

Evaluating and Improving Flight Derivation from ADS-B Data

Jonah Goode*

PlaneQuery.com, USA

*Corresponding author: goodejonah@gmail.com

Abstract

Deriving flights from ADS-B data is challenging due to inaccurate messages and poor network coverage, particularly near airports. This paper presents and evaluates an ADS-B to flights algorithm that uses the readsb legs algorithm for flight segmentation and a neural network for takeoff and landing airport determination. The proposed algorithm is compared against existing algorithms and published flight datasets using ADS-B data from ADSB.lol, OpenSky, and ADS-B Exchange. Evaluation is performed against flight records from the Bureau of Transportation Statistics (BTS), FAA SWIM SFDPS, and EUROCONTROL. The proposed flight segmentation algorithm performs similarly to OpenSky's published flights. The neural network consistently outperforms OpenSky's airport determination method and a nearest-airport approach. Using OpenSky ADS-B data, the neural network improves landing airport accuracy from 90.89% to 95.51% on BTS, 89.00% to 94.62% on SFDPS, and 82.97% to 90.31% on EUROCONTROL.

Keywords: ADS-B; Flights; Neural Network

1. Introduction

Companies such as FlightAware and FlightRadar24 derive flights from multiple sources (Figure 1).

FlightAware's Data Sources

FlightAware began as a free interface into a single data feed from the FAA. Today, FlightAware provides the most accurate and comprehensive flight tracking by fusing data from thousands of sources around the world.

Data Sources

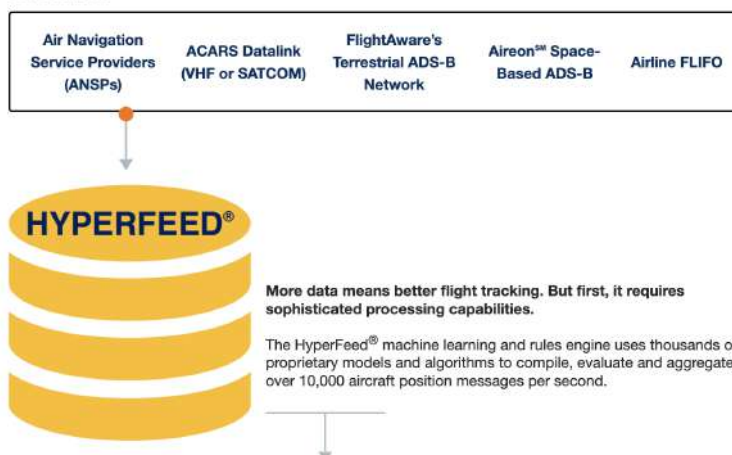


Figure 1. <https://www.flightaware.com/about/datasources/>

ADS-B networks like ADSB.lol are the most accessible. Government sources have imperfect flight data. They may censor or require the user to censor some flights. OpenSky and others derive flights using only ADS-B.

2. Background and Related Work

In this paper, a “flight” is when an aircraft makes a takeoff and landing. When an aircraft lands before taking off again and landing at the same airport, that is considered 2 separate flights.

2.1 Theory

ADS-B networks have incomplete coverage and the messages they receive are not always accurate. These issues are discussed in [1, 2, 3] and elsewhere. For ADS-B to flight derivation a challenge is that ADS-B networks have worse coverage near the ground. Some flights have 0 messages when the aircraft is on the ground. Therefore algorithms must guess which airport the aircraft took off or landed at. Figure 2 demonstrates this as well as inaccurate `on_ground`, `barometric altitude`, `baro_altitude_ft` and `ground speed`, `ground_speed_kt` message data.

2.2 Existing ADS-B to flights datasets

2.2.1 OpenSky

I used `flights_4` table data.

2.2.2 ADSB Exchange

Data is released for the first of every month, partitioned by `takeoff_time`.

2.3 Existing ADS-B to flights related published algorithms

2.3.1 OpenSky

The heuristic algorithm is described in Section 3.2 of Strohmeier et al. [4]. I found the algorithm described does not match the output of the `flights_4` table since at least 2024.

2.3.2 Sun et al.: Flight Extraction and Phase Identification

Sun, Ellerbroek, and Hoekstra [5] use a machine-learning algorithm to segment ADS-B message sequences into flights using the time gap between messages.

2.3.3 plane-notify

Sweeney [6] uses a state machine algorithm based on `on_ground`, `altitude` and `time gap` between messages. Airports are chosen by selecting the nearest airport to the identified takeoff/landing message using Ourairports.com data, 4.4.1.

2.3.4 OpenAP [7], Traffic [8]

OpenAP uses fuzzy logic to detect flight phases; however, it does not have functionality to segment a sequence of ADS-B messages into separate flights. Traffic uses a heuristic based on `callsign` and `time gap` between messages.

2.3.5 ReadsB Legs

`readsB` [9] uses a state machine incorporating hysteresis on `altitude`, `speed` and `time gap` between messages to segment messages into “legs”. Legs are not flights but are close. The legs feature

