOQA: Flight Operations Quality Assurance, GT: Georgia Tech, KATL: Hartsfield-Jackson Atlanta International Airport, KBOS: Boston Logan International Airport, KDEN: Denver International Airport, KDFW: Dallas Fort Worth International Airport, KDTW: Detroit Metropolitan Wayne County Airport, KLAS: Harry Ried International Airport (Las Vegas, Nevada): KLAX: Los Angeles International Airport, KLGA: LaGuardia Airport (Elmhurst, New York), KMCO: Orlando International Airport, KMSP: Minneapolis-Saint Paul International Airport, KRDU: Raleigh-Durham International Airport, KSEA: Seattle-Tacoma International Airport, KSFO: San Francisco International Airport, KSLC: Salt Lake City International Airport, MERRA: Modern-Era Retrospective analysis for Research and Applications, N/A: not applicable.

				BADA 4				BADA 3			
GT Flight ID	**Aircraft Type**	**City Pair**	**Total Fuel Burn FOQA (lb)**	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather
GT2409424	A319-100	KSFO-KLAX	5,249	4,986	5,236	5.02%	0.25%	5,612	6,003	-6.92%	-14.37%
GT2449334	A319-110	KLAX-KSFO	5,195	4,837	4,894	6.89%	5.79%	5,533	5,587	-6.51%	-7.55%
GT2438990	A320-200	KSEA-KSLC	9,789	9,253	9,501	5.47%	2.93%	9,904	10,557	-1.18%	-7.85%
GT2400565	A321-200	KBOS-KATL	15,283	14,583	13,891	4.58%	9.11%	16,262	15.151	-6.41%	0.86%
GT2406987		KDFW-KATL	10,350	10,099	11,897	2.43%	-14.94%	10,962	12,962	-5.91%	-25.24%
GT2407502		KATL-KDFW	12,423	12,216	11,635	1.67%	6.35%	13,034	12,244	-4.92%	1.45%
GT2414166		KATL-KBOS	13,685	13,376	15,009	2.25%	-9.68%	13,800	15,878	-0.85%	-16.03%
GT2419953		KATL-KDEN	19,092	18,265	16,557	4.34%	13.28%	20,901	18,444	-9.47%	3.39%

				BADA 4				BADA 3			
GT Flight ID	Aircraft Type	City Pair	Total Fuel Burn FOQA (lb)	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather
GT2425557		KMSP-KLAS	19,788	20,289	18,117	-2.53%	8.44%	20,289	19,424	-2.53%	1.84%
GT2427137		KATL-KRDU	7,921	7,840	7,786	1.02%	1.71%	8,850	8,908	-11.72%	-12.46%
GT2440699		KDEN-KATL	15,061	15,109	16,852	-0.32%	-11.89%	16,011	17,830	-6.30%	-18.38%
GT2441139		KMSP-KDEN	11,034	10,874	10,696	1.45%	3.06%	11,764	11,519	-6.61%	-4.40%
GT2447364		KATL-KLAX	26,959	26,417	26,755	2.01%	0.76%	27,697	28,390	-2.74%	-5.31%
GT2448788		KDFW-KATL	11,739	11,164	12,306	4.90%	-4.83%	11,739	12,999	0.00%	-10.74%
GT2407293	A321-232	KLAS-KSLC	7,470	8,062	7,263	-7.92%	2.77%	8,062	8,496	-7.92%	-13.73%
GT2421526		KATL-KLAS	26,165	25,464	25,153	2.68%	3.87%	27,236	26,243	-4.09%	-0.30%
GT2421991		KATL-KSLC	28,700	26,963	22,782	6.05%	20.62%	29,870	24,612	-4.08%	14.24%
GT2424237		KATL-KLGA	12,375	12,177	13,160	1.60%	-6.34%	12,948	14,525	-4.63%	-17.37%
GT2425148		KSLC-KATL	19,655	19,953	22,840	-1.52%	-16.20%	20,187	23,698	-2.71%	-20.57%
GT2428735		KATL-KSLC	24,656	24,574	24,320	0.33%	1.36%	24,622	24,408	0.14%	1.01%
GT2429097		KMSP-KATL	14,918	14,860	15,225	0.39%	-2.06%	15,366	15,269	-3.00%	-2.35%
GT2440383		KLGA-KATL	12,039	11,975	12,312	0.53%	-2.27%	12,790	13,189	-6.24%	-9.55%
GT2446243		KSLC-KATL	20,896	20,728	23,967	0.81%	-14.70%	20,734	23,243	0.78%	-11.23%
GT2419261		KDEN-KATL	12,780	12,913	15,020	-1.04%	-17.52%	15,796	17,468	-23.59%	-36.68%
GT2439124		KDEN-KMSP	8,271	8,048	9,755	2.69%	-17.94%	10,402	12,052	-25.76%	-45.71%
GT2410840	B737-800	KSEA-KATL	26,396	25,967	28,412	1.63%	-7.64%	22,998	26,975	12.87%	-2.19%
GT2415788		KATL-KMSP	14,058	13,748	12,971	2.20%	7.73%	13,792	12,583	1.89%	10.49%
GT2423030		KLAS-KSLC	5,522	5,529	5,767	-0.12%	-4.43%	5,411	5,797	2.02%	-4.99%

