Airline Route Similarity Via Trajectory-Based Dynamic Time Warping

Steven Lu Chen

April 22, 2025

Contents

1 Introduction

Air traffic networks are typically studied through the lens of geography, airline schedules, and aggregate passenger demand. Although these factors explain why certain routes exist, they do not fully capture how aircraft behave on those routes - their ascent profiles, cruising patterns, descent structures, and overall spatial dynamics. In this project, we take a different approach.

We ask the question, 'Which international routes are most behaviorally similar to the busiest U.S. trunk routes?' Instead of comparing routes by origin-destination labels or great-circle distance, we use multivariate telemetry data to represent each flight as a trajectory through space, time, and motion. Each flight is encoded as a fixed-length time series capturing standardized position, heading, velocity, and altitude - and similarity is computed via Dynamic Time Warping (DTW), a method designed to compare temporal sequences that may be misaligned in time.

By applying this approach to a filtered slice of historical ADS-B data from OpenSky, we discover a series of international analogs to iconic U.S. domestic routes like SFO–LAX, LAX–JFK, and ORD–LGA. Some analogs are intuitive — e.g., San Francisco-Toronto as a behavioral twin to SFO–LAX. Others are more surprising, such as the domestic Turkish route between Izmir and Ankara emerging as a telemetry analog to DCA-BOS, or Kathmandu-Kolkata closely matching ATL-MCO.

The results show that route similarity is not always geographical. By comparing how flights move, not just where they go, we surface latent structural patterns in global airspace - driven by traffic density, topography, regulatory constraints, and air traffic control behavior.

2 Trunk Route Selection and Dataset Constraints

We define trunk routes as high-frequency, strategically important corridors that form the backbone of the U.S. domestic airspace. These routes typically connect major airline hubs, serve dense population centers, and exhibit consistent operational patterns shaped by air traffic control, terrain, and network scheduling.

To construct a set of reference trajectories, we selected ten U.S. trunk routes spanning a range of distances, regions, and airline hubs. Each route was represented using a curated set of 10 commercial flights from **July 15, 2019**, drawn from the OpenSky Network's archived ADS-B telemetry. These flights were manually identified using airline callsigns (e.g., UAL, DAL, AAL) and flight numbers known to operate the specified trunk on that day. The resulting vectors were aggregated to form a representative behavioral profile for each route.

Our selection was guided by a combination of:

- Historical flight frequency
- Hub connectivity for the three major U.S. carriers (United, Delta, American)
- Geographic and operational diversity (costal, transcontinental, mid-continent, and short-haul)

Trunk Route	**Description**
San Francisco–Los Angeles	Ultra-short-haul shuttle between major West Coast cities
Denver–Chicago	Central inter-hub connector for United Airlines
Dallas–Atlanta	Dense Southern corridor connecting American and Delta hubs
Washington/National–Boston	Northeast business corridor with congested airspace
Chicago–New York/LaGuardia	Historic route connecting major financial centers
Los Angeles–New York/JFK	Coast-to-coast American flagship route with long cruise segment
Houston–Chicago	Mid-continent route connecting major United hubs
San Francisco–Seattle	West Coast route with terrain and descent variability
Atlanta–Orlando	Leisure-heavy Southeast pairing with aggressive climb/descent
Los Angeles–Atlanta	Long-haul domestic route linking Delta's two major hubs

Table 1: Selected U.S. trunk routes used as reference vectors

In parallel, we extracted a large comparison set of international flights from the same date to serve as candidate analogs. Each flight was represented as a multivariate trajectory and compared against the ten trunk aggregates using Dynamic Time Warping (DTW). To ensure quality and comparability, we imposed the following constraints on both trunk and candidate flights:

- Valid callsigns only: All flights had to use a standard three-letter ICAO prefix (e.g., AFR, ACA, ETD) corresponding to a major international airline. Private jets, regional carriers, and malformed callsigns were excluded.
- Complete trajectory coverage: Flights were only included if both the initial and final altitudes (standardized) were below the dataset-wide mean, ensuring that the full takeoff and landing arcs were captured.
- Minimum displacement: Flights had to show at least 100 km of geodesic movement between start and end points, removing local loops, positioning legs, or clipped telemetry.
- Flight duration cap: Any flight exceeding 8 hours in duration was excluded to avoid long-haul outliers. The longest flight of the trunk routes is LAX-JFK, which takes approximately 6 hours.
- Trajectory length standardization: All vectors were resampled to a fixed length of 200 time steps using interpolation over elapsed time, enabling direct DTW comparison.