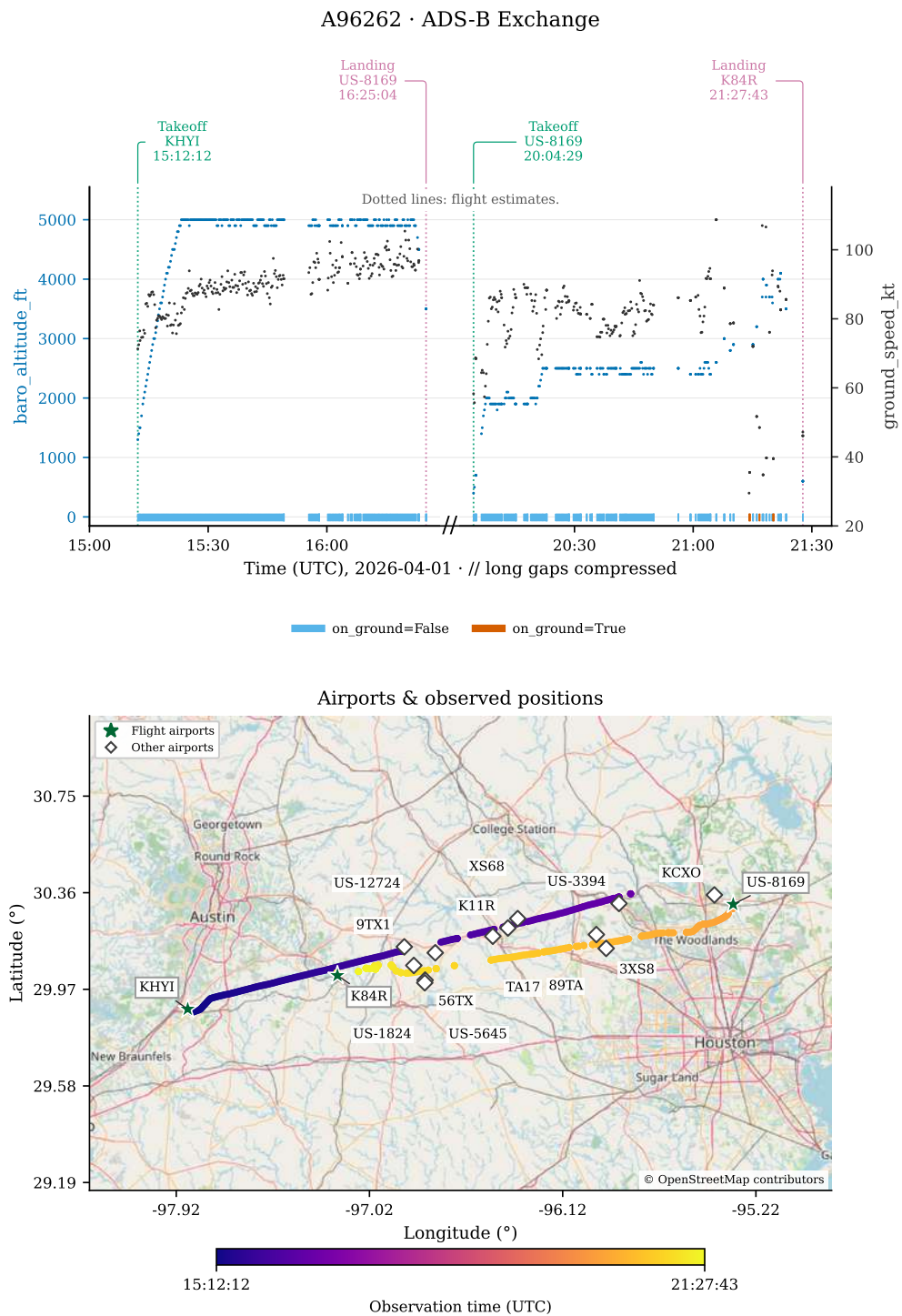


Figure 2. Hard ADS-B -> Flight Airport determination example. The nearest airport, US-3394 to the first flight's endpoint is not the correct landing airport nor is the more major airport, KCXO. <https://globe.adsbexchange.com/?icao=a96262&lat=30.170&lon=-97.400&zoom=7.2&showTrace=2026-04-01&leg=1>

is available on websites that use tar1090 [10] including globe.adsbexchange.com, map.opensky-network.org, adsb.lol, and others.

2.4 Previous ADS-B to Flights evaluation

Goens et al. [11] use EUROCONTROL Airport Operator Data Flow (APDF) aggregate flight counts at 10 major airports to evaluate an ADS-B to flights dataset derived as follows from OpenSky ADS-B data.

The assignment process starts by grouping the state vectors based on the aircraft’s Mode S transponder code (field icao24) and the flight’s ID (callsign). Each group is further segmented if there is a gap greater than 30 minutes between subsequent state vectors, with each segment assigned a splitnumber. [11, p. 4]

They found general agreement with some discrepancies e.g “The two major exceptions are Istanbul (LTFM) and Paris Charles de Gaulle (LFPG). Coverage at these airports is limited in 2022 and 2023”.

3. ADS-B Networks

3.1 ADSB.lol

This network releases publicly downloadable historical ADS-B data from 2023-02-16 to present on GitHub.

3.2 ADS-B Exchange

This network releases publicly downloadable full days of ADS-B data on the first of every month. ADS-B Exchange has better coverage than the other 2 networks. This can be seen in Table 4 in the better performance and lower percentages of gold ICAO addresses with zero detected flights.

3.3 OpenSky

I used ADS-B data from the `state_vectors_data4` table.

4. Reference Flight Datasets

A ground truth dataset for this problem would contain all flights for a given aircraft over a specified time period. Ideally, it would include flights that are challenging to derive from ADS-B data. I found the following publicly available datasets to be the most suitable, although all contain both expected and unexpected errors in flight segmentation and airport determination. A few examples of these errors are provided. Discrepancies between surveillance-derived and flight event datasets were also found by Hotle et al. [12].

4.1 Bureau of Transportation Statistics (BTS) On-Time Performance data

Historical Flight Data from 1987 to present day (with at least a 2 month lag) is publicly downloadable from https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGJ. It contains flight data provided by major US airlines in FAA airspace. Flight rows have `WHEELS_OFF` and `WHEELS_ON` columns, which I use for `takeoff_time` and `landing_time`, respectively.

On 2026-03-01, BTS records incorrectly assign flights AS960 (PHOG–KSMF) and AS959 (KSMF–PHOG) to registration N205AS (ICAO address a1a4a7). The correct aircraft is N205HA, a1a546.

4.2 FAA SWIM SFDPS

The Federal Aviation Administration (FAA) System-Wide Information Management (SWIM) Flight Data Publication Service (SFDPS) is downloadable in real time. I recorded this feed for 6 months. Types of flights that may not be in this feed but do broadcast on ADS-B include visual flight rules (VFR) flights and military flights. I derive flights from the raw messages.