				BADA 4				BADA 3			
GT Flight ID	Aircraft Type	City Pair	Total Fuel Burn FOQA (lb)	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather	Total Fuel Burn - MERRA (lb)	Total Fuel Burn - APT (lb)	% Error MERRA	% Error APT Weather
GT2423371		KSLC-KLAS	6,301	6,343	6,341	-0.68%	-0.65%	6,040	6,030	4.13%	4.30%
GT2434309		KATL-KMSP	14,571	14,331	12,926	1.65%	11.29%	13,988	12,081	4.00%	17.09%
GT2439960		KATL-KSEA	29,081	29,154	28,602	-0.25%	1.65%	27,299	27,293	6.13%	6.15%
GT2442148		KSLC-KLAS	5,932	5,935	6,074	-0.05%	-2.39%	5,597	5,869	5.64%	1.06%
GT2402954	B737-900	KLAS-KMSP	17,739	17,190	16,959	3.09%	4.4%	16,408	15,769	7.5%	11.1%
GT2404403		KSEA-KMSP	16,760	15,237	17,438	9.09%	-4.04%	15,528	18,593	7.35%	-10.94%
GT2409132		KSEA-KMSP	16,706	15,179	16,758	9.15%	-0.31%	15,319	17,324	8.31%	-3.70%
GT2443552		KMSP-KSEA	18,590	17,796	16,774	4.27%	9.77%	17,796	17,146	4.27%	7.77%
GT2447664		KDTW-KATL	9,959	9,174	8,557	7.88%	14.08%	9,674	8,730	2.86%	12.34%
GT2449664		KSLC-KSEA	10,445	9,747	9,507	6.69%	8.99%	10,167	9,821	2.67%	5.98%
GT2404615	B757-200	KATL-KDTW	11,324	10,765	11,135	4.94%	1.67%	10,744	11,192	5.13%	1.17%
GT2447376		KRDU-KATL	7,967	7,828	7,442	1.74%	6.59%	7,776	7,309	2.40%	8.26%
GT2447839		KATL-KMCO	8,416	8,458	8,716	-0.50%	-3.56%	7,905	8,367	6.07%	0.58%
GT2417499	B757-300	KATL-KSEA	38,179	39,578	39,781	-3.66%	-4.19%	N/A	N/A	N/A	N/A
GT2437635		KLAS-KATL	29,157	30,150	32,084	-3.41%	-10.04%	N/A	N/A	N/A	N/A
GT2434848		KMCO-KATL	10,435	11,371	10,033	-8.97%	3.85%	N/A	N/A	N/A	N/A

Across the evaluated flights, the AEDT-modeled fuel burn shows a consistent pattern of acceptable deviation from FOQA values, with performance depending on both the aircraft type and weather source. BADA 4 generally tracks FOQA fuel burn more closely than BADA 3, which tends to under- or over-estimate consumption depending on the route. BADA 4 showed the closest overall agreement, especially when paired with MERRA-2, which tended to minimize bias across climb, cruise, and descent segments. BADA 3 performed acceptably for most short- and medium-haul routes but displayed larger deviations for flights where strong meteorological gradients played a role. MERRA-2 based runs consistently show lower error magnitudes than the APT weather cases, although certain routes with strong winds or altitude profile mismatches show larger negative biases, especially for long-haul segments such as KDEN–KATL and KDEN–KMSP. In several short-haul pairs, including KSLC–KLAS and KLAS–KSLC, the modeled fuel aligns closely with FOQA, with errors often below 1%. Larger

discrepancies appear for flights with greater vertical variability or complex meteorological conditions, where APT weather occasionally produces over-predictions exceeding 10%. Overall, the comparison indicates that the AEDT combined with BADA 4 and MERRA provides the most stable agreement with FOQA data, while still exhibiting route-specific deviations that highlight the sensitivity to weather and climb–descent behavior.

The cumulative-fuel plots (Figure 33 and Figure 34) illustrate this behavior clearly: the AEDT tracks FOQA well at low altitudes, while cruise-phase differences accumulate gradually, leading to over- or under-estimation depending on the weather source. MERRA-2 usually produced smaller overall errors than airport weather, where several long-range flights showed persistent negative biases tied to missing instantaneous local weather information.

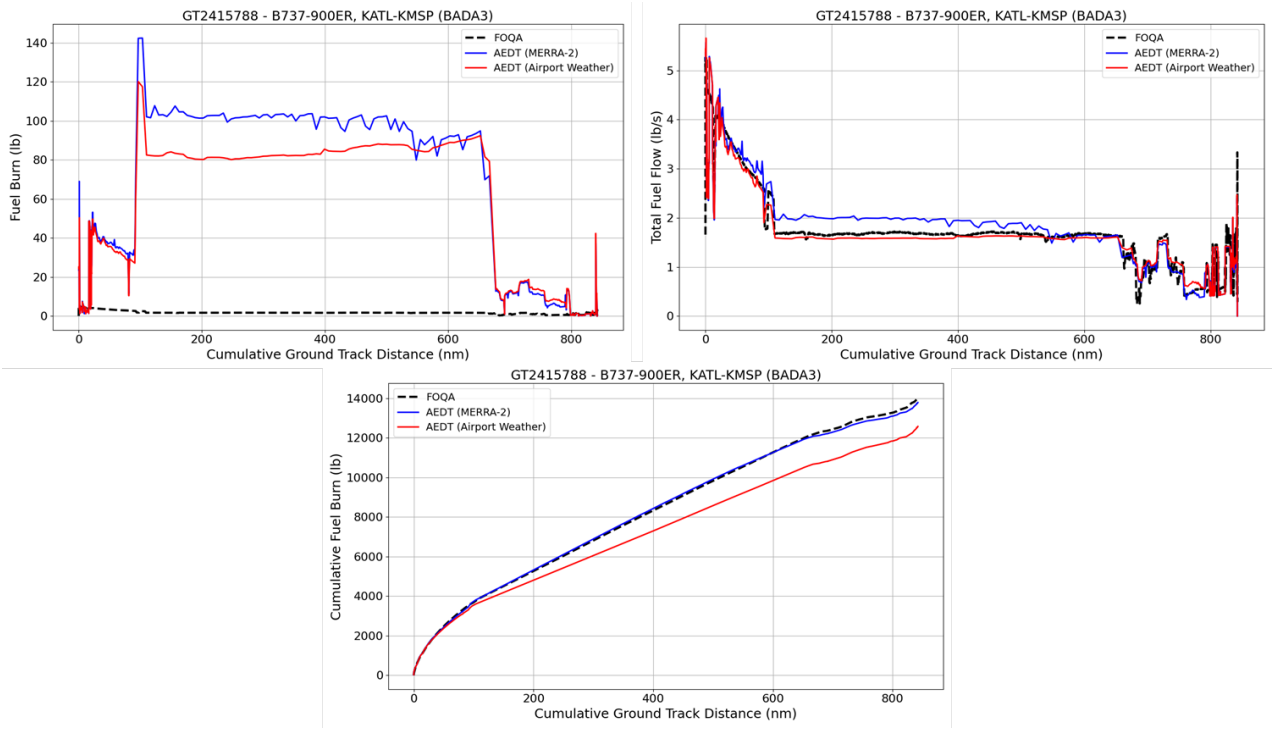


Figure 33. Fuel-burn and fuel-flow comparison for flight GT2415788 (B737-900ER, Hartsfield-Jackson Atlanta International Airport [KATL]– Minneapolis-Saint Paul International Airport [KMSP]) modeled in the AEDT using the Base of Aircraft Data (BADA) 3 performance model. FOQA: Flight Operations Quality Assurance.