These filters reduced the dataset from an initial 50 million telemetry rows to 22,000 high-quality flight vectors suitable for comparison. Chinese flights were explicitly excluded from results due to known telemetry corruption, sparse low-altitude coverage, and evidence of systemic spatial obfuscation in ADS-B broadcasts.

3 Trajectory Representation

Each flight in the dataset is represented as a multivariate time series encoding its position and motion through space and time. To enable meaningful similarity comparisons between flights of varying lengths, we resample each trajectory to a fixed length of 200 time steps, producing a uniform 200×6 matrix per flight. Each row in this matrix corresponds to a single timestamp, and each column encodes one of the following six features:

Feature	**Description**
Δ Latitude	Offset from origin latitude, in degrees
Δ Longitude	Offset from origin longitude, in degrees
Altitude	Standardized altitude (in feet)
Velocity	Standardized ground speed (in knots)
$\sin(\theta)$	Sine of true heading (in radians)
$\cos(\theta)$	Cosine of true heading (in radians)

Table 2: Flight trajectory representation: each flight is encoded as a 200×6 time series.

To compute positional offsets, we subtract the initial latitude and longitude from each point in the trajectory. This transformation centers each flight at the origin $(0, 0)$, allowing for shape comparison that is invariant to absolute location. True heading is transformed into its sine and cosine components to preserve directionality while avoiding angular discontinuities near $0/360°$.

Velocity and altitude are standardized across the entire dataset to ensure equal weighting during similarity computation. This prevents higher-magnitude features from dominating the DTW distance metric. Each final flight vector is therefore a shape- and behavior-preserving abstraction of its telemetry, suitable for direct pairwise comparison via DTW.

4 Measuring Trajectory Similarity

To quantify the behavioral similarity between flights, we use Dynamic Time Warping (DTW) - a classic algorithm for measuring distance between temporal sequences that may vary in speed or alignment. Unlike pointwise metrics such as Euclidean distance, DTW is robust to phase shifts, local deformations, and timing inconsistencies. This makes it particularly well-suited to comparing aircraft trajectories, where flights of similar structure may differ slightly in timing due to weather, routing, or air traffic control.

Each flight is encoded as a multivariate time series of shape 200×6, with dimensions corresponding to position (displacement in latitude/longitude), velocity, altitude, and heading. To compute similarity between two flights A and B, we apply DTW over their respective time series, treating each time step as a 6-dimensional vector.

Formally, given two trajectories:

$$A = (a_1, a_2, \dots, a_n),\ B = (b_1, b_2, \dots, b_n),\ a_i, b_j \in \mathbb{R}^6$$

DTW finds a warping path $P = (p_1, p_2, \dots, p_L)$ with $p_k = (i_k, j_k)$ such that:

- The path starts at $(1, 1)$ and ends at (n, m),
- The indices increase monotonically
- The total cost is minimized:

$$\mathrm{DTW}(A, B) = \min_P \sum_{k=1}^{L} ||a_{i_k} - b_{j_k}||_2$$

In practice, we use a fast approximation of DTW optimized for multidimensional time series. The output is a scalar distance value, with lower scores indicating higher trajectory similarity. For each of the ten U.S. trunk routes, we compute the DTW distance between its aggregated vector and each international candidate flight. We then retain the top-k most similar callsigns as potential analogs. To avoid bias from over-alignment of incomplete trajectories, we apply DTW only to flights that include the full takeoff-to-landing arc, as verified by altitude constraints at the endpoints.