For the aircraft in Figure 2, SFDPS reported these flights: KHYI → KCXO, 15:10–16:41; and KCXO → KHYI, departing at 20:13 with no recorded arrival. This may have contributed to FlightRadar24 similarly incorrectly reporting a landing at KCXO (Figure 8).

4.3 Eurocontrol Aviation Data Repository for Research (ADRR)

ADRR publishes historical flights data of commercial and business aviation flights with a 1 year delay in release [13]. On 2024-06-01, 3c6573 diverted from its EDDF → KDFW flight to KOKC, then flew KOKC → KDFW. ADRR records KDFW as the first flight’s destination and omits the subsequent non EU airspace flight, as expected (globe.adsbexchange.com/?icao=3c6573).

4.4 Supporting Data

4.4.1 Ourairports.com [14]

Airports dataset with airport_elevation_ft, lat, lon and other information including a qualitative type field that labels airports as small, medium or large.

4.4.2 Micronics Registration Data

BTS and Eurocontrol flights do not have ICAO addresses. The flights are identified by aircraft registration. Therefore, a registration -> ICAO mapping must be created. ADSB.lol ADS-B messages include registration data from Micronics [15], a database constructed from government sources and individual contributors.

Table 1. Aircraft categories of gold reference flights

Category / description	BTS 2026-03-01 <i>n</i> = 15, 951	SFDPS 2026-03-01 <i>n</i> = 23, 013	Eurocontrol 2024-06-01 <i>n</i> = 28, 050
	(%)	(%)	(%)
A0: No category information	0.24	0.03	1.66
A1: Light (e.g., Cessna 172)	0.09	5.73	1.34
A2: Small (e.g., CRJ-200)	5.50	20.54	7.99
A3: Large (e.g., Boeing 737 / Airbus A320)	90.82	67.21	75.66
A4: High-vortex large (Boeing 757)	1.66	1.18	0.43
A5: Heavy (e.g., Boeing 777)	1.10	2.46	12.14
A6: High-performance aircraft	0.00	0.05	0.00
A7: Rotorcraft	0.11	0.01	0.09
B0–D2: Other reported categories	0.03	0.01	0.07
Missing category (null / blank)	0.46	2.78	0.61

Table 2. Airport characteristics of gold reference flights

Airport characteristic	BTS 2026-03-01 <i>n</i> = 15, 951	SFDPS 2026-03-01 <i>n</i> = 23, 013	Eurocontrol 2024-06-01 <i>n</i> = 28, 050
	TO / LD (%)	TO / LD (%)	TO / LD (%)
Airport size / type			
Small airport	0.00 / 0.00	4.33 / 4.39	0.36 / 0.35

Continued on next page

Airport characteristic	BTS	SFDPS	Eurocontrol
Medium airport	8.92 / 9.26	20.90 / 18.92	20.93 / 21.10
Large airport	91.08 / 90.74	74.72 / 76.67	78.66 / 78.51
Unknown / unresolved airport	0.00 / 0.00	0.04 / 0.02	0.00 / 0.00
No IATA or ICAO	0.00 / 0.00	1.35 / 1.40	0.07 / 0.08

Only small regional and/or general aviation airports have no assigned IATA or ICAO codes.

5. Methods

5.1 Flight Segmentation

The basis for the proposed flight segmentation algorithm uses the readsb legs algorithm. This proposed algorithm outperforms the other existing algorithms in Section 2.3, Table 3.

Leg boundaries do not necessarily correspond to takeoff/landing time. Figure 3 illustrates this distinction. The proposed algorithm assigns `takeoff_time` to the first airborne message in the leg and `landing_time` to the first subsequent ground message in the leg. If no subsequent ground message is available, it uses the last message in the leg.

Another issue is that legs have no concept of the end of the available data. If it is given data starting at 00:00 while the aircraft is broadcasting, this is where the first leg starts and therefore where the first takeoff is identified. Similarly, if the aircraft is still flying at the end of the ADS-B message sequence, e.g. 23:59, a landing will be produced. In the flight segmentation algorithm, I calculate flights for a `target_day` but provide ADS-B messages from `target_day - 1` through `target_day + 1`. The flights produced are filtered to retain only those with a `takeoff_time` on `target_day`. I assume flights are less than 24 hours.

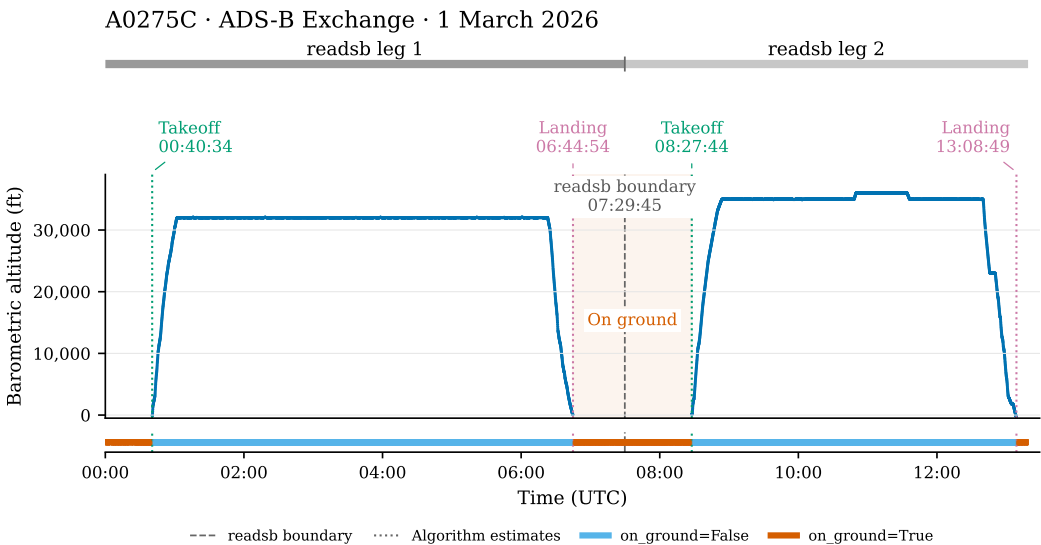


Figure 3. a0275c is on ground but still broadcasting messages. The leg boundary time is because readsb creates a new leg if the aircraft has spent more than 45 minutes on the ground. <https://globe.adsbexchange.com/?icao=a0275c&lat=35.923&lon=-94.062&zoom=4.3&showTrace=2026-03-01&leg=1×tamp=1772325571>