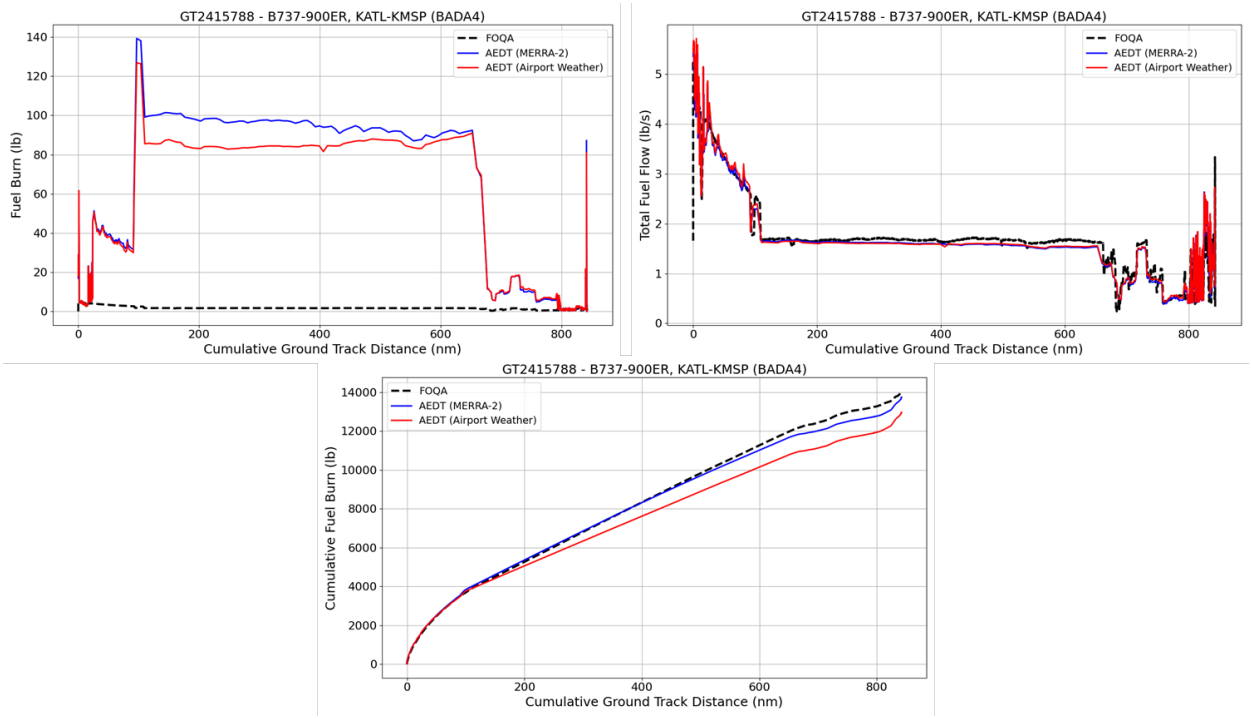


Figure 34. Fuel-burn and fuel-flow comparison for flight GT2415788 (B737-900ER, Hartsfield-Jackson Atlanta International Airport [KATL]– Minneapolis-Saint Paul International Airport [KMSP]) modeled using the Base of Aircraft Data (BADA) 4 performance model.

The representative case for the B737-900ER (KATL–KMSP) (Figure 33 and Figure 34) show how both models capture major performance features. In both the figures, the top-left image shows instantaneous fuel burn along the cumulative ground-track distance for FOQA, AEDT with MERRA-2 weather, and AEDT with airport weather. The top-right image presents the corresponding total fuel-flow profiles. The bottom image shows cumulative fuel burn, illustrating how differences in cruise-level fuel flow and weather inputs propagate over the full flight. The AEDT with MERRA-2 tracks FOQA more closely than the airport-weather case, though both capture the major climb, cruise, and descent features. The AEDT reproduces FOQA's climb-spike, stabilizes through cruise, and follows the same curvature in cumulative burn, even with reduced data resolution. BADA 4 narrows the gap further and maintains closer alignment with FOQA fuel-flow rates through the cruise segment, while BADA 3 shows a slightly larger spread. Overall, the results show that the AEDT combined with BADA 4 and MERRA-2 provides the most consistent agreement with FOQA, and that model-to-FOQA differences are driven more by route-specific atmospheric conditions and thrust modeling than by trajectory preprocessing.

SPI Tool Issues Reported

During the modeling process, several limitations in the SPI tool and its interaction with the AEDT were identified. These issues affected how FOQA trajectories could be imported and how thrust and fuel-burn behavior were represented during simulation. Each issue and the corresponding mitigation strategy are summarized below.

- Trajectory Size Limitation in SPI Imports: The SPI tool was unable to import FOQA trajectories containing more than roughly ten thousand points. Several flights in the dataset exceeded this threshold because FOQA operates at a one-second resolution throughout the mission. When these large files were passed to the AEDT, the import process failed before modeling could begin.
- Negative Thrust Values in BADA 4 Cruise and Descent: When full-resolution FOQA data (after reducing the points to be below 10,000) are input in the SPI tool, the AEDT consistently produced highly negative thrust values during cruise and parts of descent for BADA 4 cases. This behavior did not match FOQA and is not physically meaningful for commercial aircraft operating in steady cruise.

Cause Identified: The AEDT has difficulty processing extremely high-resolution (1 Hz) input, especially when paired with BADA 4, which uses a more detailed thrust model. Small fluctuations in altitude and ground speed at high resolution can produce unstable thrust estimates.

Mitigation applied: The team introduced a controlled down-sampling routine to reduce trajectory resolution while preserving high-value segments. Points below 10,000 ft were retained at full resolution to maintain climb and final-approach fidelity. Points between 10,000 and 30,000 ft were kept at 12-s intervals, and cruise-level points above 30000 ft were sampled at 60-s intervals. After the trajectory sampling was reduced, negative thrust values in cruise disappeared. Some negative values continued to appear in the descent phase, but these were isolated, moderate in magnitude, and did not affect cumulative fuel-burn results. The comparison of this reduced sample with the original is shown in Figure 35.

