

Towards Fuel Consumption Prediction in Trombone Procedures Using ADS-B Data

Lidia Serrano-Mira ^{*,1} Eduardo S. Ayra,^{1,2} Luis Pérez Sanz ¹ Javier A. Pérez-Castán ¹ and Raúl Alberto Linta¹

¹Universidad Politécnica de Madrid, Madrid, Spain

²Agencia Española de Seguridad Aérea (AESA), Madrid, Spain

*Corresponding author: lidia.serrano@upm.es

Abstract

Reducing uncertainty about arrival trajectories before departure could help airlines plan fuel more efficiently. However, tactical air traffic control (ATC) decisions make the distance flown in trombone procedures difficult to anticipate. This study aims to reduce uncertainty in pre-departure estimates of the distance flown during trombone procedures through probabilistic modeling, thereby supporting operational flight plan (OFP) preparation. Historical ADS-B trajectories from Barcelona–El Prat Airport (LEBL), collected mainly through the OpenSky Network between March 2025 and March 2026, are combined with ATS flight plan, traffic-demand and weather data. Supervised machine learning (ML) classification and regression models are first evaluated, achieving a maximum accuracy of 60.4% and a minimum mean absolute error (MAE) of 7.09 NM, respectively. A stochastic hybrid automaton guided by ML models is then developed to simulate aircraft trajectories, surrounding traffic and tactical ATC decisions. Monte Carlo simulations are used to propagate operational uncertainty through the hybrid automaton, yielding estimated probability distributions of the distance flown. On the test dataset, the automaton achieves an MAE of 5.38 NM, with 80% and 90% prediction intervals covering the actual distance in 80.4% and 90.0% of flights, respectively. A theoretical fuel-planning assessment using the EUROCONTROL BADA model suggests potential reductions of 55–85 kg in planned fuel per flight for the aircraft models studied, depending on the operational conditions and assumptions considered, with possible additional cost-to-carry savings. These findings suggest that the proposed approach offers a promising way to reduce pre-departure uncertainty about the distance flown during trombone procedures. Translating these probabilistic distance estimates into fuel consumption estimates could help reduce uncertainty in fuel calculations during OFP preparation, although operational implementation would require airline-specific aircraft performance models and forecasts of weather and traffic conditions.

Keywords: ADS-B; Fuel Consumption; Trombone Procedures; Probabilistic Modeling

Abbreviations: ATCO: Air Traffic Controller, ATC: Air Traffic Control, OFP: Operational Flight Plan, ML: Machine Learning, FPL: ATS Flight Plan, LEBL: Barcelona-El Prat Airport, TMA: Terminal Maneuvering Area, IAF: Initial Approach Fix, IF: Intermediate Fix, NM: Nautical Mile

1. Introduction

The aviation sector faces the dual challenge of accommodating the increasing demand for air traffic while reducing its environmental impact. The International Civil Aviation Organization (ICAO) has established a target of net-zero carbon dioxide (CO₂) emissions by 2050 for international aviation [1]. Reducing fuel consumption through more efficient operations represents one of the contributions towards achieving this goal.

Arrival operations have been identified as a potential source of operational inefficiencies, as traffic sequencing and congestion can result in additional distance traveled, flight time, fuel consumption, and local emissions [2]. As traffic concentration around major airports continues to increase, arrival

management must therefore accommodate sequencing and capacity requirements while limiting unnecessary trajectory extensions and their associated fuel impact. In this context, Performance-Based Navigation (PBN) has enabled the implementation of structured arrival procedures aimed at facilitating traffic sequencing while improving trajectory predictability and operational efficiency. Among these, Point Merge and trombone procedures provide Air Traffic Controllers (ATCOs) with route flexibility to accommodate delay and adjust the distance flown according to sequencing needs [3, 4].

Point Merge is an arrival sequencing method based on one or more arc-shaped sequencing legs approximately equidistant from a common merge point. Aircraft may remain on these legs to absorb the delay required for sequencing and are subsequently cleared “direct to the merge point” once adequate separation with the preceding aircraft has been achieved. In contrast, trombone procedures are PBN-based arrival routes consisting of parallel route segments and a sequence of predefined waypoints. Aircraft are expected to follow the published route unless the ATCO issues a “direct” to a waypoint [4]. This allows ATCOs to adjust the distance flown according to the separation and sequencing needs. The flexibility provided by trombone procedures is particularly useful for arrival sequencing, but it also introduces variability in the distance actually flown by each aircraft. Depending on traffic conditions and tactical Air Traffic Control (ATC) interventions, an aircraft may receive an early shortcut or fly a substantially larger part of the published procedure. Consequently, the trajectory (and therefore the fuel consumption associated with the arrival) is not fully known at the pre-flight planning stage.

Both procedures have been investigated in the literature from different operational perspectives. Regarding Point Merge, early studies [5, 6] focused on its ability to systematize arrival sequencing and reduce the dependence on radar vectoring. Subsequent research has investigated the integration of PM with continuous descent operations (CDOs), as well as its impact on trajectory efficiency, considering aspects such as distance flown, flight time, fuel consumption, and emissions [7, 8, 9]. In contrast, the literature specifically addressing trombone procedures is more limited. Sáez et al. [10] investigated the use of trombone structures for arrival traffic synchronization and the implementation of continuous descent operations. More recently, Liu et al. [4] compared the performance of PM and trombone procedures under different traffic demand conditions. Previous research has mainly focused on the operation and efficiency of trombone procedures, while less attention has been paid to uncertainty in the distance aircraft will fly when following these procedures. In particular, the extent to which pre-departure information can reduce this uncertainty remains an open question.

Reducing this uncertainty before departure is particularly relevant to airline fuel planning. Commercial airlines prepare Operational Flight Plans (OFPs) for their scheduled flights, generally using flight planning systems that estimate fuel requirements based on the planned route, forecast weather conditions, aircraft performance, and airline operational policies. These OFPs are generated systematically for a large number of flights and are based on assumptions about the expected departure and arrival procedures, which may differ from those actually flown. Therefore, on the day of operation, the flight crew must verify the validity of the planned route and the fuel calculations, and the captain may decide to carry additional fuel if the amount specified in the OFP is considered insufficient for the expected operating conditions.

In this context, the European Union Aviation Safety Agency (EASA) guidance [11] specifies that, for a STAR or transition that includes a trombone procedure, the fuel required for the reasonably expected route should be included in the trip fuel calculation. Additional fuel for a possible extension of the trombone may be included in contingency fuel or, when a delay is anticipated, in extra fuel.