The readsb legs algorithm has errors which the flight segmentation algorithm will have too. For example, Figures 4 and 5 show that 2 flights are identified as being 1 leg due to incorrect on_ground values. Consequently, the segmentation algorithm produces only an incorrect flight, 18:47:13.875, KSCK -> 20:23:32.531, KSCK over this time period for a60529.

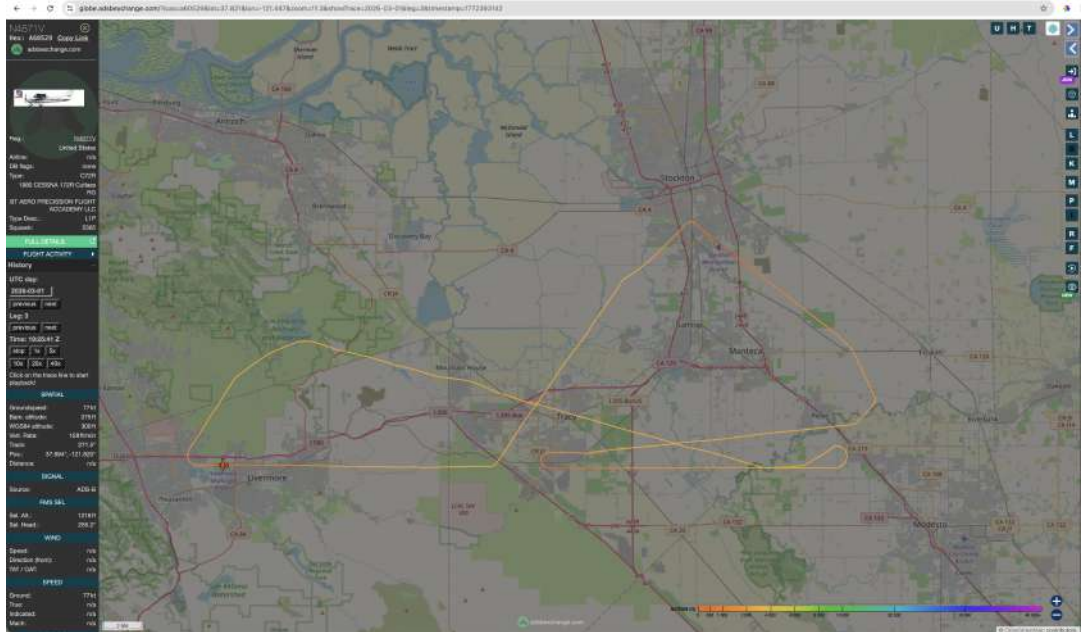


Figure 4. A60529 legs failure, <https://globe.adsbexchange.com/?icao=a60529&lat=37.893&lon=-121.255&zoom=14.7&showTrace=2026-03-01&leg=3×tamp=1772390834>

5.2 Airport Neural Network

Gated Recurrent Unit (GRU) Neural Networks have been used in similar problems [16, 17, 18, 19]. For determining the takeoff airport of a flight, the first 10 minutes of messages after the takeoff_time are used. For performance considerations, features are generated on 32 of the messages. If there are more than 32 messages in the first 10 minutes, then quadratic spacing is used to select more messages closer to the takeoff_time. Using the lat, lon of the takeoff message, all airports within 200 km are considered as candidates, limited to the 32 nearest if applicable due to performance considerations.

For each airport candidate, features for the selected 32 messages are derived from time, lat, lon, baro_altitude_ft, ground_speed_kt, track_deg, baro_rate_fpm, and on_ground. No absolute features of the aircraft or airport are used. lat, lon, and track_deg are made relative to each candidate airport, Figure 6. The altitude above the airport, baro_altitude_ft - airport_elevation_ft, is also a feature. Determining the landing airport is the same process but uses the 10 minutes of messages before the landing_time.

The model was trained using ADSB.lol ADS-B data and airport labels from BTS and SFDPS flight data for February 1–20, 2026, and validated on data from the same sources for February 21–27, 2026.

5.3 Evaluation

To compare a test flight dataset to a gold flight dataset, we consider segmentation and airport accuracy separately. A test flight matches a gold flight if its takeoff_time and landing_time are within 60 minutes of the gold flight's values. This value was chosen arbitrarily; tolerances greater than 30