- Limitations of BADA 3 for Certain Aircraft Types: B757-300 and B767-400 flights could not be modeled using BADA 3 because corresponding performance models do not exist within that version of the database. These aircraft were therefore excluded from BADA 3 comparisons.
- Takeoff-Weight Adjustments Not Reflected in BADA 3: Adjusting aircraft takeoff weight had no effect on BADA 3 fuel-burn results. This is because BADA 3 does not incorporate takeoff weight into its thrust or drag formulations.
- Manual Assignment of Equipment IDs: The SPI tool requires each input trajectory to be linked to a BADA aircraft model via an equipment ID. This required manual editing of input file for all the flights (as mentioned in the Methodology section).

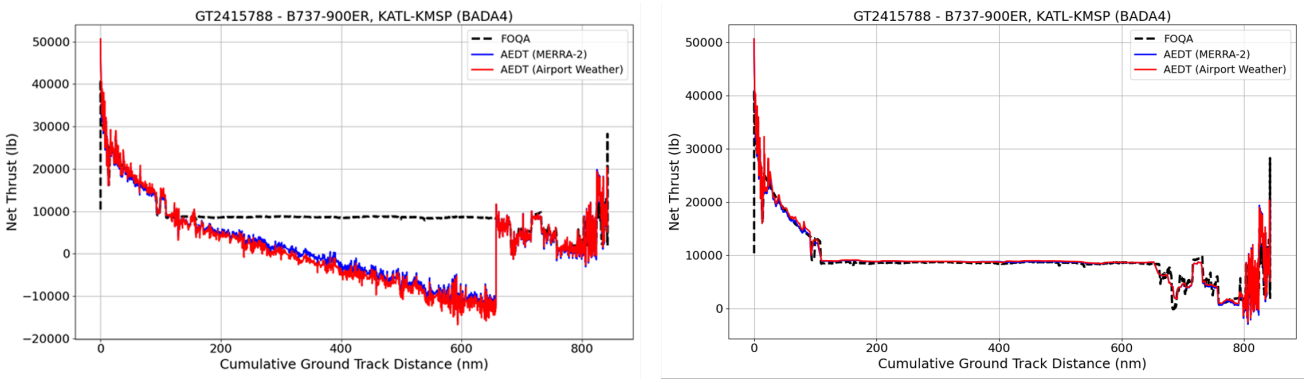


Figure 35. Net-thrust comparison for flight GT2415788 (B737-900ER, Hartsfield-Jackson Atlanta International Airport [KATL]– Minneapolis-Saint Paul International Airport [KMSP]) modeled with Base of Aircraft Data (BADA) 4. The left image shows results using full-resolution (1Hz) FOQA input. The right image shows the same flight after applying reduced-resolution trajectory sampling.

Milestones

None.

Major Accomplishments

- Conducted a correlation analysis of en-route fuel burn using real-world flight data, including the TTD and FOQA datasets.
- Developed linear regression/machine learning models to predict fuel burn using airport and aircraft performance features.

Publications

None.

Outreach Efforts

None.

Awards
None.

Student Involvement
Graduate research assistants, Hyungu Choi, Nitya Maruthuvakudi Venkatram, and Yash Vinod from the School of Aerospace Engineering participated in this task. Hyungu was involved with the creation of surrogate models for fuel burn prediction. Nitya and Yash performed the analysis for the clustered core path generation.

Plans for Next Period
Documentation and closeout of work performed under Task 3.

Task 4 – System Testing and Evaluation of AEDT
Georgia Institute of Technology

Objective
To provide the best possible environmental impact modeling capabilities in the AEDT, the FAA Office of Environment and Energy (OEE) continues to develop the AEDT by improving existing modeling methods and data and adding new functionalities. The FAA OEE seeks an independent effort in system testing to evaluate the accuracy, functionality, and capabilities of the AEDT, and support future model development. The objective of this task is to provide the FAA with high-quality systematic testing and evaluation of the capabilities of the AEDT, and to identify gaps in the tools' functionality and areas for further development.

Research Approach
Modeling Noise with Mixed Ground Types, Terrain, and Manmade Structures
New Mixed Ground Types, Terrain, and Manmade Structures (MGT/S) features, which include capabilities for enhanced terrain shielding, creation of various manmade structures and land cover definition, are being implemented for a future release of the AEDT. Thus, testing for these features commenced with focus primarily towards proper functioning of the features in terms of expected results and usability. Additionally, feature runtime was investigated, and user feedback was given in order to improve the aforementioned capabilities.

The structural creation feature allows for generation of building rows, individual buildings, and sound barriers, each of which was utilized individually and in combination in order to capture noise reflection and diffraction effects. For the study, several annualizations were created for modeling arrival and departure operations for the Brainerd Lakes Regional Airport (KBRD) in Brainerd, Minnesota (see Figures 36, 37, and 38). The annualizations included various structural combinations with and without land cover and terrain data.

The results showed that accuracy was only impacted by land cover data. The various structure types individually and in concert with each other without land cover data showed no difference in the shape of the noise contours, whereas inclusion of land cover data resulted in substantial differences for the noise contours. One of the tests included a large building far from the runway and showed an issue with the 25 dB contour being inside of the building rather than outside like all of the other structural tests without land cover. This finding was reported and was investigated further outside of the testing phase.

Runtime was investigated by examining start and end times for each case within the AEDT study logs. Findings showed that cases in which land cover was incorporated resulted in higher runtimes. Quantity of structures had little effect on runtime across cases; however, some variability still exists for runtimes within the same annualization. As in, if a case was run multiple times, different runtimes were shown. All findings were reported and backup files and logs were delivered as part of the results package.

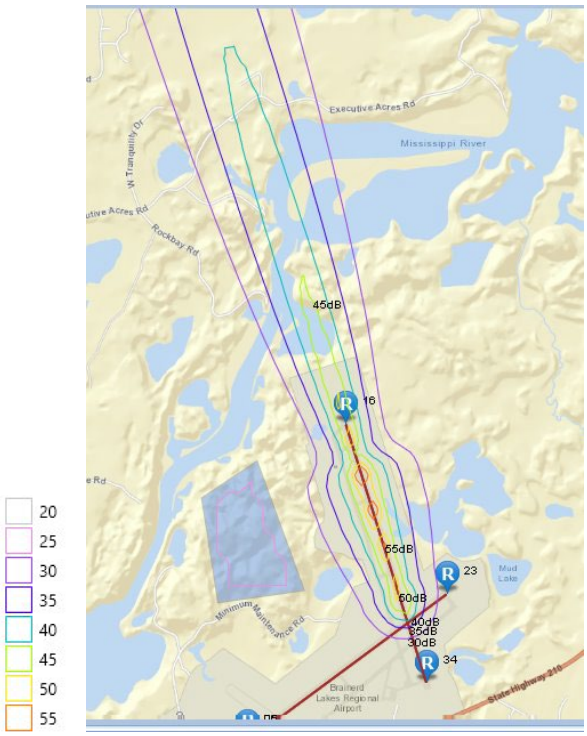


Figure 36. Large building far from runway (no land cover) at Brainerd Lakes Regional Airport.