However, carrying additional fuel increases aircraft mass, resulting in additional fuel burn (“carrying fuel to carry fuel”), emissions, and operating costs. In 2016, Hao et al. [12] analyzed the impact of

this “cost-to-carry” effect, estimating that fuel loaded due to operational unpredictability may result in additional costs of \$14–\$54 per flight (equivalent to approximately \$120–\$452 million per year in domestic operations in the US). These findings highlight the potential economic implications of improving fuel planning under operational uncertainty.

Accordingly, the goal of this work is to assess whether pre-departure information can be used to reduce uncertainty in estimating the distance flown by aircraft during the approach phase in procedures involving trombone maneuvers using probabilistic models. Rather than providing a deterministic prediction for individual flights, the proposed approach aims to characterize the probability distribution of different trombone extensions across a set of flights, providing information that could support the large-scale preparation of OFPs at the pre-dispatch stage, before the flight crew makes its final decision on the amount of fuel to be carried. To this end, historical ADS-B trajectories are used to characterize trombone operations, first evaluating the limitations of deterministic prediction approaches and subsequently developing a probabilistic framework based on information available before departure. Finally, the study investigates how uncertainty in the distance flown translates into variability in fuel consumption, assessing its potential implications for flight planning.

This paper is organized as follows. Section 2 describes the case study and ADS-B dataset. Section 3 presents the methodology and results of the deterministic and probabilistic approaches to predicting the distance flown during arrival procedures involving trombone maneuvers, including the stochastic hybrid automaton. Section 4 evaluates the implications for fuel consumption and flight planning. Section 5 discusses the main limitations of the study. Finally, Section 6 presents the conclusions and future research work.

2. Case Study and ADS-B Data

2.1 Barcelona–El Prat Scenario

Barcelona–El Prat Airport (LEBL) is the second busiest airport in Spain in terms of air traffic, with a total of 360,786 aircraft movements in 2025 [13]. Its high traffic demand, particularly during certain periods of the day, requires the use of sequencing procedures to efficiently manage arrival flows into the terminal maneuvering area.

In 2018, as part of the first phase of ENAIRE’s BRAIN project [14], the implementation of trombone procedures began in the Barcelona Terminal Maneuvering Area (TMA), with the aim of sequencing arrivals and optimizing air traffic flows into LEBL.

Barcelona–El Prat Airport has three physical runways, two of which are parallel, designated 02/20, 06L/24R, and 06R/24L. The preferred daytime configuration is known as “West Parallel Runways”, in which arrivals predominantly use runway 24R and departures use runway 24L. However, certain larger aircraft may use runway 24R for departure when a longer take-off distance is required. For noise abatement reasons, the preferred nighttime configuration is known as “North Crossed Runways”, with arrivals on runway 02 and departures from runway 06R.

From an ATC perspective, trombone procedures at LEBL are managed by the Approach ATCOs (APP, feeders and final ATCOs). Feeder ATCOs are responsible for managing and sequencing aircraft entering the TMA, establishing the traffic sequence on each side of the trombone procedure, adjusting speeds, issuing shortcuts, and transferring aircraft to the final ATCO with appropriate separation. They also validate or, in most cases, modify the initial sequence generated by the Arrival Manager system (AMAN) according to traffic and weather conditions.

The final ATCO merges arrival flows coming from the Initial Approach Fix (IAF), previously sequenced by the feeders and approaching from either side of the trombone inbound leg, into a single

sequence towards the Intermediate Fix (IF), maintaining the required separation (typically 2.5 NM). The ATCO issues clearances to turn from the outbound legs towards the intermediate approach. Furthermore, the final ATCO may adjust the rate at which the aircraft enters the trombone procedure (metering) in coordination with the feeders, depending on traffic and weather conditions.

Consequently, although trombone procedures follow predefined routes, the actual distance flown depends on traffic sequencing and tactical ATC decisions, particularly shortcut clearances and turn timing. This variability is analyzed using historical ADS-B trajectories.

2.2 ADS-B Dataset and Trajectory Processing

This section describes the data collection and processing procedures, as well as the variables used to characterize arrival operations and estimate the distance flown during trombone procedures.

2.2.1 Data Collection and Processing

The ADS-B trajectories recorded in the vicinity of Barcelona–El Prat Airport were analyzed over a period of approximately one year, from March 2025 to March 2026. Raw data (undecoded messages) were retrieved from the OpenSky Network database [15] using queries performed with the pyopensky library [16]. The data were then processed using the pyModeS [17] and traffic [18, 19] libraries, together with custom-developed algorithms, to rebuild and clean the trajectories and extract additional information from the available ADS-B messages. To ensure data consistency and reduce data volume, the trajectories were resampled at 7-second intervals.

Next, spatial and altitude filtering was applied, retaining records within a 50 NM (nautical miles) radius of Barcelona Airport Reference Point (ARP) and below 10,000 feet AMSL (Above Mean Sea Level). Records were grouped by 24-bit ICAO address and callsign, with trajectories segmented when temporal gaps greater than 10 minutes were detected. Subsequently, only trajectories whose last recorded point was located within 15 NM of the ARP were selected, resulting in a total of 176,731 trajectories for analysis.

Finally, the dataset was enriched with information from multiple sources, including supplementary ADS-B data, meteorological observations, and Air Traffic Services (ATS) flight plan (FPL) information. Historical METAR reports, retrieved from the Iowa State University database [20], were used to characterize weather conditions during arrival operations. Additionally, an indicator of traffic demand in the approach area was derived from ADS-B trajectories by estimating the number of aircraft observed in the area over a 10-minute time window [21], following a methodology similar to that used for EUROCONTROL’s sector occupancy metric. As a result, a dataset was obtained that integrates operational, meteorological, and aircraft trajectory information and was used to develop and evaluate predictive models.

2.2.2 Study Variables

Based on the data processing described above, a set of variables was obtained to characterize the operational conditions of each flight. These variables were grouped into two main categories: operational and flight-planning characteristics (FPL features) and meteorological conditions (weather features), as summarized in Table 1.

Since the objective of this study is to reduce pre-departure uncertainty regarding the distance flown during the approach phase in procedures involving trombone maneuvers, it is essential to distinguish between information available during flight planning and information requiring estimates or forecasts. Although variables such as aircraft type and estimated time of arrival are available before departure, actual traffic occupancy in the approach area and weather conditions at arrival remain uncertain.

Table 1. Input variables considered for trombone distance prediction.