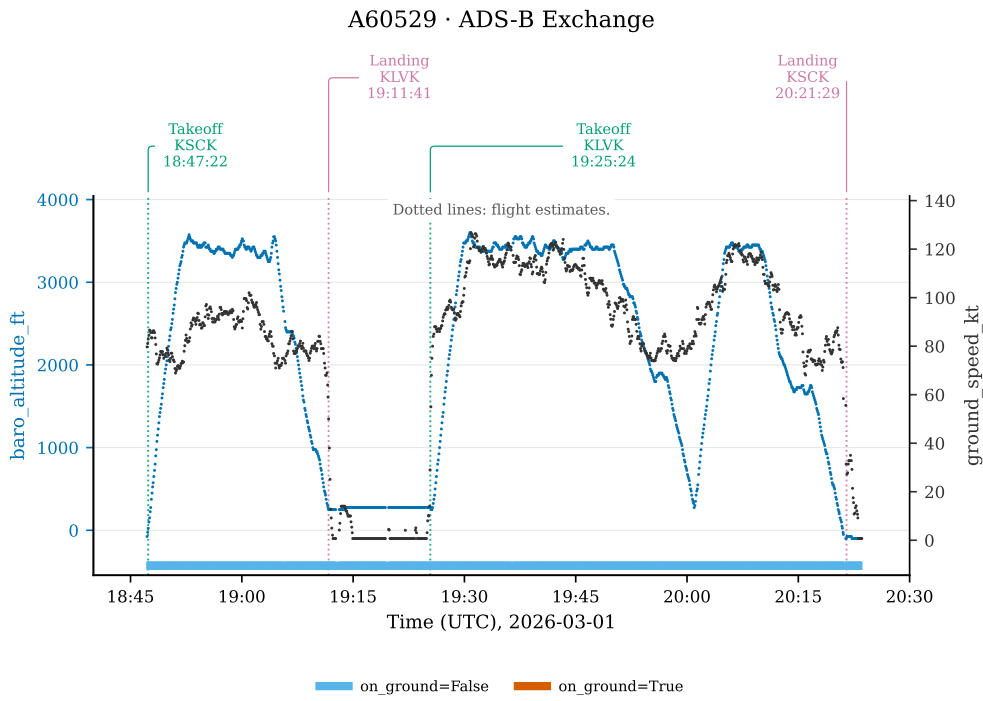


Figure 5. ADS-B messages of Figure 4. `on_ground` is always false, which is incorrect from 19:11:41 to 19:25:24.

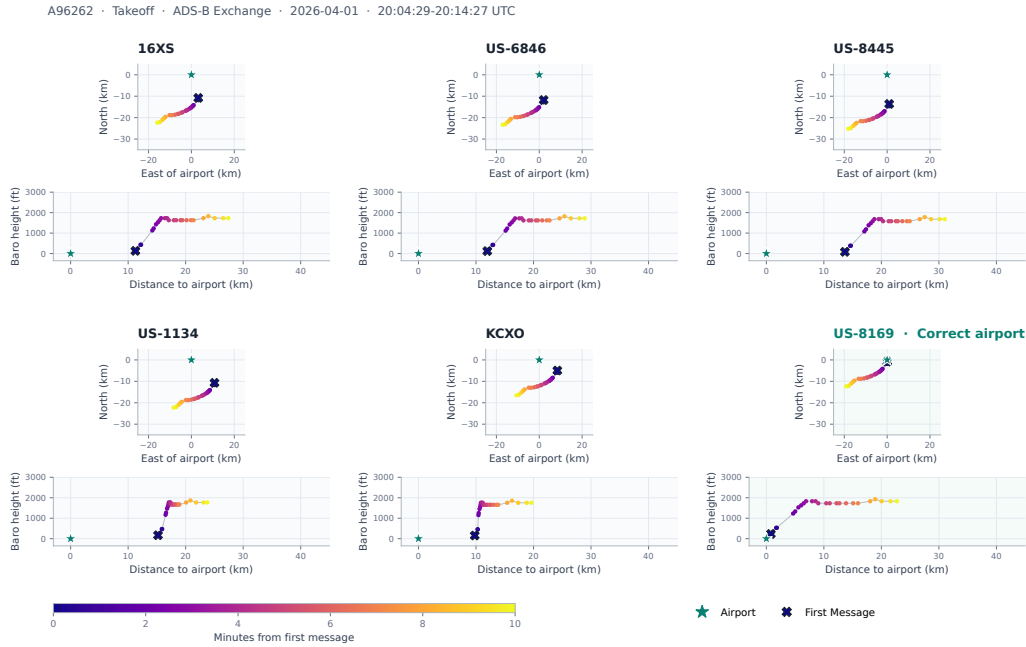


Figure 6

minutes minimally change metrics, as shown in Table 8. A gold flight can only be matched to 1 test flight.

6. Results

TO: takeoff; **LD:** landing. **Zero flights ICAOs (%):** The percentage of ICAOs in the gold set that have zero flights in the test set. This occurs when the ADS-B source contains no or insufficient messages for the test source to output a flight. **Airport acc.:** Calculated on matched flights; null airports are considered incorrect. **Airport null rate:** Calculated on matched flights. **Time bias:** The mean signed difference between the test and gold takeoff/landing times, calculated on matched flights.

Table 3. Algorithm comparison using OpenSky ADS-B

Method	Test flights T	Segmentation P / R / F1 (%)	Airport acc. TO / LD (%)
BTS 2026-03-01 Gold ICAOs: 4,719 Gold flights: 15,951			
Sun	17,821	80.31 / 89.72 / 84.76	93.80 / 89.53
plane-notify	16,804	87.79 / 92.48 / 90.07	93.99 / 91.95
Traffic	22,785	62.59 / 89.40 / 73.63	90.79 / 85.29
SFDPS 2026-03-01 Gold ICAOs: 9,117 Gold flights: 23,013			
Sun	26,530	74.43 / 85.81 / 79.72	90.30 / 87.74
plane-notify	25,513	79.54 / 88.18 / 83.64	90.48 / 89.50
Traffic	33,536	59.85 / 87.22 / 70.99	87.02 / 83.99
Eurocontrol 2024-06-01 Gold ICAOs: 7,875 Gold flights: 28,050			
Sun	33,757	62.69 / 75.44 / 68.48	82.39 / 77.78
plane-notify	23,656	81.07 / 68.37 / 74.18	88.63 / 85.48
Traffic	47,862	43.88 / 74.88 / 55.34	80.95 / 76.20

