



Project Peregrine: Enhancing Public Sector Research in Global Transparency and Anti-Corruption Through Open Aviation Data

Peter Kirechu and Jake Glass

Center for Advanced Defense Studies (C4ADS)

Abstract

ADS-B technology is expanding the amount of flight data made available, in the open source, for public sector research. Nowhere is this benefit more visible than in anti-corruption and global transparency research where ADS-B derived flight data is helping expose the airborne movements of corrupt global elites, dodgy public officials, and a vast array of illicit and criminal actors including wildlife traffickers and conflict financiers.

1 Introduction

Open aviation data is an under-utilized source of information that can significantly complement the work of anti-corruption activists, conflict researchers, and investigative journalists around the world. By law and for safety reasons, most aircraft are outfitted with conventional radar systems (traditionally surface-based radar technologies such as Mode-S systems that have been progressively replaced by Automatic Dependent Surveillance Broadcast (ADS-B) technology). ADS-B systems transmit real-time GPS-based positions to both ground stations and other aircraft, which improve overall aircraft visibility and reduce the risk of air collisions.

But beyond these improvements in aviation safety, ADS-B technology has expanded the amount of flight data made available, in the open source, for public sector research. Nowhere is this benefit more visible than in anti-corruption and global transparency research where ADS-B derived flight data is helping expose the airborne movements of corrupt global elites, dodgy public officials, and a vast array of illicit and criminal actors including wildlife traffickers and conflict financiers.

Outside state governments, open flight data is generally only collected by hobbyists and aircraft enthusiasts, or by commercial companies who collect, filter, and strip the data, primarily to track commercial flight arrivals and departures. Some commercial providers also offer anonymity to state governments, well-resourced oligarchs, or other aircraft owners willing to pay high premiums to block viewing of their transit activities. This filtering process creates significant gaps in the data currently available to the public – especially in areas such as offshore tax havens, geographic trafficking corridors, and global conflict zones where the public sector interest in transparency is most acute.

However, the Center for Advanced Defense Studies (C4ADS), along with its partner ADS-B Exchange, is developing an unfiltered open source flight tracking and analysis capability customized specifically for activists and journalists to add to their investigative toolkit. This platform is issue-agnostic and can be used widely for a variety of research and investigative tasks beyond the basic search and historical data viewing currently offered by most commercial providers. Additionally, through this collaborative effort, both C4ADS and ADS-B Exchange are applying machine-based learning algorithms that are designed to identify specific typologies in aircraft movement that suggest potentially illicit air transit activities and behavior.

2 ADS-B Applications in Public-Sector Research

C4ADS uses data and technology in innovative ways to combat transnational illicit networks. To do so, we focus around “vector vulnerabilities,” or the points along the intersection of the licit and illicit worlds, where criminal actors are forced to nest their illicit activities within the licit systems of trade, transport, and finance, thereby leaving behind a discoverable digital trail. Over the past several months, C4ADS has narrowed in on the movement of stolen assets and proceeds of corruption as a specific vector vulnerability where open flight data (including ADS-B derived flight information) has proven most useful to global transparency and anti-corruption research.

Corrupt public officials, illicit actors, and well-resourced elites often hide assets in secrecy jurisdictions and offshore tax havens such as Switzerland, the Bahamas, the British Virgin Islands, Labuan, Panama, and Beirut. They also frequent these locations, often by private air travel, generating traceable data points that are captured through open-source flight monitoring platforms that researchers can use to recreate their movements, in part or in whole.

In the past two months, C4ADS researchers delved into the airborne movements of Equatorial Guinea’s Vice President Teodoro Nguema Obiang Mangue as well as Cameroon’s long-running President Paul Biya, and discovered a pattern of flight movements that converge on several European destinations, including Switzerland, where both public officials reportedly hold luxury assets allegedly acquired through illicit means (Kirechu, 2019). The tracking of the arrivals and departure activities of reportedly crooked public officials is just one of many investigative options made possible by the proliferation of ADS-B flight data. It is, however, just one of many analytic capabilities under development at C4ADS.