Feature	Unit	Type	Example	State Features	Information Source
<i>Flight Plan Features</i>					
Aircraft Model	[-]	Cat.	A320	Static	Airline (OFF)
Callsign	[-]	Cat.	IBE1234	Static	Airline (OFF)
Estimated time of arrival (ETA)	UTC	Num.	21:20:07	Static	Airline (OFF)
Day of the week	[-]	Num.	2	Static	Airline (OFF)
Week of the year	[-]	Num.	30	Static	Airline (OFF)
Runway Configuration	[-]	Cat.	24L	Static	FPL
STAR Initial Point	[-]	Cat.	BISBA	Static	Airline (OFF)
IAF	[-]	Cat.	TEBLA	Static	Airline (OFF)
City Pair	[-]	Cat.	XXXX-XXXX	Static	Airline (OFF)
Wake Turbulence	[-]	Cat.	M	Static	Airline (OFF)
APP Occupancy	Aircraft	Num.	14	Static	Historic and Forecasted Data
<i>Weather Features</i>					
Surface Wind Direction	°	Num.	340	Static	Forecasted Data
Surface Wind Speed	kt	Num.	12	Static	Forecasted Data
Aircraft Wind Direction	°	Num.	270	Static	Forecasted Data
Aircraft Wind Speed	kt	Num.	30	Static	Forecasted Data
METAR Temperature	°C	Num.	32	Static	Forecasted Data
CAVOK	[-]	Cat. (bool)	True	Static	Forecasted Data

In this study, some input variables were obtained from historical observations and considered representative of the conditions expected during flight operations. Therefore, it was assumed that their values were known without error before departure, excluding the uncertainty associated with their forecasts. This assumption enables independent analysis of the uncertainty in the distance flown and the uncertainty in the input variables. Consequently, the results reflect the performance of the model under this assumption. The target variable is the distance flown by each aircraft from the IAF to the Touchdown Zone (TDZ). This measure is used to characterize the distance flown during approach procedures involving trombone maneuvers.

In addition, other variables derived from the ADS-B trajectories were used, including latitude, longitude, timestamp, barometric altitude, ground speed, track, turn rate, and vertical rate. These variables enable the characterization of aircraft trajectories and support the development of the models in this study.

3. Methodology

This section evaluates the ability of supervised ML models to predict the target variable using pre-departure information. The classification and regression approaches are first assessed. Their limited predictive performance motivates the development of a probabilistic framework based on an ML-guided stochastic hybrid automaton characterizing the variability in the distance flown and reducing uncertainty in pre-departure estimates.

3.1 Trombone Distance Prediction Based on Machine Learning Models

The first part of the study focuses on evaluating the predictive capability of ML models for the target variable. First, a classification problem is considered, in which the trajectories are grouped into three categories (short, medium, and long trombone procedures). Subsequently, a regression approach is considered, with the distance flown during the approach procedure as the target variable.

3.1.1 Classification Model

In the first approach, each flight is assigned to one of three categories based on the distance flown from the IAF: short approaches (≤ 30 NM), medium approaches (30–40 NM), and long approaches (> 40 NM). Figure 1 illustrates these categories in red, yellow, and green, respectively, identifying the areas of the trombone procedure where aircraft turn from the outbound leg toward the intermediate approach. The distance thresholds defining these categories were established based on operational criteria and calibrated to obtain a similar number of flights in each category. Short approaches correspond to direct routings or trombone maneuvers that are not fully executed due to low traffic demand. The medium category, although covering the smallest spatial area, accounts for the largest proportion of flights. Finally, long approaches correspond to the longest trombone maneuvers, generally observed during periods of high traffic density.

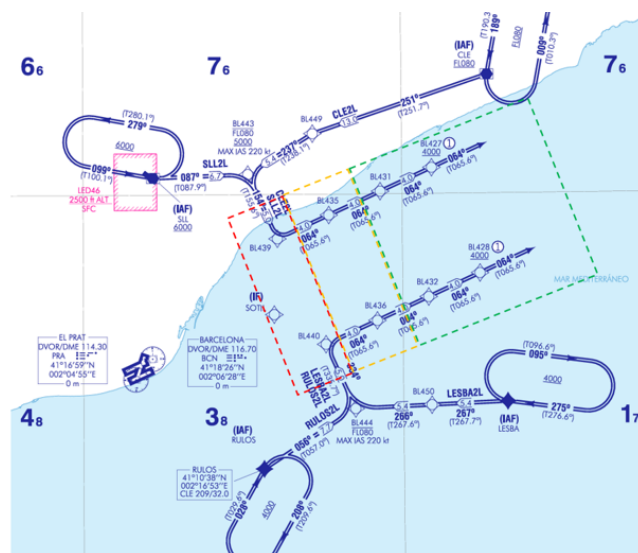


Figure 1. Classification of approaches according to trombone extension. LEBL RWY 24L Transition Chart. Source: AIP Spain.

To analyze the influence of pre-departure information on the predictive performance of ML models, two sets of input variables were considered. In the first experiment, only operational and flight planning variables (FPL Features) were used. Subsequently, meteorological variables were incorporated to determine whether this additional information could improve the prediction of the distance flown during the approach procedure. All variables are detailed in Table 1.

For each set of input variables, different classification algorithms were trained and evaluated using the PyCaret library [22]. Model performance was assessed using accuracy, precision, recall and F1-score. Table 2 presents the results obtained using FPL Features alone and the complete set of features (FPL + Weather).

The results indicate that the classification models have limited predictive capacity to estimate the distance flown using information available before flight dispatch. Although Gradient Boosting achieves

Table 2. ML Model Results for the Classification Problem.

Model	FPL Features				FPL+Weather			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
1 Gradient	0.593	0.590	0.588	0.591	0.604	0.603	0.589	0.604
2 Random Forest	0.584	0.580	0.579	0.581	0.596	0.593	0.590	0.594
3 AdaBoost	0.571	0.564	0.569	0.565	0.580	0.575	0.571	0.576
4 Extra Trees	0.550	0.547	0.554	0.548	0.575	0.574	0.564	0.574
5 LDA	0.521	0.522	0.540	0.521	0.546	0.557	0.540	0.549
6 Ridge Classifier	0.519	0.507	0.535	0.506	0.545	0.540	0.536	0.541
7 Decision Tree	0.504	0.504	0.512	0.504	0.526	0.526	0.523	0.526
8 Logistic Regression	0.528	0.518	0.533	0.518	0.498	0.483	0.491	0.480
9 Dummy (baseline)	0.336	0.113	0.333	0.113	0.336	0.113	0.333	0.169

the best performance (an accuracy of 59.3% using FPL features and 60.4% when weather variables are included), the inclusion of meteorological variables leads to only minor and inconsistent improvements across the other algorithms. This limited accuracy suggests that dividing the trajectories into three predefined categories, despite being supported by operational criteria, oversimplifies the actual operation of the procedure. Nevertheless, these findings motivate the exploration of regression approaches to assess whether predicting the distance flown as a continuous variable can better reflect the variability of the problem without the constraints imposed by fixed intervals.