5 Results

At the core of this analysis is the identification of international routes whose flight trajectories most closely resemble those of major U.S. trunk routes. The five most similar flights for each trunk route are shown in the tables below, ranked by increasing DTW score (i.e., decreasing similarity distance). Each result includes the matched flight's callsign, inferred origin-destination pair, and the DTW distance to the trunk route's aggregate vector.

Before presenting the top-k analogs for each trunk route, we note two exclusions from the final analysis due to data quality concerns.

First, the SFO–SEA trunk route was excluded entirely. While initially included, its resulting DTW scores were significantly higher than those of all other trunk routes, with no analogs scoring below 550. Visual inspection of these trajectories revealed highly inconsistent shapes and disproportionately small values for Δ longitude. We hypothesize that this led to poor alignment in DTW and reduced the reliability of any matches derived from this trunk.

Second, one match under the LAX–ATL trunk - a LOT Polish Airlines flight from Warsaw to Seoul - was excluded manually. Although it passed altitude filters, the flight was clearly truncated in the dataset, resulting in a partial trajectory that did not match the expected duration or displacement of a valid long-haul flight. Its DTW score was anomalously low relative to other, more plausible analogs.

These removals were made to preserve the interpretability and integrity of the remaining results.

Flight Number	**Route**	**DTW Score**
Gulf Air 161	Bahrain–Riyadh	345.4
Vueling 1302	Barcelona-Alicante	343.2
Gulf Air 163	Bahrain–Riyadh	341.6
Saudia 1131	Dammam-Riyadh	340.4
Air Canada 469	Ottawa–Toronto	325.6

Table 3: Top-5 DTW analogs for SFO–LAX

Flight Number	**Route**	**DTW Score**
Aeroflot 1009	Kaliningrad-Moscow/Sheremetyevo	244.8
British Airways 984G	London/Heathrow-Berlin/Tegel	289.4
Wizz Air 451	Eindhoven-Warsaw	290.4
British Airways 982G	London/Heathrow-Berlin/Tegel	290.4
Aeroflot 1294	Kaliningrad-Moscow/Sheremetyevo	291.0

Table 4: Top-5 DTW analogs for DEN–ORD

Flight Number	**Route**	**DTW Score**
Vueling 6598	Madrid-Barcelona	283.8
Air Europa 7707	Madrid-Barcelona	286.5
Iberia 2100	Madrid-Barcelona	291.3
Iberia 1730	Madrid-Barcelona	293.6
Iberia 1546	Madrid-Barcelona	294.5

Table 5: Top-5 DTW analogs for DFW–ATL

Flight Number	Route	DTW Score
Turkish Airlines 86E	Izmir-Ankara	386.3
KLM Royal Dutch 33N	Amsterdam-Billund	405.5
Qantas 816	Melbourne-Canberra	408.3
Air New Zealand 602	Queenstown-Wellington	416.3
Lufthansa 178	Frankfurt-Berlin/Tegel	416.3

Table 6: Top-5 DTW analogs for DCA–BOS

Flight Number	Route	DTW Score
LOT Polish Airlines 268	Amsterdam-Warsaw	161.8
KLM Royal Dutch 85R	Amsterdam-Warsaw	172.2
Ukraine International 808	Prague-Kyiv/Boryspil	174.0
KLM Royal Dutch 53W	Amsterdam-Warsaw	191.9
easyJet 6983	Edinburgh-Copenhagen	195.1

Table 7: Top-5 DTW analogs for ORD–LGA

Flight Number	Route	DTW Score
Air Canada 758	San Francisco-Toronto	351.3
Air Canada 786	San Francisco-Toronto	430.6
Air Canada 788	San Francisco-Toronto	475.5
Air Canada 738	San Francisco-Toronto	489.1
Air Canada 792	San Francisco-Toronto	497.0