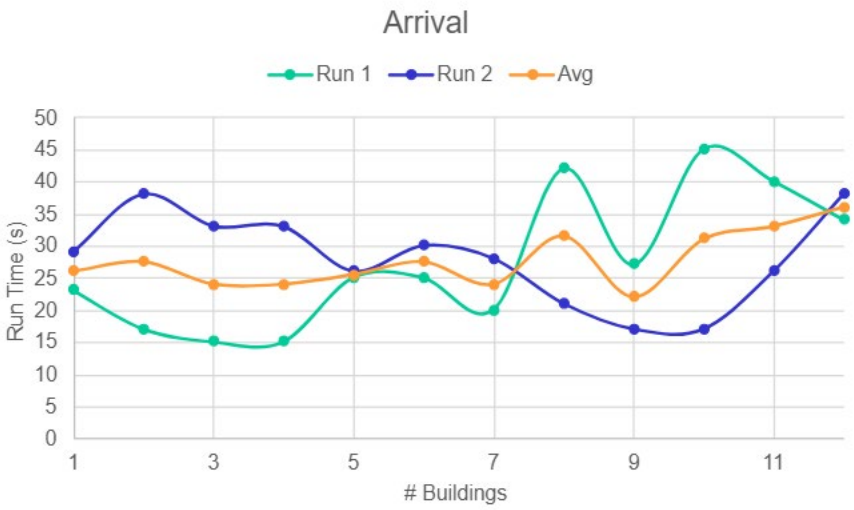


Figure 37. Runtime for arrival operations vs. building quantity.

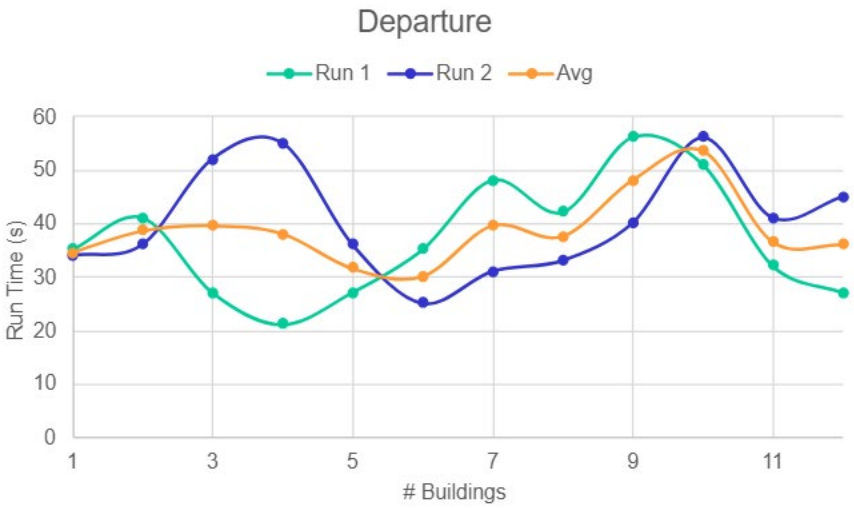


Figure 38. Runtime for departure operations vs. building quantity.

Supersonic Aircraft Modeling Testing
New SST aircraft modeling capabilities were implemented into the AEDT which required testing to ensure proper functionality of said features. The test consisted of six airport pairs with runway-to-runway operations and supersonic aircraft definitions. A user guide was provided with steps to follow in order to reproduce initial testing results and verify proper functioning of each step in the testing process.

One of the issues seen during the testing was in the SPI file import of the airport pair AEDT Standard Input File (ASIF) files in step 8 and its sub-steps. Partial sensor import was achievable in step 8.2 only after first changing the anpAircraftID in the provided .xml files. Prior to this, "unable to parse data" or "bad data integrity" errors would arise. Changing the anpAircraftID to be consistent with what was established in the GUI for the supersonic aircraft definition solved the partial sensor import issue for all airport pairs except for the Dubai International Airport (OMDB) - Suvarnabhumi Airport (VTBS) pair. This airport pair had a separate issue where two different airport codes were present in the SQL definition for one of the airports. By going in and deleting the older airport code, the problem was resolved. The SPI tool import was still an issue so proceeding was only possible by restoring the provided backup study with all sensor path operations already imported.

Other minor issues were documentation-related, such as step 9 of the document being misplaced due to the fact that the "tracks" option doesn't show up until a metric result has been defined. Another issue was that operation groups must be established prior to partial sensor path importing in step 8.2 for proper functioning. This information was missing from the documentation, and thus a recommendation to include this information was given. The final recommendation given was to create default airport settings for heavily used airports and airport pairs for enhanced ease of use when conducting studies.

Milestones
None.

Major Accomplishments
None.

Publications
None.

Outreach Efforts
None.

Awards
None.

Student Involvement
Graduate research assistants, Santusht Sairam and Fahraan Badruddin from the School of Aerospace Engineering participated in this task. Santusht and Fahraan were involved with ideation and modeling of test cases, analysis of results, and creation of any relevant bug reports.

Plans for Next Period
Continue system testing efforts to support ongoing AEDT development.

Task 5 – Supersonic Transport Modeling in the Aviation Environmental Design Tool
Georgia Institute of Technology

Objectives
The SST modeling task has two primary objectives:

1. Develop a predictive model for the propulsion system, encompassing thrust, fuel flow, engine dimensions, weight, and noise generation, and additionally accounting for variations in Mach number, altitude, and throttle setting, to enable comprehensive analysis of performance metrics such as fuel burn, gross weight, takeoff field length, and environmental factors including noise and emission.
2. Generate parametric drag polar models for the optimized vehicle that capture aerodynamic performance across the entire flight envelope as a function of Mach number, altitude, and lift coefficient.