3.1.2 Regression Model

To avoid the loss of information derived from the division of trajectories into three categories, the problem is reformulated using a regression approach. In this second approach, the target variable is directly defined as the distance flown by each aircraft during the approach procedure.

The same gradual methodology used in the classification problem is applied. First, the models are trained using only FPL features, and then meteorological variables are incorporated. The performance of the different models is evaluated using MAE (mean absolute error), RMSE (root mean squared error), and R². The results are presented in Table 3.

Table 3. ML Model Results for the Regression Problem.

Model	FPL Features			FPL+Weather		
	MAE(NM)	RMSE(NM)	R ²	MAE(NM)	RMSE(NM)	R ²
1 Gradient Boosting	7.09	11.35	0.240	7.09	11.28	0.249
2 Random Forest	7.30	11.60	0.207	7.17	11.40	0.234
3 Extra Trees	7.42	11.81	0.177	7.21	11.51	0.219
4 Ridge	7.84	11.85	0.172	7.82	11.83	0.174
5 Linear Regression	7.84	11.88	0.168	7.82	11.86	0.171
6 MLP	7.70	11.68	0.195	8.30	12.28	0.109
7 Dummy (baseline)	9.38	13.03	0.000	9.38	13.03	0.000

As in the previous classification stage, the regression models show limited predictive capacity, with Gradient Boosting achieving the best performance, MAE of 7.09 NM and an R² of 0.249. Although

this model reduces the error by 24.4% compared to the static baseline (9.38 NM, Dummy model), the RMSE is 11.28 NM which highlights the limitations of conventional supervised learning techniques in predicting the distance flown using only information available before flight dispatch. The inclusion of meteorological variables does not provide a substantial improvement (R^2 of 0.249 compared to 0.240). These findings suggest that the dominant source of uncertainty in trombone procedures is not primarily associated with the information contained in the FPL, but rather with tactical ATC decisions and real-time traffic sequencing. This motivates the transition from purely deterministic approaches towards a stochastic framework that accounts for dynamic decision-making.

3.2 Probabilistic Trajectory Modeling: Stochastic Hybrid Automaton

To represent the impact of tactical decisions made by APP ATCOs and the interactions between aircraft and ATC services, an alternative approach based on a hybrid automaton was adopted. Hybrid systems have been applied in air traffic management to combine continuous aircraft dynamics with discrete control decisions, particularly for conflict resolution and separation assurance [23]. In this framework, discrete modes represent different phases of operation, while continuous variables describe the physical state of the system. Within each mode, flow equations govern the evolution of continuous variables, whereas invariants define the conditions under which the system can remain in that mode. Transitions between modes are governed by guards, which are logical conditions based on continuous variables or external decision-making models. When a transition occurs, reset functions specify how the continuous variables are updated, if necessary, to reflect the new operating mode.

Given the complexity of the system, defining the automaton entirely through differential equations and discrete conditions would require extensive modeling to obtain an approximate representation of its behavior. Therefore, supervised Machine Learning (ML) models were employed to replace part of the explicit mathematical modeling, allowing the behavior of the system to be approximated more efficiently. However, this approach alone remains insufficient to adequately represent the variability and uncertainty of the operational environment. To overcome this limitation, a stochastic simulation framework based on the Monte Carlo method was incorporated, enabling repeated evaluations of the hybrid automaton under different input scenarios and perturbations. The resulting model is therefore formulated as a Machine Learning-Guided Stochastic Hybrid Automaton.

The uncertainty of the system is propagated through Monte Carlo simulations with 1,000 iterations in one-hour windows of real traffic data. In each simulation, the aircraft's initial conditions are modified, including the time and entry point into the TMA, speed, track, heading, fly-by turn anticipation, trajectory accuracy or navigation error, runway, IAF, initial point of the STAR, and wind speed and direction or meteorological conditions. Changes in STAR routing, runway configuration, weather conditions, or the presence of additional traffic are also introduced with lower probabilities.

To this end, probability density functions derived from historical data and the literature are used to model these variables and generate synthetic inputs for each Monte Carlo simulation. Figure 2 illustrates the overall data-processing workflow, with the hybrid automaton at its core, supported by several external models and modules that provide input and influence its behavior.

First, an initial classifier based on XGBoost (Trombone Probability Model, see Figure 2) estimates the probability that an aircraft will follow a trombone procedure or receive a direct routing (i.e., without flying any outbound leg), based on the prevailing operational conditions. The performance metrics of this classifier are presented in Table 4. The purpose of this model is to account for aircraft that have already received direct routing clearances before entering the modeled area. Let $P(T)$ denote the probability of following a trombone procedure. The model then performs $N \times P(T)$ simulations in which the aircraft is initially cleared to follow the trombone procedure, and $N \times (1 - P(T))$ simulations

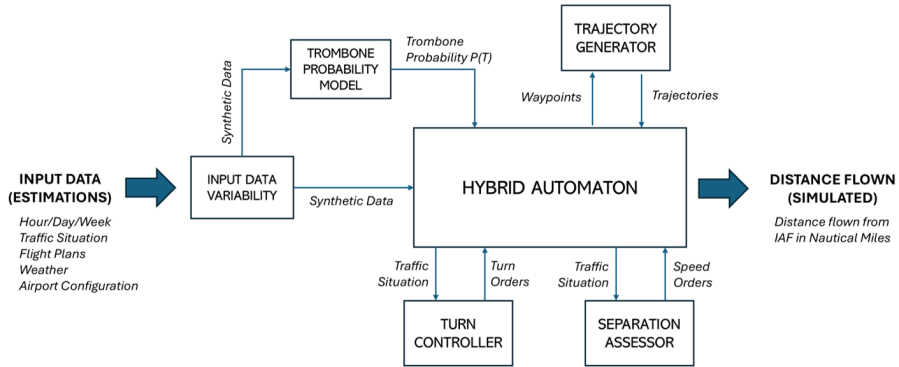


Figure 2. Data Processing Workflow in the Hybrid Automaton

in which it follows a direct routing, where N is the total number of Monte Carlo simulations.