Table 4. Flight segmentation and airport accuracy by ADS-B source

Method / ADS-B	Test flights T	Segmentation P / R / F1 (%)	Zero flights ICAOs (%)	Airport acc. TO / LD (%)
BTS 2026-03-01 Gold ICAOs: 4,719 Gold flights: 15,951				
Nearest / ADSB.lol	16,628	89.02 / 92.80 / 90.87	1.04	98.12 / 92.35
NN / ADSB.lol				99.39 / 95.53
Nearest / OpenSky	16,878	88.83 / 93.99 / 91.33	0.74	92.72 / 86.83
NN / OpenSky				96.89 / 91.73
Nearest / ADSBX	17,276	88.67 / 96.03 / 92.20	0.68	99.90 / 95.82
NN / ADSBX				99.99 / 96.21
SFDPS 2026-03-01 Gold ICAOs: 9,117 Gold flights: 23,013				
Nearest / ADSB.lol	25,552	80.65 / 89.55 / 84.87	4.49	95.99 / 91.51
NN / ADSB.lol				98.17 / 94.90
Nearest / OpenSky	25,663	80.56 / 89.84 / 84.95	4.68	89.29 / 84.79
NN / OpenSky				94.90 / 90.73
Nearest / ADSBX	26,566	80.00 / 92.35 / 85.73	4.14	99.52 / 95.32
NN / ADSBX				99.71 / 95.83
Eurocontrol 2024-06-01 Gold ICAOs: 7,875 Gold flights: 28,050				
Nearest / ADSB.lol	23,165	70.18 / 57.96 / 63.49	2.79	80.01 / 71.14
NN / ADSB.lol				88.59 / 77.77
Nearest / OpenSky	27,025	83.03 / 80.00 / 81.49	1.82	86.04 / 80.46
NN / OpenSky				92.43 / 88.32

Continued on next page

Flight segmentation and airport accuracy by ADS-B source (continued)

	Test flights		Segmentation	Zero flights	Airport acc.
Method / ADS-B	T		P / R / F1 (%)	ICAOs (%)	TO / LD (%)
Nearest / ADSBX	28,564	85.43 / 86.99 / 86.20		0.18	96.80 / 93.26
NN / ADSBX					97.41 / 94.82

Since ADS-B Exchange only provides a single day of data, all test_srcs were run with only the target_day, 2026-03-01, thus there are lower segmentation scores than for Table 5.

152

Table 5. Flight evaluation by reference dataset

	Test flights	Segmentation	Zero flights	Airport acc.	Airport null rate	Time bias
Method	T	P / R / F1 (%)	ICAOs (%)	TO / LD (%)	TO / LD (%)	TO / LD (min)
BTS 2026-03-01 Gold ICAOs: 4,719 Gold flights: 15,951						
OpenSky				83.24 / 75.29	12.86 / 10.98	
OpenSky-Nearest	16,963	90.10 / 95.82 / 92.87	0.81	84.03 / 73.86	0.00 / 0.39	2.12 / -1.80
OpenSky-NN				90.10 / 87.27	0.00 / 0.00	
Nearest				92.95 / 90.89	0.00 / 0.00	
NN	16,748	92.26 / 96.87 / 94.50	0.81	97.00 / 95.51	0.00 / 0.00	1.01 / 0.64
SFDPS 2026-03-01 Gold ICAOs: 9,117 Gold flights: 23,013						
OpenSky				77.25 / 72.80	17.52 / 10.70	
OpenSky-Nearest	25,388	81.90 / 90.35 / 85.92	6.36	78.61 / 71.50	0.00 / 0.47	2.44 / -4.30
OpenSky-NN				86.53 / 85.28	0.04 / 0.00	
Nearest				89.71 / 89.00	0.00 / 0.00	
NN	25,378	83.54 / 92.12 / 87.62	5.34	95.19 / 94.62	0.01 / 0.00	1.05 / -1.86
Eurocontrol 2024-06-01 Gold ICAOs: 7,875 Gold flights: 28,050						
OpenSky				77.77 / 78.77	19.93 / 15.98	
OpenSky-Nearest	27,407	89.03 / 86.99 / 88.00	3.28	84.06 / 77.57	0.00 / 0.04	2.70 / -3.53
OpenSky-NN				89.16 / 86.59	0.37 / 0.38	
Nearest				87.24 / 82.97	0.00 / 0.00	
NN	25,722	87.71 / 80.43 / 83.92	3.49	93.32 / 90.31	0.12 / 0.04	1.88 / -1.67

Table 6. Eurocontrol algorithm airport accuracy on training and nontraining airports (2024-06-01)

Method	Training		Nontraining		Drop (pp)
	Training <i>n</i> TO / LD	accuracy (%) TO / LD	Nontraining <i>n</i> TO / LD	accuracy (%) TO / LD	
Nearest	1,108 / 1,212	97.65 / 97.11	21,454 / 21,350	86.70 / 82.17	10.95 / 14.94
NN		98.92 / 98.93		93.03 / 89.82	5.89 / 9.11

Some airports in Eurocontrol are also present in the training data such as London Heathrow (EGLL)

153

Table 7. Flight evaluation by reference dataset

Method	Test flights T	Segmentation P / R / F1 (%)	Zero flights ICAOs (%)	Airport acc. TO / LD (%)	Airport null rate TO / LD (%)	Time bias TO / LD (min)
BTS 2026-03-01 Gold ICAOs: 4,719 Gold flights: 15,951						
ADS-B Exchange	15,086	94.31 / 89.19 / 91.68	1.84	100.00 / 99.97	0.00 / 0.00	0.61 / 0.91
SFDPS 2026-03-01 Gold ICAOs: 9,117 Gold flights: 23,013						
ADS-B Exchange	22,929	84.48 / 84.17 / 84.33	9.28	99.75 / 99.75	0.00 / 0.00	0.44 / -1.40

The SWIM airport data for the example in Figure 2 is incorrect.