Research Approach
Building on the ASCENT SST surrogate-model development implemented previously, the same end-to-end modeling workflow has been retained for this term's extension. The earlier study focused on validating regression-based propulsion (fuel flow) and aerodynamic (drag coefficient) models against Flight Optimization System (FLOPS) for six long-range international routes, and demonstrated that the methodology with continuous neural-network-based fuel-flow models and Total Energy Method (TEM)-based thrust estimation substantially reduced prediction errors, with final fuel-burn deviations remaining within approximately 6% across all evaluated origin-destination pairs. The combined method, which integrates the updated fuel-flow surrogate, TEM-based thrust estimation, and piecewise aerodynamic regression with flight-phase-specific logic (including dedicated descent thrust handling) is as shown in Figure 39.

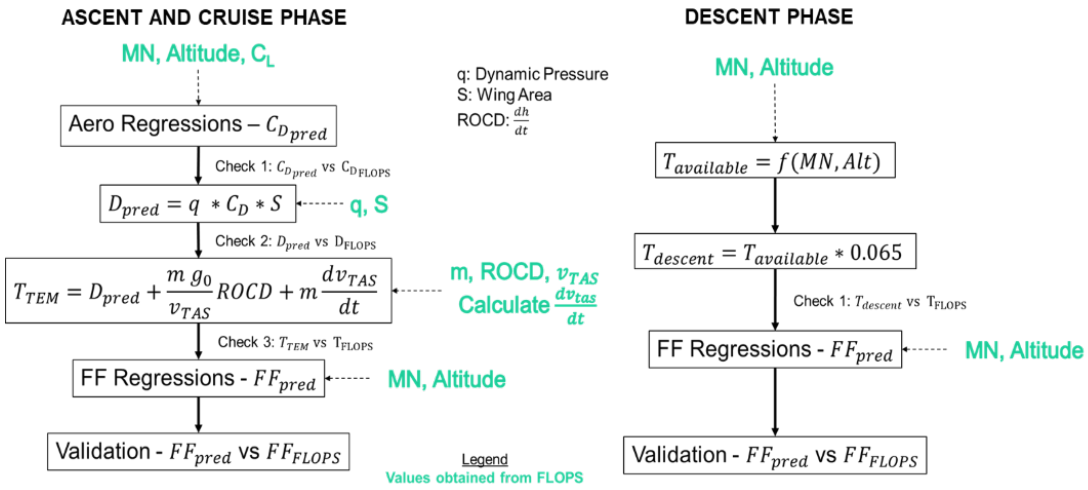


Figure 39. Modeling approach for ascent, cruise, and descent phases.

A brief summary of the implementation logic is as below:

1. **Fuel-Flow Regression Model:** Neural-network surrogate (2-layer Gaussian, 10 neurons each, LR = 0.01), trained on engine-deck data and validated against route-specific FLOPS outputs.
2. **Thrust Estimation Logic**: TEM formulation used to compute thrust directly from climb/descent energy balance and drag regressions.
3. **Aerodynamic Regression Model**: Drag coefficient modeled through regime-specific Mach–C_L sub-domains using cardinal grids and bi-quadratic interpolation, removing previous domain discontinuity artifacts.
4. **Descent-Phase Special Handling:** Thrust available predicted via polynomial regression with fixed idle-thrust factor applied, replacing TEM during descent to prevent over-prediction at low thrust conditions.

Because the foundational methodology has already been verified and documented, the current work focuses exclusively on the new route evaluation, applying the same modeling pipeline with no codebase or logic alterations. Trajectory discretization, weight propagation, altitude-Mach scheduling, and detailed takeoff and landing profiles followed the same implementation assumptions described in the earlier report. As before, detailed takeoff was modeled up to 15,000 ft and detailed landing from 3,333 ft to touchdown to ensure smooth physical continuity into and out of the surrogate-modeled cruise segment. FLOPS mission results were used as the validation reference for both total mission fuel burn as well as point-wise fuel-flow and thrust behavior along the ground-track distance.

A new city pair has been evaluated between KJFK and KLAX. Unlike the previously studied intercontinental missions, this route represents a domestic transcontinental scenario. In order to investigate performance trade-offs between design performance intent and practical regulatory compliance, each direction was modeled under two distinct cruise conditions as shown in Table 14.

Table 14. Modeling cases for SST route between John F. Kennedy International Airport (KJFK) and Los Angeles International Airport (KLAX).

Case	Direction	Cruise Mach	Motivation
A	KJFK → KLAX	1.70	Demonstrates peak design mission performance
B	KJFK → KLAX	1.15	Demonstrates regulatory-compliant domestic operation
C	KLAX → KJFK	1.70	Design Mach under prevailing tailwind conditions
D	KLAX → KJFK	1.15	Regulatory-compliant eastbound mission

In the previous phase of this study, the Supersonic Transport Route Tool (SSTRT) was used to optimize the mission routes by balancing time savings against fuel burn using a tunable weighting parameter α. That optimization framework allowed for over-water supersonic segments whenever possible to satisfy regulatory limits. The SSTRT therefore generated custom mission trajectories for each origin-destination pair, and those trajectories fed into the fuel-flow surrogate modeling workflow described earlier.

In the present phase, because the selected test route (KJFK–KLAX, bidirectional) flies entirely over continental landmass, supersonic cruise is not allowed under current FAA/ICAO regulations. As a result, no route optimization was performed, and the reference trajectory was defined using the Great Circle Distance (GCD). These routes are as shown in Figure 40. The Latitude-longitude for the FPPs are interpolated using the Great Circle path,

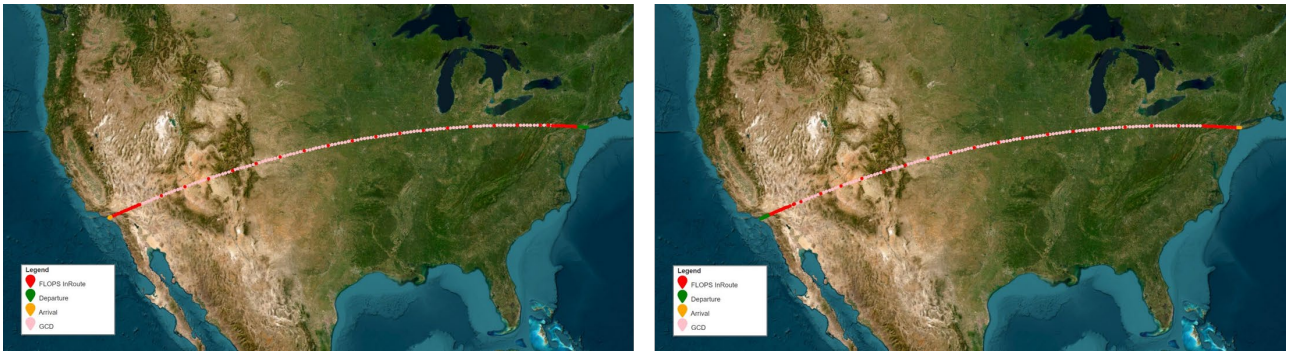


Figure 40. Route Visualizations - KJFK→KLAX (left) KLAX→KJFK (right).