The spatial and temporal interactions during the simulation are handled by three modules directly connected to the hybrid automaton. The trajectories to be followed are built by the Trajectory Generator (see Figure 2), which combines historical aircraft trajectories passing through predefined waypoints with trajectories generated using the BlueSky library [24, 25]. The performance metrics of the Trajectory Generator are presented in Table 4. A time-independent reference trajectory is generated by combining trajectories from both sources, providing an approximation of the flight path the aircraft is expected to follow. The actual simulated trajectory, however, depends on the aircraft dynamics, speed, and interactions with other aircraft. Furthermore, since aircraft frequently bypass the waypoints designated as "Initial Point" (shown in Figure 3), the model also reproduces this behavior at the same frequency observed in actual operations (48%).

Decision-making regarding aircraft maneuvers and separation is handled by two external ATCOs. First, turn clearances are determined by a pseudo-controller (Turn Controller, see Figure 2), formulated as a discrete-time predictive model (see performance metrics in Table 4). Every 5 seconds of simulation, an XGBoost model estimates the probability that the ATCO will issue a turn clearance based on the simulated traffic situation (see Figure 2, Traffic Situation), including arrival sequence, preceding aircraft, minimum separation requirements, and wind conditions. A turn order (Turn Orders in Figure 2) is issued when the cumulative probability exceeds a threshold of 0.50, calibrated using real operational data through a maximum likelihood estimation approach by fitting a Student's t -distribution. Second, aircraft separation is managed by a Separation Assessor (see performance metrics in Table 4), which predicts the future position of each aircraft along its planned trajectory and issues discrete speed adjustment commands (Speed Orders, see Figure 2) in increments of 5 kt (acceleration or deceleration) to maintain the minimum separation between aircraft, while complying with the indicated airspeed (IAS) restrictions specified in the approach chart. As a result of the simulation process, the system calculates the final distance flown (Distance Flown, see Figure 2).

To describe the geometry and operating states of the automaton, two discrete states are defined: trombone and direct. In the trombone state, the aircraft flies towards the IAF, then proceeds to the waypoints designated as "Initial Point" (see Figure 3) and subsequently to the four downwind waypoints (Downwind Points in Figure 3), after which it continues along the outbound leg until a turn clearance is issued (see Figure 3, trajectory C). In the direct state, the aircraft proceeds directly towards the IF and completes the approach (see Figure 3, trajectories B and D). Within each state, waypoints from the published approach chart are assigned discretely according to the runway

configuration and aircraft position. A single transition is defined between the two states, from trombone to direct, corresponding either to a direct routing clearance issued before reaching the IAF or to a turn clearance issued while flying along the outbound leg (see Figure 3, trajectories B and A, respectively).

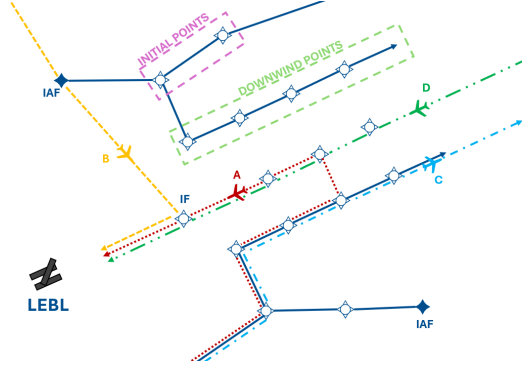


Figure 3. Trajectories Considered in the Hybrid Automaton.

Figure 4 provides a more detailed representation of the internal structure of the hybrid automaton and its interaction with external models. Its mathematical formulation is given by

$H = (Q, X, f, Init, Inv, E, Guard, Reset)$, where Q represents the discrete states, X is the state vector, f denotes the flow functions, $Init$ defines the initial conditions, Inv represents the invariants, E denotes the transitions, $Guard$ defines the guard conditions, and $Reset$ represents the reset functions applied when transitioning between states.

Figure 4 illustrates in detail how transitions between discrete states and their continuous evolution are analytically managed. The state vector $X = (Lat, Lon, h, v, \chi) \in \mathbb{R}^5$ consists of latitude (Lat), longitude (Lon), altitude (h), ground speed (v), and track (χ). Upon initialization of the state vector and the probability $P(T)$, the initial condition (q_0, x_0) is established, determining whether the system starts in Q_1 (Trombone State: Start Point $\rightarrow$ IAF $\rightarrow$ Initial Points $\rightarrow$ Downwind Points) or Q_2 (Direct State: Start Point $\rightarrow$ IF $\rightarrow$ End Approach). During the simulation, in either state, the turn rate is dynamically adjusted using information provided by the Trajectory Generator, while speed adjustments are determined by the Separation Assessor ($\Delta v = f(Traffic, t)$, as shown in Figure 4).

The transition from Q_1 to Q_2 occurs when an event is triggered by a turn clearance issued by the Turn Controller when the guard condition is met (Guard: Turn Order > 0.5 , in Figure 4). However, this transition depends on a constraint imposed by a separation invariant (Invariant: Minimum Separation, in Figure 4), which accounts for conflict situations (potential losses of the required separation minima) during the initial and intermediate approach phases. If the model predicts that executing the turn clearance would result in a future loss of separation, the transition is blocked, even if the decision threshold has been exceeded.

Table 4 presents the most relevant performance metrics for each of the intermediate models that support the hybrid automaton.

For model development, the entire dataset was divided into three subsets by randomly selecting days from the observed period. This approach ensured that flights operating simultaneously remain within the same subset. The three subsets, denoted as A, B, and C, accounted for 40%, 40%, and 20% of the data, respectively, with a total of 176,731 trajectories. Subset A was used to train and test the Turn Controller and Separation Assessor. Subset B was used to train the Initial State Predictor, estimate the probability density functions representing the uncertainty associated with the initial