Table 8: Top-5 DTW analogs for LAX–JFK

Flight Number	Route	DTW Score
KLM Royal Dutch 1308	Toulouse-Amsterdam	224.8
Brussels Airlines 42B	Toulouse-Brussels	258.2
Jetstar 724	Hobart-Sydney	263.4
KLM Royal Dutch 41H	Amsterdam-Ålesund	271.7
Brussels Airlines 87E	Toulouse-Brussels	274.6

Table 9: Top-5 DTW analogs for IAH–ORD

Flight Number	Route	DTW Score
Air India 248	Kathmandu-Kolkata	170.3
Lufthansa 2M	Hamburg-Munich	173.8
Swiss 119N	Zurich-Florence	181.5
Lufthansa 4M	Hamburg-Munich	183.8
AirAsia 1727	Penang-Singapore	187.4

Table 10: Top-5 DTW analogs for ATL–MCO

Flight Number	Route	DTW Score
KLM Royal Dutch 907	Amsterdam-Moscow/Sheremetyevo	306.7
Aeroflot 2695	Amsterdam-Moscow/Sheremetyevo	309.7
Aeroflot 2551	Amsterdam-Moscow/Sheremetyevo	334.1
Air France 1752	Paris/CDG-Kyiv/Boryspil	336.2

Table 11: Top-4 DTW analogs for LAX–ATL

The resulting analogs reveal several intuitive and insightful patterns. Short-haul trunk routes, such as SFO–LAX and DCA–BOS, consistently matched to similarly short regional routes abroad - including Bahrain–Riyadh and Melbourne–Canberra. These pairings reflect not only similar durations, but comparable climb-cruise-descent profiles and ATC behaviors.

Longer trunk routes showed more structural than geographic alignment. LAX–JFK, for instance, matched repeatedly to San Francisco–Toronto flights, suggesting that DTW prioritizes flight shape and temporal structure over geographic direction. Medium-haul routes such as DEN–ORD and ATL–MCO found analogs with surprisingly high fidelity, including Kaliningrad–Moscow and Hamburg–Munich, respectively.

While interpreting the DTW scores, we observed a consistent trend: shorter flights tended to receive lower DTW scores than longer ones. This is expected given the mechanics of DTW. For fixed-length trajectory vectors (in our case, 200×6), shorter flights exhibit smaller absolute changes per timestep, leading to lower per-step alignment errors during DTW. Additionally, short-haul routes often have simpler and more consistent phase structures (climb–cruise–descend) with less ATC-induced variation, making them easier to align.

Importantly, this effect does not compromise the integrity of our results, since we are only identifying the top-k most similar international analogs within each trunk route, not comparing DTW scores across trunks. However, it does suggest caution in interpreting raw DTW values as absolute indicators of similarity when route lengths differ substantially.

6 Conclusion

This project explored a novel approach to comparing airline routes: not by geography, scheduling, or airline identity, but by the shape and behavior of actual flights in the sky. Using Dynamic Time Warping (DTW) on multivariate telemetry time series, we identified international analogs to ten high-density U.S. trunk routes based purely on flight trajectory similarity.

The results surfaced a number of intuitive and compelling patterns - short-haul shuttles like SFO–LAX matched regional pairs such as Bahrain–Riyadh and Ottawa–Toronto, while longer routes like LAX–JFK found analogs in structurally similar transcontinental flights like San Francisco–Toronto. The analogs highlighted how behavioral similarity in airspace usage can transcend geography, reflecting broader structural patterns in global aviation.

Throughout the project, we addressed key challenges such as trajectory preprocessing, temporal standardization, data quality filtering, and the interpretability of DTW scores. We also identified and excluded problematic results when necessary, demonstrating the importance of both quantitative and qualitative validation in trajectory analytics.

While born of curiosity, this framework has applications beyond just that. It could be extended to anomaly detection, route planning, or simulation by identifying flight behaviors that deviate from expected patterns. Future work might incorporate clustering techniques, dimensionality reduction, or learned trajectory embeddings to further enrich route representations and uncover latent structure in global airspace.