3 Actor Attribution Challenges in Flight Data

There are several challenges in developing a flight data analysis platform that connects the known world of aircraft surveillance broadcasts to behavior of persons and organizations in the open source. Such a tool must allow analysts to efficiently search and wade through a large historical archive of processed ADS-B data to gain analytical insights, and that requires heavy optimization to make terabytes of raw data useful for research. ADS-B must also be paired with outside data sources that can logically tie the technical data provided by Mode-S broadcasts – International Civil Aviation Organization (ICAO) and path information – with registration numbers and aircraft ownership information. On its own, Mode-S transponder data is of limited use to investigators without links to known entities.

Civil aircraft registration data, including identifying information for air-frames and the contact information for their corporate or personal owners, is thus extremely useful in grounding position data in the licit world of our investigations. As a practical investigative tool, however, it

is stymied by limited public availability and corporate ownership structures that often obfuscate ultimate beneficial ownership (UBO). Not all countries offer this data in a public and search-enabled form, and given the variety of formats in use for the information when it is made available, consolidating a global database is a data challenge. Tracing changes in ownership over time – and ascertaining how to best store and visualize those in a consistent ontology – is difficult even with FAA-provided data on US-registered aircraft.

When aircraft registration data is lacking or obfuscated, open source investigation using social media accounts can provide similar insights into relationships between aircraft and owners or users. Not every photo an investigator finds on Facebook is likely to be a dictator standing next to their wholly-owned private jet used exclusively for their travel, but instances like these do exist. Once we have established a relationship between an individual or organization of interest and an aircraft, we may be able to track their travel using freely available ADS-B data. In the case of politically exposed persons (PEPs), this connection also lets us tie news events, such as diplomatic appearances or public meetings, with flight segments. When we can identify known flights for public meetings, we can analyze flights to destinations of interest or visits that do not appear in public reporting to gain context around suspicious activity.

An ideal manifestation of an investigative tool for flights would also include the ability to trace ultimate beneficial ownership of aircraft in corporate registries. This integration in particular must be a cross-cutting dataset; the use of holding companies in secrecy jurisdictions or simply companies registered in a different place from where the aircraft is registered often makes Ultimate Beneficial Owner (UBO) discovery a challenge.

4 Actor Attribution Challenges in Flight Data

The past decade has seen a significant expansion in publicly-accessible corporate registers, allowing open source investigators to more effectively match aircraft to owners, managers, and operators. In the United States, the Federal Aviation Administration (FAA) provides public access to aircraft ownership records through an online portal that allows users to search for an aircraft’s registration number as well as other actor-revealing identifiers. FAA records contain specific information on aircraft owners to include names and addresses, and in some cases provide historical information on past owners. These data points are critical to the historical mapping of an aircraft’s ownership record within and across the United States. While some illicit actors often register their aircraft through corporate service providers, trusts, and/or foundations, others commonly use ownership companies that are traceable through publicly accessible registries.

Take the case of Samer Foz, a Syrian oligarch and favored commercial affiliate of Syrian President Bashar Al-Assad who was sanctioned on June 11, 2019, as a war profiteer by the United States Department of Treasury’s Office of Foreign Assets Control (OFAC) (OFAC, 2019). Samer Foz was sanctioned alongside siblings Hosen and Amer Foz, and together the Foz family operates dozens of companies in Lebanon, Turkey, and Syria. (Dressel, 2019; OFAC, 2019) At least one of his linked companies, Alaman for Jets Limited – registered in the British Virgin Islands – was identified as the owner of a Dassault Falcon 2000, registration number M-SFOZ (ICIJ, 2015)

While plane spotting websites Plane Spotter and Jetphotos placed the aircraft in Germany, Turkey, and Italy from 2016 to 2018, limited information about its flight histories is freely available through commercial flight data providers (M-SFOZ, 2016; M-SFOZ, 2017; M-SFOZ, 2018). These data providers also rarely catalogue historical ownership information limiting the amount of forensic ownership discovery possible without incurring extensive costs.

However, like most aircraft, the M-SFOZ is registered to a national civilian aircraft registry (in this case the Isle of Man), whose civilian registry is open and accessible to the public (S-FOZ, 2014). A search through the Isle of Man aircraft registry confirms the aircraft's dates of registration, its registered owners (current and former), its previous registration numbers, and its former jurisdictions of registration (Turkey, France, and Lebanon) (F-ORAX, 2004; F-HFLX, 2014; TC-ANG, 2019). Forensic tracing of the aircraft's ownership records reveals additional ownership information in secrecy jurisdictions including Seychelles, the British Virgin Islands, and the United States (N78NT, 2006; M-TANA, 2011). Most of the ownership information in these jurisdictions is hidden from public view though the circuitous change in ownership creates a potentially viable detection heuristic for aircraft that display similar ownership change behavior.