Results and Discussion

The overall mission-level error, as shown in Table 15, remains within acceptable limits for both conditions, consistent with previous results. The surrogate appears slightly optimistic for Mach 1.7 and slightly conservative for Mach 1.15, but both remain within the target error window (< 3%). These errors are calculated without ignoring the duplicate points between the flight segments. FLOPS is path independent (directionally) and uses only the distance between the OD pair (GCD in this case) and airport latitude-longitude coordinates.

Table 15. Total fuel burn predictions comparison. FLOPS: Flight Optimization System, KJFK: John F. Kennedy International Airport, KLAX: Los Angeles International Airport.

Route	Actual (lbm)	FLOPS-Calculated (lbm)	% Error	Predicted (lbm)	% Error
KJFK ↔ KLAX (M1.7)	66,491.00	66,185.91	0.46%	65,566.38	1.39%
KJFK ↔ KLAX (M1.15)	70,868.00	70,526.18	0.48%	72,827.00	−2.76%

Plots of error in thrust and fuel-flow rate versus distance (Figure 41) indicate that all non-transient points remain below approximately 8% error. Outliers occur only at (1) the first data point, and (2) a duplicated point between segments associated with continuity (an artifact of FLOPS) (as can be seen the altitude plot on Figure 42).

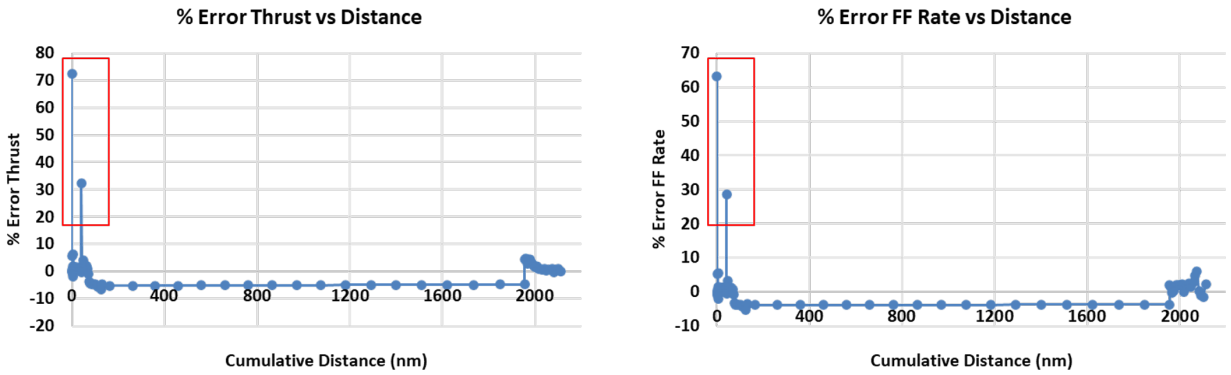


Figure 41. Percent error thrust (left) and percent error fuel flow (FF) rate (right) vs. ground track distance (nmi).

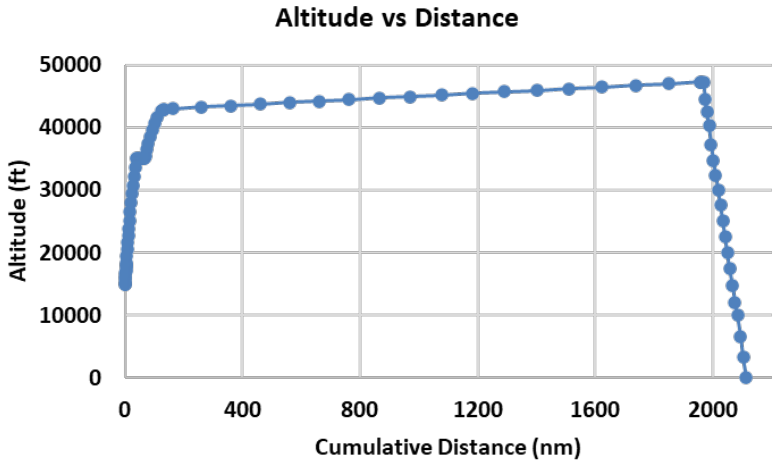


Figure 42. Altitude (ft) vs. ground track distance (nmi).

Histograms comparing raw versus de-duplicated error (Figure 43) confirm that other than the first point, the duplicate point, all other points have error lower than 8%.