154

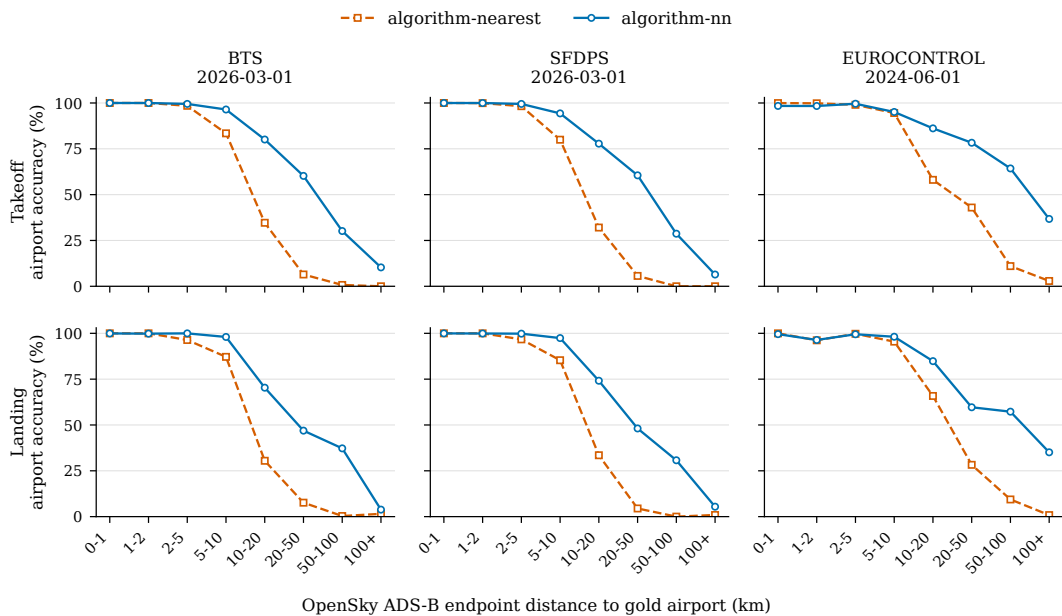


Figure 7

7. Discussion

7.1 Flight Segmentation Algorithms are too good for the BTS and SFDPS datasets

In Table 4, for BTS and SFDPS, recall goes up as the quality of the ADS-B network increases, e.g., recall in ADS-B Exchange > OpenSky > ADSB.lol. Precision decreases. While the flights in the BTS and SFDPS datasets are generally very accurate, they are missing many correct flights. Therefore, as more correct flights are detected by the better networks, precision decreases. Eurocontrol does not have this issue due to the dataset being in worse network coverage, resulting in less accurate ADS-B-derived flights.

While there are examples of the reference datasets producing correctly segmented flights with wrong airports (Figure 2) I don't believe it is a significant occurrence, that substantially affects evaluation conclusions.

7.2 Flight Segmentation

Table 5 shows that the flight segmentation of OpenSky and the proposed algorithm is similar. The proposed algorithm slightly outperforms on BTS and SFDPS, while OpenSky outperforms on Eurocontrol. There is no definitive evidence that the proposed flight segmentation algorithm is better or worse than OpenSky.

7.3 OpenSky's poor airport accuracy

OpenSky is using an OurAirports.com dataset from before 2022. I found the effect of this on accuracy to be <2%. As described in Strohmeier et al. [4], OpenSky outputs null for airports that are far in altitude or distance from the first/last message of a flight segment. Its non-null airport accuracy is 95.53/84.58, 93.66/81.52, 97.12/93.74 for takeoff/landing on BTS, SFDPS, and Eurocontrol, respectively. This remains lower than the neural network for BTS and SFDPS but is higher for Eurocontrol.

The neural network has worse performance on the matched flights of the OpenSky dataset than on the matched flights of the proposed algorithm. I found that this is due to OpenSky’s takeoff/landing times being further from the first/last message of the sequence; thus, the neural network had a less useful 10 minutes/32 points of information.

7.4 Features not used

7.4.1 Flight Segmentation

Changes in callsign are not used. Some aircraft do not change callsign between consecutive flights. Some aircraft such as General Aviation aircraft always use their registration as callsign. Like all ADS-B fields callsign may be inaccurate.

7.4.2 Airport Neural Network

Most flights for a given ICAO address are continuous, meaning the takeoff airport of a flight is the same as the landing airport of the previous flight. The takeoff or landing airport may be easier to determine as shown in Figure 2. There are cases where flights are not continuous, such as when an ICAO address is reassigned to another aircraft, when military aircraft reuse ICAO addresses, or when aircraft use Privacy ICAO Addresses (PIA).

7.5 ADS-B Exchange Results

Compared with OpenSky and the proposed algorithm, ADS-B Exchange has lower recall but higher precision: it misses more reference flights, but a greater proportion of the flights it reports are correct.

7.6 LLM algorithm improvement attempt

Asking ChatGPT-6 Astra XHigh Codex to improve the flight segmentation algorithm produced an algorithm that scored better on the evaluation metrics but would also produce spurious flights lasting 5 minutes. It found a way to overfit to the test data even though I explicitly told it not to.

8. Conclusion

The airport neural network outperformed a nearest airport approach on the proposed algorithm flight segments and OpenSky flight segments. It also outperformed the OpenSky airports.

It does not use features that could hurt generalization.

Appendix 1. Supplementary tables

Table 8. Matching-tolerance sensitivity against BTS on 2026-03-01, with a fixed scoring window of 01:00–23:00 UTC (inclusive). Both methods use OpenSky ADS-B; Algorithm denotes `algorithm-nn`. All metrics are percentages.

Tolerance (min)	OpenSky			Algorithm		
	Precision	Recall	F1	Precision	Recall	F1
60	90.10	95.82	92.87	92.26	96.87	94.50
45	89.62	95.31	92.38	92.23	96.84	94.48
30	88.75	94.38	91.48	92.12	96.72	94.37
10	80.39	85.49	82.86	88.00	92.39	90.14

Appendix 2. Supplementary figures

204



Figure 8



Figure 9. <https://www.adsbexchange.com/>

Acknowledgement

I thank Martin Strohmeier for encouraging me to submit an abstract to the symposium and for his valuable feedback.

205

206

207

Open data statement

All code and the model are in github.com/PlaneQuery/planequery-data. Flights derived from ADSB.lol ADS-B data using the proposed flight segmentation algorithm and neural network from 2024-01-01 to 2026-10-04 are available on huggingface.co/datasets/ggman12/planequery-data-flights.