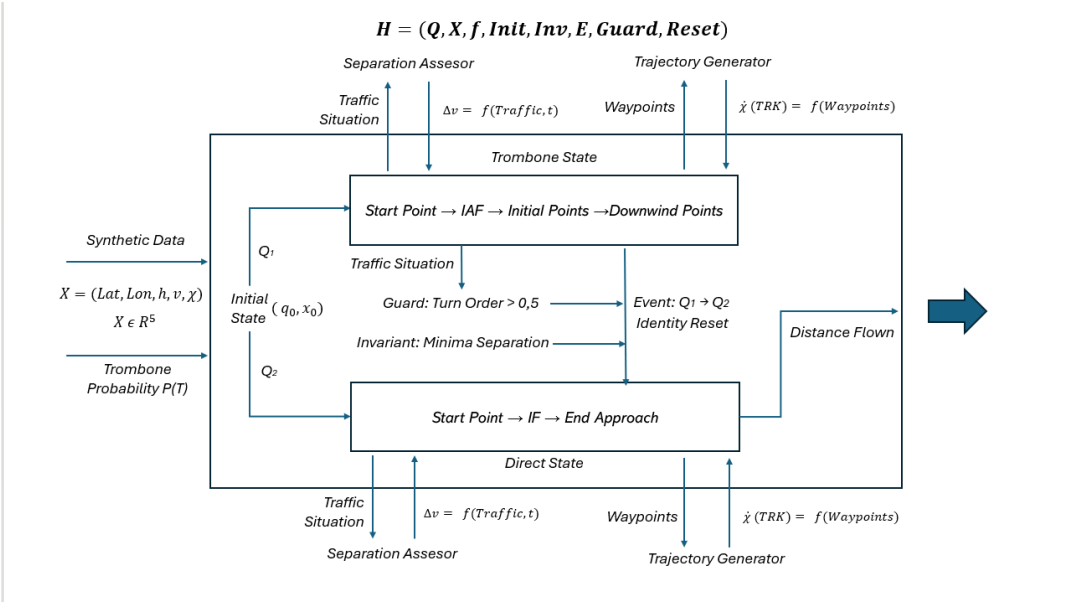


Figure 4. Internal Structure of the Hybrid Automaton.

Table 4. Performance Evaluation of the Intermediate Models Supporting the Hybrid Automaton.

Model	Evaluation Criterion	Metric	Value
Trajectory Generator	Comparison between actual flight trajectories and generated trajectories passing through the same waypoints	Mean Lateral Error	0.61 NM
Trombone Probability Model	Binary classification between trombone procedures and direct routings based on operational conditions	Accuracy	0.797
Turn Controller	Difference in the distance at which the real ATCO and the model issue the turn clearance	MAE	1.47 NM
Separation Assessor	Prediction of aircraft separation over a 60-second horizon	MAE	0.20 NM

variables, and calibrate different thresholds and discrete parameters of the automaton. Both subsets A and B were divided further into training and testing sets using a 70/30 split. Finally, subset C was reserved exclusively for testing and validating the overall model. This data splitting strategy aims to reduce the risk of overfitting when the different models are combined, although it reduces the number of flights available for each individual task.

The model generates a histogram representing the distribution of distances flown across the simulations. As shown in Figure 5, the simulated distances tend to concentrate around a particular value. However, under complex traffic conditions, when the model cannot reliably distinguish between direct routing and trombone procedures, the resulting distribution becomes bimodal, with two distinct peaks corresponding to flights following trombone procedures and those receiving direct routing. To evaluate prediction accuracy, both the mean and the median are considered. While the mean is sensitive to outliers, the median provides a more robust measure of the central tendency of the distribution.

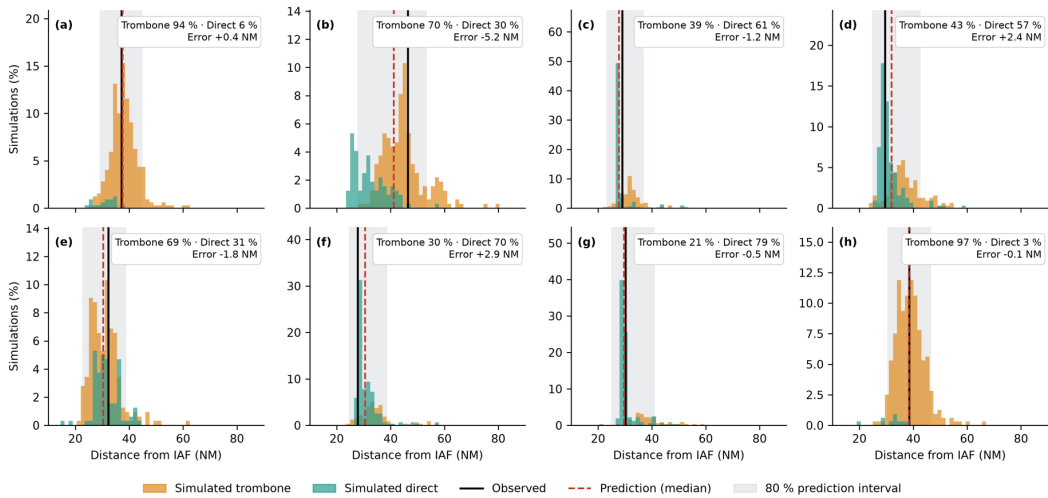


Figure 5. Randomly Selected Examples of Model-Generated Histograms Compared with the Actual Values.

In the validation performed in subset C, the hybrid automaton achieves a MAE of 5.38 NM in the distance flown from the IAF, or from the closest point when the IAF is not reached, compared to 7.09 NM obtained by a direct Gradient Boosting regression model (Section 3.1.2) using the same input variables (see Table 5). A baseline model that always predicts the median distance flown during the trombone procedure yields an MAE of 8.90 NM. The improvement is significant and particularly pronounced for direct approaches. The 80% prediction interval contains the actual distance flown for 80.4% of flights, while the 90% prediction interval achieves 90.0% coverage, indicating that the model is well calibrated, except for the longest trajectories. The residual distribution is better described by a Johnson SU distribution than by a normal distribution due to its heavier tails, with parameters $\gamma = 0.485$, $\delta = 1.348$, $\xi = 1.534$ NM, and $\lambda = 7.005$ NM.

Table 5. Comparison of Hybrid Automaton Results with Previous Models.

	MAE(NM)	RMSE(NM)	R ²
Hybrid Automaton (Median)	5.38	7.61	0.457
Hybrid Automaton (Mean)	5.40	7.41	0.485
Gradient Boosting (Regression Model)	7.09	11.28	0.249

These results demonstrate that stochastic hybrid architectures allow the uncertainty inherent in arrival procedures involving trombone maneuvers to be modeled and bounded with significantly greater accuracy. By integrating continuous flight dynamics with discrete state transitions governed by traffic interactions and ATC decisions, the hybrid automaton overcomes some of the limitations of static ML models.

However, some variability remains unmodeled due to the subjective nature of tactical ATC decision-making, as traffic management strategies may differ depending on the ATCO on duty. Furthermore, the modular architecture of the hybrid automaton allows for future improvements by independently enhancing its modules, such as the Turn Controller and the Separation Assessor.