In the absence of adequate historical ownership data, open flight data is the remaining investigative reserve for open source researchers. In recent years, global adoption of ADS-B technology has expanded the reservoir of aircraft transit data that is available for both technical and non-technical public sector analysis. Challenges abound in the application of ADS-B data to public sector research, but in the case of Samer Foz and the M-FOZ Dassault 2000 aircraft, open flight tracking data revealed the aircraft's transit activities in Lebanon, Italy, the United Kingdom, and Syria, some of which occurred even after his OFAC sanctions designation. The ability to identify and track the flight activities of designated aircraft and their owners has expanded the threshold of what is possible for open source researchers.

The next frontier of open source research stretches beyond active and historical flight viewing and is instead focused on the development of detection heuristics that identify anomalous behavior potentially linked to illicit activities.

5 Machine-Aided Activity Recognition

Machine learning is not a passive process that can automatically detect every single anomalous flight pattern, but it can be a useful tool in lead generation by extrapolating from known entities and patterns of behavior to previously undetected ones. Project Peregrine seeks to combine perceptual models from our analysts and other open source investigators with machine learning techniques – for example, path clustering, flight data anomaly detection, and classification of unknown landing sites or runways. We are building a platform that integrates these basic machine learning tools and also allows analysts on their own to quickly develop and test models that comb through a large archive of historical flight data.

For cases in which we track politically exposed persons, we can use public reporting and the results of previous investigations to identify destinations or flight paths of interest that have been associated with known illicit activity, along with aircraft controlled or used by their organizations, governments, or collaborators. A simple rule-based heuristic in the platform can then alert analysts of new leads. A timed job automatically traverses flight data for the aircraft of interest and triggers an alert when a specific flight path occurs. Running this type of heuristic engine over historical ADS-B data can illuminate patterns of behavior for PEPs ranging from dictators to potentially corrupt mid-level government officials. Investigators can analyze these flight paths in concert with news, social media, corporate records, property registries, other public reporting in order to identify locations of asset flight or illicit operations.

In addition to working from known aircraft of interest and using their flight tracks to identify patterns of behavior, applying machine learning and analyst-developed models on large Mode-S datasets allows us to identify patterns of behavior and match aircraft of interest – which will warrant further investigation – to these patterns. The main challenge we face in platform

development is how best to make this method of analysis available directly to tech-enabled analysts who themselves are not data scientists.

Training a machine learning model that matches a known pattern of behavior against troves of historical data is typically implemented programmatically in a distributed computing cluster, a process which requires developer and data scientist support unavailable to most investigators. We have settled on an imperfect but still powerful solution of developing models of “suspicious” behavior and running them on historical and live data, then enriching Mode-S flight tracks with our catalogued “suspicious” aircraft behavior tags. For our purposes, these models so far encompass automated detection of surveillance flights and flight path clustering. We plan to extend these capabilities to include detection and classification of unknown landing sites and GPS spoofing/jamming. This way, analysts can access prepared information derived from the raw data by us as developers, while preparing their own rule-based models to identify new “suspicious” aircraft activity for future investigations. More detailed rule-based analysis available in the application will include custom pattern recognition, including abrupt changes in trajectory, changes in flight information broadcast, or uncharacteristic descents/ascents to/from unrecognized runways.

The automatic recognition of anomalous flight conduct goes beyond corruption research and can be applied to the identification of geographic trafficking corridors where small aircraft, potentially carrying narcotics, take off or land in remote/unregistered airfields. This analysis capability is also applicable to conflict zones where flight data can substantially enhance evidence-based reporting on air operations conducted during conflict.

Elsewhere, automatic recognition of anomalous flight activity has been used by investigative journalists to identify surveillance activity across the United States. In 2017, BuzzFeed journalists analyzed Mode-S data provided by FlightRadar24 using a random-forest model trained to detect sharp banks in flight (Aldhaus, 2017). They successfully used this model to identify FBI and DHS surveillance flights over the United States. Not only did they identify previously-known federally-operated planes, their model’s in-flight behavior analysis allowed them to flag twenty aircraft registered to local governments in Arizona and California, and to the Ohio State Highway Patrol.