Reproducibility statement

All raw data except for the historical SFDPS data is downloadable. ADRR and OpenSky require approval and are intended for academia. However, `flights_4` and `state_vectors_data4` are publicly downloadable for 2026-03-01.

References

- [1] Matthias Schäfer, Martin Strohmeier, Matthew Smith, Markus Fuchs, Vincent Lenders, and Ivan Martinovic. “OpenSky Report 2018: Assessing the Integrity of Crowdsourced Mode S and ADS-B Data”. In: *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. 2018, pp. 1–9. DOI: 10.1109/DASC.2018.8569833.
- [2] Xavier Olive, Jan Krummen, Benoit Figuet, and Richard Alligier. “Filtering Techniques for ADS-B Trajectory Preprocessing”. In: *Journal of Open Aviation Science* 2.2 (Mar. 2025). DOI: 10.59490/joas.2024.7882. URL: <https://journals.open.tudelft.nl/joas/article/view/7882>.
- [3] Ben Burwell. *Falsehoods Programmers Believe About Aviation*. FlightAware. June 2, 2025. URL: <https://flightaware.engineering/falsehoods-programmers-believe-about-aviation/> (visited on 10/26/2026).
- [4] M. Strohmeier, X. Olive, J. Lübke, M. Schäfer, and V. Lenders. “Crowdsourced air traffic data from the OpenSky Network 2019–2020”. In: *Earth System Science Data* 13.2 (2021), pp. 357–366. DOI: 10.5194/essd-13-357-2021. URL: <https://essd.copernicus.org/articles/13/357/2021/>.
- [5] Junzi Sun, Joost Ellerbroek, and Jacco Hoekstra. “Flight Extraction and Phase Identification for Large Automatic Dependent Surveillance–Broadcast Datasets”. In: *Journal of Aerospace Information Systems* 14 (Aug. 2017), pp. 566–572. DOI: 10.2514/1.I010520.
- [6] Jack Sweeney. *Plane Notify*. GitHub repository. URL: <https://github.com/jxck-s/plane-notify>.
- [7] Junzi Sun, Jacco M. Hoekstra, and Joost Ellerbroek. “OpenAP: An Open-Source Aircraft Performance Model for Air Transportation Studies and Simulations”. In: *Aerospace* 7.8 (2020). ISSN: 2226-4310. DOI: 10.3390/aerospace7080104. URL: <https://www.mdpi.com/2226-4310/7/8/104>.
- [8] 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>.
- [9] readsb. *readsb: ADS-B Decoder Swiss Knife*. URL: <https://github.com/wiedehopf/readsb> (visited on 10/26/2026).
- [10] tar1090. *tar1090: Web Interface for ADS-B Decoders*. URL: <https://github.com/wiedehopf/tar1090> (visited on 10/26/2026).
- [11] Quinten Goens, Rainer Koelle, Enrico Spinielli, Tatjana Bolić, Andrew Cook, and Martin Strohmeier. “The Open Performance Data Initiative: A Foundation Supporting the European Open Science Alliance for ATM Research”. In: *Proceedings of the 14th SESAR Innovation Days*. Rome, Italy, 2024. URL: https://www.sesarju.eu/sites/default/files/documents/sid/2024/papers/SIDs_2024_paper_100%20final.pdf.
- [12] Susan Hotle, Kyle Titlow, Mehdi Hashemipour, and Ed Strocko. “State-of-the-Art Analysis of U.S. Flight Event and Surveillance Data Coverage and Future Research Directions”. In: *Transportation Research Record* 2680.5 (2026), pp. 519–532. DOI: 10.1177/03611981251378487. eprint: <https://doi.org/10.1177/03611981251378487>. URL: <https://doi.org/10.1177/03611981251378487>.

- [13] EUROCONTROL. *EUROCONTROL Aviation Data Repository for Research – Metadata*. Data Description, Version 8 FNI Doc25. EUROCONTROL, Sept. 2026. URL: <https://www.eurocontrol.int/sites/default/files/2026-10/eurocontrol-aviation-data-repository-research-metadata.pdf>. 253
254
255
- [14] OurAirports. *OurAirports Open Data*. Public domain airport dataset, updated daily. 2026. URL: <https://ourairports.com/data/> (visited on 10/25/2026). 256
257
- [15] Micronics. *Micronics Aircraft Database*. <https://www.micronics.de/aircraft-database/>. Accessed: 2026-10-26. 2026. 258
259
- [16] Lan Ma, Xianran Meng, and Zhijun Wu. “Data-Driven 4D Trajectory Prediction Model Using Attention-TCN-GRU”. In: *Aerospace* 11.4 (2024). ISSN: 2226-4310. DOI: 10.3390/aerospace11040313. URL: <https://www.mdpi.com/2226-4310/11/4/313>. 260
261
262
- [17] Hesam Shafienya and Amelia C. Regan. “4D flight trajectory prediction using a hybrid Deep Learning prediction method based on ADS-B technology: A case study of Hartsfield–Jackson Atlanta International Airport (ATL)”. In: *Transportation Research Part C: Emerging Technologies* 144 (2022), p. 103878. ISSN: 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2022.103878>. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X22002911>. 263
264
265
266
267
- [18] Hesam Shafienya and Amelia Regan. “4D Flight Trajectory Prediction based on ADS-B data: A comparison of CNN-GRU models”. In: *2022 IEEE Aerospace Conference (AERO)*. 2022, pp. 01–12. DOI: 10.1109/AERO53065.2022.9843822. 268
269
270
- [19] Xin Ma, Linxin Zheng, and Xikang Lu. “4D trajectory prediction and conflict detection in terminal areas based on an improved convolutional network”. In: *PLOS ONE* 20.2 (Feb. 2025), pp. 1–16. DOI: 10.1371/journal.pone.0317549. URL: <https://doi.org/10.1371/journal.pone.0317549>. 271
272
273
274