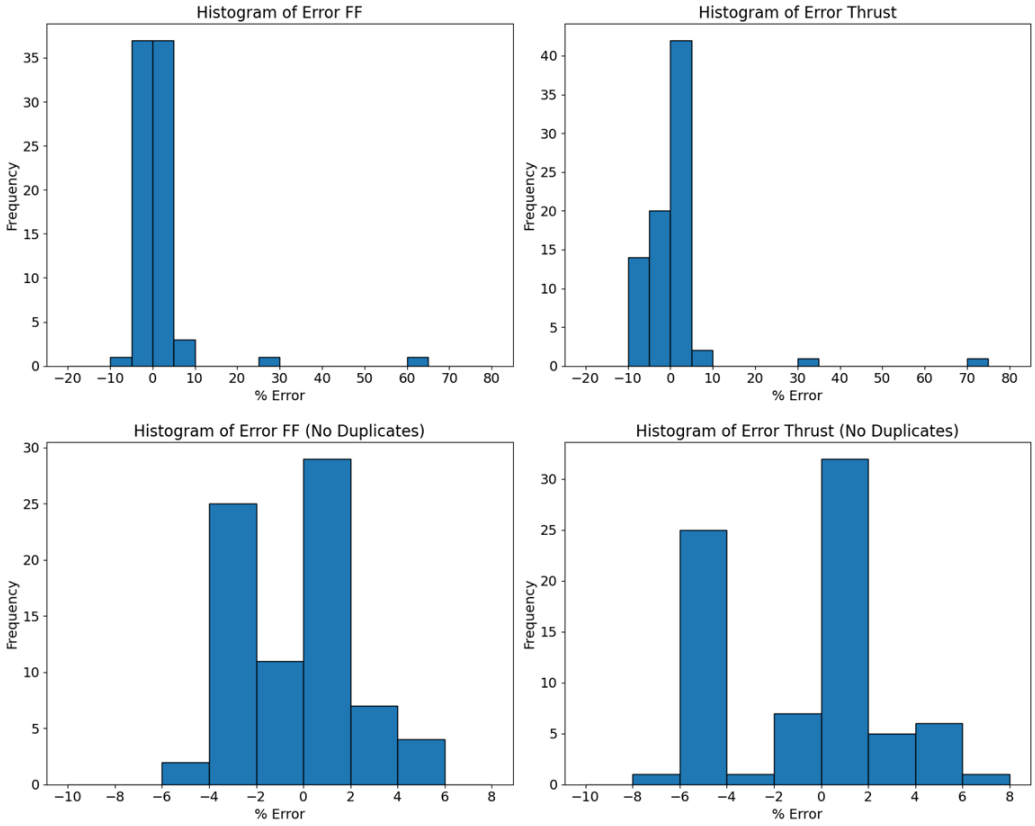


Figure 43. Raw and de-duplicated error histograms for fuel flow (FF) rate and thrust.

Detailed Takeoff and Landing Segments

Takeoff transition to the cruise mission matched closely, demonstrating good continuity between FPP and regression-based segments (Figure 44). Differences in detailed landing segments (Figure 45) are attributed to configuration modeling differences in FLOPS. FLOPS detailed landing assumes gear and full flaps from transition altitude, whereas the sizing run retains near-cruise aerodynamic configuration, leading to temporary true airspeed divergence. This behavior is consistent with earlier findings and stems from FLOPS internal modeling rather than the surrogate itself.

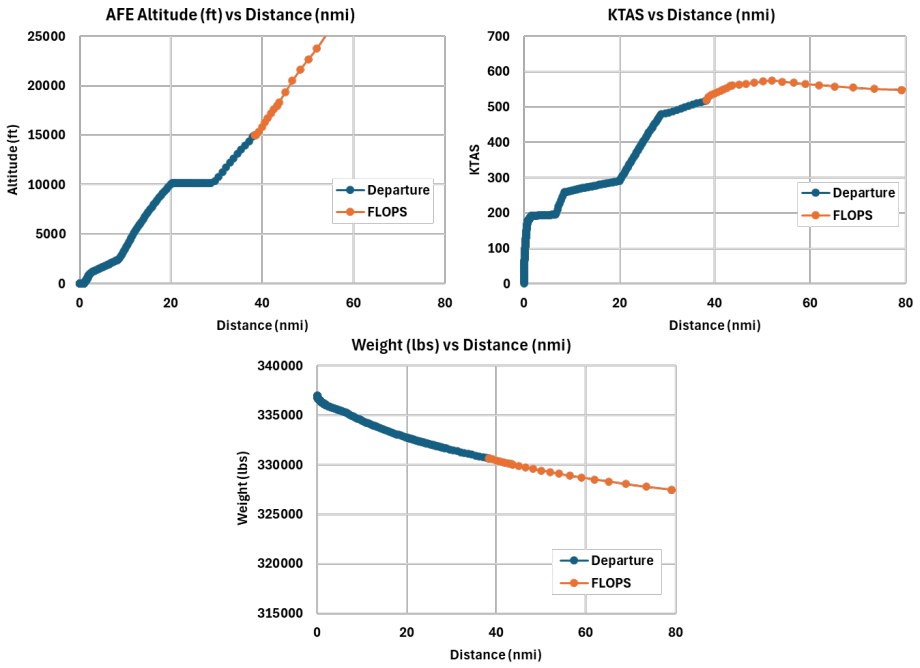


Figure 44. Detailed take-off segments (FPP) till 15,000 ft above field elevation (AFE). KTAS: knots true airspeed, FLOPS: Flight Optimization System.

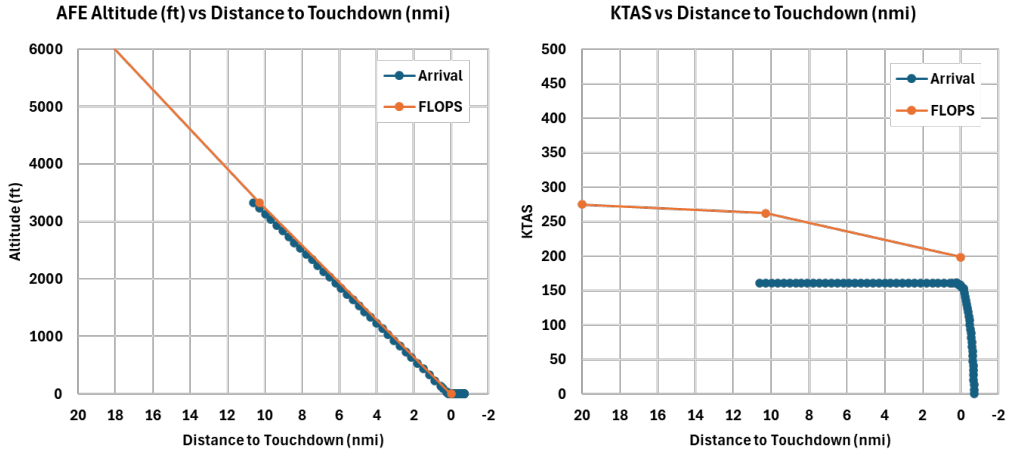


Figure 45. Detailed landing segment (FPP) from 3,333 ft above field elevation (AFE). KTAS: knots true airspeed.

To conclude, the surrogate-based fuel-flow modeling approach continues to demonstrate robust predictive behavior even when applied to land-constrained. Mission-level fuel burn accuracy remained within ±3%, and point-wise error remained below ±8% outside initialization transients.

Milestones

None.

Major Accomplishments
Complete Data Package Handover

Publications
None.

Outreach Efforts
Participated in biweekly calls and attended biannual ASCENT meetings.

Awards
None.

Student Involvement
Graduate research assistant, Nitya Maruthuvakudi Venkatram from the School of Aerospace Engineering participated in this research. Nitya was involved with the generation of the propulsion and aero models, modeling of test cases, and analysis of results.

Plans for Next Period
None.