Together, machine-aided activity recognition and traditional ownership identification methodologies allow researchers to match aircraft flight activities to ownership data (where available), invariably closing the loop between aircraft movement and both owners, users, and operators. This data can then be transformed into high risk watchlists used by investigators to automatically identify and track planes of interest depending on the characteristics of the aircraft, its movements, and/or its known users.

When combined with geographic filters and real time alerting capabilities, all of this data can be harnessed to create robust rule engines that provide historic, current, and in the future, predictive recognition of illicit or suspicious aircraft conduct across time and space. Altogether, this approach expands the breadth of public research tools and capabilities that are available in the open source, and at limited to no cost, to global transparency and accountability research.

6 References

Aldhouse, P. (August 7, 2017). We Trained A Computer To Search For Hidden Spy Planes. This Is What It Found. BuzzFeed. Retrieved from: <https://www.buzzfeednews.com/article/peteraldhaus/hidden-spy-planes>

Dressel, T. (September 16, 2019). The Henchman’s Dilemmas: The Logic of the Samer Foz Network. *Blog*. Center For Advanced Defense Studies. Retrieved from: <https://c4ads.org/blog>

Erol, E. *Image: M-SFOZ*. (April 16, 2016). Jet Photos. Retrieved from: <https://www.jetphotos.com/photo/8277870>.

F-HFLX. (December 10, 2014). Extract from Register of Aircraft. Directorate General of Civil Aviation. Retrieved from: https://immat.aviation-civile.gouv.fr/immat/servlet/aeronef_liste.html

F-ORAX. (March 19, 2004). Extract from Register of Aircraft. Directorate General of Civil Aviation. Retrieved from: https://immat.aviation-civile.gouv.fr/immat/servlet/aeronef_liste.html

International Consortium of Investigative Journalists (ICIJ). *Alaman for Jets*. Offshore Leaks Database. Retrieved from: <https://offshoreleaks.icij.org/nodes/10155564>

Kirechu, P. (July 30, 2019). Assets in Flight: Cameroon's Absentee President. *C4ADS Blog*. Center for Advanced Defense Studies (C4ADS). Retrieved from: <https://c4ads.org/blogposts/2019/7/30/assets-in-flight-2>

Kirechu, P. (July 9, 2019). Assets in Flight: Tracing Africa's Corrupt Ruling Elite in the Skies. *C4ADS Blog*. Center for Advanced Defense Studies. Retrieved from: <https://c4ads.org/blogposts/2019/7/9/assets-in-flight>

M-SFOZ. (June 10, 2014). Isle of Man Aircraft Registry. *ARDIS*. Retrieved from: <https://ardis.iomaircraftregistry.com/register/696/view>

M-TANA. (June 16, 2011). Isle of Man Aircraft Registry. *ARDIS*. Retrieved from: <https://ardis.iomaircraftregistry.com/register/376/view>

M-ZLAM. (April 17, 2015). Isle of Man Aircraft Registry. *ARDIS*. Retrieved from: <https://ardis.iomaircraftregistry.com/register/787/view>

N78NT. (November 22, 2006). Federal Aviation Administration, *N-Number Inquiry*, Retrieved from: https://registry.faa.gov/aircraftinquiry/NNum_Results.aspx?NNumbertxt=78NT

Spoladore, A. (May 21, 2017). *Image: M-SFOZ*. Jetphotos. Retrieved from: <https://www.jetphotos.com/photo/8624557>

Schmidt-Blindenhöfer, T. (August 16, 2018). *Image: M-SFOZ Alaman For Jets Ltd Dassault Falcon 2000*. PlaneSpotters. Retrieved from: <https://www.planespotters.net/photo/877707/m-sfoz-alaman-for-jets-ltd-dassault-falcon-2000>

TC-ANG. Turkish Civil Register. (January 11, 2019). Retrieved from: <https://www.ole-nikolajsen.com/TURKISH>

US Department of the Treasury. Treasury Designates Syrian Oligarch Samer Foz and His Luxury Reconstruction Business Empire (June 11, 2019). Office of Foreign Assets Control. Retrieved from: <https://home.treasury.gov/news/press-releases/sm704>

US Department of the Treasury. *Graphic: Samer Foz Network and Lebanon Oil-based Companies*. (June 11, 2019). Office of Foreign Assets Control. Retrieved from: <https://home.treasury.gov/sites/default/files/136/Syria>