4. Fuel Application

The probabilistic estimates of the distance flown are used to assess their potential implications for fuel planning. Rather than replacing the aircraft-specific performance models used by airlines to prepare OFPs, this study employs a theoretical model to quantify fuel consumption differences associated with the distance flown during trombone procedures. Operational implementation would require airline-specific performance models and flight-dependent data, particularly aircraft mass.

To perform these estimates, two types of aircraft, A and B, representative of the traffic in LEBL, are considered. Using EUROCONTROL’s BADA model [26], fuel consumption is estimated under three reference mass conditions (low, medium, and high, according to BADA criteria) to assess the sensitivity of the results to this variable. Furthermore, flights are classified into three categories according to their route distance (origin–destination): short-haul (less than 500 NM), medium-haul (between 500 and 1,000 NM), and long-haul (more than 1,000 NM), based on the distribution of distances observed in the analyzed dataset. The results of these estimates are presented in Table 6.

Table 6. Potential Fuel Savings (kg) by Aircraft Type, Mass, and Flight Distance Category.

Route Distance (haul) / Mass	Aircraft A			Aircraft B		
	Short	Medium	Long	Short	Medium	Long
Low	63.2	72.9	65.1	57.2	65.9	58.8
Medium	68.8	79.3	70.8	61.1	70.5	62.9
High	73.2	84.4	75.4	64.2	74.1	66.1

The results indicate potential fuel savings ranging from 57.2 to 84.4 kg per flight under the assumptions considered. For both aircraft types, the estimated savings increase with aircraft mass, with the highest values observed for medium-haul flights. An additional benefit associated with the cost-to-carry effect should also be considered, as carrying less fuel further reduces fuel consumption. Although the savings per flight are relatively modest, their cumulative effect across an airline’s annual operations could represent significant economic and environmental benefits in terms of CO₂ emissions.

5. Model Limitations

This study presents some limitations related to traffic selection, meteorological conditions, and modeling assumptions. General aviation, military operations, helicopters, and commercial flights performing holding patterns were excluded, representing 1.8% of the initial sample. The influence of departures from the non-preferred runway (RWY 24R) on arrival sequencing was also not explicitly considered.

Meteorological conditions were represented using METAR temperature, wind speed and direction, and CAVOK, without fully accounting for phenomena such as thunderstorms that may affect traffic management. Furthermore, some input variables were assumed to be known exactly before departure, excluding forecast uncertainty. Fuel estimates were based on theoretical aircraft performance models and reference masses, without access to actual aircraft mass or observed fuel consumption.

Finally, tactical ATC decisions remain a source of variability that is difficult to model. Discussions with LEBL APP ATCOs highlighted differences in speed control, radar vectoring, and shortcut clearances depending on individual ATCO judgment. Decision-support tools could contribute to more consistent and predictable traffic management in trombone procedures.

6. Conclusions

Fuel represents one of the major operating costs for airlines, making its efficient planning particularly relevant. During the preparation of OFPs, airlines must estimate fuel requirements without knowing the actual trajectories that aircraft will follow. In arrival procedures involving trombone maneuvers, tactical ATC decisions introduce uncertainty regarding the distance flown. This study investigates the potential to reduce this uncertainty using information available before flight dispatch, based on approximately one year of ADS-B data (from March 2025 to March 2026).

The results show that conventional supervised ML models, applied to both classification and regression problems, exhibit limited predictive capability, with a minimum MAE of 7.09 NM. In contrast, the proposed ML-guided stochastic hybrid automaton combines continuous flight dynamics with discrete ATC decisions and aircraft interactions, reducing the MAE to 5.38 NM. This represents an improvement of 24.1% compared to the best regression model and 39.5% relative to the baseline. Furthermore, the model provides well-calibrated probability distributions, achieving coverage rates of 80.4% and 90.0% for nominal 80% and 90% prediction intervals, respectively.

The application of these results to theoretical fuel consumption estimates using the BADA model suggests potential reductions in planned fuel ranging from 57.2 to 84.4 kg per flight for the two narrow-body aircraft types studied, under the assumptions considered. Additional benefits could arise from the cost-to-carry effect, as carrying less fuel further reduces fuel consumption. Although these savings are relatively modest on an individual flight basis, their systematic application across an airline's operations could provide relevant economic and environmental benefits. However, operational implementation would require the aircraft-specific performance and fuel consumption models used by airlines.

The main methodological contribution of this study lies in explicitly accounting for the inherently non-deterministic nature of the ATC environment, influenced by tactical ATCO decisions, traffic density, and other operational factors. Rather than attempting to predict a single trajectory for each aircraft, the proposed approach characterizes and bounds the uncertainty through probabilistic estimates that could support pre-departure fuel planning. This framework could also underpin decision-support tools aimed at harmonizing trombone management practices across ATCOS, thereby reducing trajectory uncertainty.

Nevertheless, the methodology presents some limitations, including the simplified representation of meteorological conditions, the lack of actual aircraft mass and fuel consumption data, and the exclusion of certain less frequent operations. These factors should be considered when evaluating the estimated fuel savings. Future research will focus on incorporating uncertainty in pre-departure input variables, improving the representation of operational conditions, and assessing the generalization of the model to other airport environments.

Author contributions

Conceptualization: L.S.-M., E.S.A., R.A.L.; Methodology and Validation: all authors; Software: R.A.L.; Formal analysis: L.S.-M., E.S.A., J.A.P.C., R.A.L.; Resources: J.A.P.C., L.P.S.; Data curation: L.S.-M., R.A.L.; Original draft: L.S.-M., R.A.L.; Review and editing: all authors; Funding acquisition and Supervision: L.P.S. All authors have read and approved the submitted version of the manuscript.

Open data statement

The processed dataset supporting this study is available in Mendeley Data at <https://doi.org/10.17632/n9btxs45b2.1>.

Reproducibility statement

The source code and algorithms used in this study are publicly available at <https://github.com/GIN-A-UPM/LEBL-Trombones>.

References

- [1] International Civil Aviation Organization. *Resolution A41-21: Consolidated Statement of Continuing ICAO Policies and Practices Related to Environmental Protection – Climate Change*. Assembly Resolution. International Civil Aviation Organization (ICAO), 2022. URL: https://www.icao.int/environmental-protection/Documents/Assembly/Resolution_A41-21_Climate_change.pdf.
- [2] Xavier Olive, Junzi Sun, Luis Basora, and Enrico Spinielli. “Environmental inefficiencies for arrival flights at European airports”. In: *PLOS ONE* 18.6 (June 2023), e0287612. DOI: 10.1371/journal.pone.0287612.
- [3] Bruno Favennec, Aymeric Trzmiel, and Karim Zeghal. *Point Merge Implementation: A Quick Guide*. Tech. rep. Edition 1.51. EUROCONTROL, Dec. 2024. URL: <https://www.eurocontrol.int/sites/default/files/2024-12/eurocontrol-point-merge-guide-v1-51.pdf>.
- [4] W. Liu, D. Delahaye, F. A. Cetek, Q. Zhao, and P. Notry. “Comparison of performance between PMS and trombone arrival route topologies in terminal maneuvering area”. In: *Journal of Air Transport Management* 115 (Mar. 2024), p. 102532. DOI: 10.1016/j.jairtraman.2023.102532.
- [5] L. Boursier, B. Favennec, E. Hoffman, A. Trzmiel, F. Vergne, and K. Zeghal. “Merging Arrival Flows Without Heading Instructions”. In: *Proceedings of the 7th USA/Europe Air Traffic Management Research and Development Seminar*. Barcelona, Spain, July 2007, pp. 414–421.
- [6] D. Ivanescu, C. Shaw, C. Tamvaclis, and T. Kettunen. “Models of Air Traffic Merging Techniques: Evaluating Performance of Point Merge”. In: *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*. Hilton Head, South Carolina, USA: AIAA, Sept. 2009. DOI: 10.2514/6.2009-7013.
- [7] A. Errico and V. Di Vito. “Study of Point Merge technique for efficient Continuous Descent Operations in TMA”. In: *IFAC-PapersOnLine* 51.9 (2018). 15th IFAC Symposium on Control in Transportation Systems CTS 2018, pp. 193–199. DOI: 10.1016/j.ifacol.2018.07.032.
- [8] H. Hardell, A. Lemetti, T. Polishchuk, and L. Smetanová. “Evaluation of the Sequencing and Merging Procedures at Three European Airports Using OpenSky Data”. In: *Engineering Proceedings* 13.1 (Jan. 2022), p. 13. DOI: 10.3390/engproc2021013013.
- [9] H. Hardell, E. Otero, T. Polishchuk, and L. Smetanová. “Optimizing air traffic management through point merge procedures: Minimizing delays and environmental impact in arrival operations”. In: *Journal of Air Transport Management* 123 (Mar. 2025), p. 102706. DOI: 10.1016/j.jairtraman.2024.102706.
- [10] R. Sáez, X. Prats, T. Polishchuk, and V. Polishchuk. “Traffic synchronization in terminal airspace to enable continuous descent operations in trombone sequencing and merging procedures: An implementation study for Frankfurt airport”. In: *Transportation Research Part C: Emerging Technologies* 121 (Dec. 2020), p. 102875. DOI: 10.1016/j.trc.2020.102875.
- [11] European Union Aviation Safety Agency (EASA). *Easy Access Rules for Air Operations*. Tech. rep. Revision 24. European Union Aviation Safety Agency, Mar. 2026. URL: <https://www.easa.europa.eu/en/document-library/easy-access-rules/easy-access-rules-air-operations>.
- [12] L. Hao, M. Hansen, and M. S. Ryerson. “Fueling for contingencies: The hidden cost of unpredictability in the air transportation system”. In: *Transportation Research Part D: Transport and Environment* 44 (May 2016), pp. 199–210. DOI: 10.1016/j.trd.2016.02.016.
- [13] AENA. *Annual Reports*. <https://www.aena.es/es/estadisticas/informes-anuales.html>. Accessed: Sep. 17, 2026.

- [14] ENAIRE. *Environmental Monitoring Report: BRAIN Project: Redesign of the Terminal Control Area of Barcelona Airport*. Tech. rep. ENAIRE, 2019. 488
- [15] Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, and Matthias Wilhelm. “Bringing Up OpenSky: A Large-scale ADS-B Sensor Network for Research”. In: *Proceedings of the 13th IEEE/ACM International Symposium on Information Processing in Sensor Networks (IPSN)*. 2014, pp. 83–94. doi: 10.1109/IPSN.2014.6846743. 489
- [16] open-aviation. *pyopensky: The Python interface for the OpenSky database*. <https://github.com/open-aviation/pyopensky>. Accessed: June 28, 2026. 490
- [17] J. Sun, H. Vu, J. Ellerbroek, and J. M. Hoekstra. “PyModeS: Decoding Mode-S Surveillance Data for Open Air Transportation Research”. In: *IEEE Transactions on Intelligent Transportation Systems* 21.7 (July 2020), pp. 2777–2786. doi: 10.1109/TITS.2019.2914770. 491
- [18] Xavier Olive. “traffic, a toolbox for processing and analysing air traffic data”. In: *Journal of Open Source Software* 4.39 (2019), p. 1518. doi: 10.21105/joss.01518. URL: <https://doi.org/10.21105/joss.01518>. 492
- [19] xoolive. *traffic: A toolbox for processing and analysing air traffic data*. <https://github.com/xoolive/traffic>. Accessed: June 28, 2026. 493
- [20] Iowa Environmental Mesonet. *IEM: Download ASOS/AWOS/METAR Data*. <https://mesonet.agron.iastate.edu/request/download.phtml>. Accessed: July 18, 2026. Iowa State University. 494
- [21] EUROCONTROL. *ATFCM Operations Manual*. Tech. rep. EUROCONTROL Network Manager, 2021. 495
- [22] PyCaret. *API Reference*. <https://pycaret.org/reference/>. Accessed: Sep. 29, 2026. 496
- [23] Cengiz Pasaoglu, Baris Baspinar, Nazim Kemal Ure, and Gokhan Inalhan. “Hybrid Systems Modeling and Automated Air Traffic Control for Three-Dimensional Separation Assurance”. In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 230.9 (2016), pp. 1788–1809. doi: 10.1177/0954410015619443. 497
- [24] Jacco M. Hoekstra and Joost Ellerbroek. “BlueSky ATC Simulator Project: An Open Data and Open Source Approach”. In: *Proceedings of the 7th International Conference on Research in Air Transportation (ICRAT)*. Philadelphia, USA, 2016. URL: <https://resolver.tudelft.nl/uuid:d1131a90-f0ea-4489-a217-ad29987689a1>. 498
- [25] TU Delft CNS/ATM. *BlueSky: The Open Source Air Traffic Simulator*. <https://github.com/TUDeft-CNS-ATM/bluesky>. Accessed: Oct. 18, 2026. 499
- [26] EUROCONTROL. *Base of Aircraft Data (BADA)*. <https://www.eurocontrol.int/model/bada>. Accessed: Aug. 27, 2023. 500