OpenSky Report 2022: Evaluating Aviation Emissions Using Crowdsourced Open Flight Data

Junzi Sun†*, Luis Basora‡, Xavier Olive‡*, Martin Strohmeier*¶‖,
Matthias Schäfer*§**, Ivan Martinovic*¶, Vincent Lenders*‖

*OpenSky Network, Switzerland
lastname@opensky-network.org

†TU Delft, The Netherlands
j.sun-1@tudelft.nl

‡ONERA, Université de Toulouse, France
firstname.lastname@onera.fr

§TU Kaiserslautern, Germany
schaefer@cs.uni-kl.de

¶University of Oxford, UK
firstname.lastname@cs.ox.ac.uk

‖armasuisse, Switzerland
firstname.lastname@armasuisse.ch

**SeRo Systems, Germany
schaefer@sero-systems.de

***Abstract*—The environmental impact of aviation has become the focus of increased concerns for policymakers around the world. The recent pandemic provided many interesting case studies on the impact of aviation on the environment. Following the initial COVID-19 containment measures and hard lockdowns, the sharp decrease in aircraft movements caused a measurably improved air quality worthy of further study.**

The OpenSky Network has acted as an important open data source for aviation research since 2013. In this paper, we analyze one year of fine-grained pre-COVID air traffic trajectories (comprising the entire year 2018) to estimate fuel consumption and pollutant emissions in the aviation industry. We compare this large-scale big data processing approach to a reduced model approach based solely on global commercial aircraft movement schedules collected from airlines and airports, aggregated by a commercial provider.

Our study quantifies the impact of commercial aviation on global emissions. The numbers reveal that aviation's CO_2 emissions contribute to 2% of global emissions and that commercial aviation contribution remains a proxy for countries' wealth.

I. Introduction

With record-breaking heatwaves and yearly temperature records, the aviation industry's climate-impacting emissions have been drawing the acute attention of policy-makers around the globe. While efforts to decrease the climate impact of flying are underway, from synthetic fuels and electric aircraft to more efficient traditional engines, these solutions will still require many years to come to fruition.

However, to tackle these sustainability challenges in a scientific and data-driven manner, we first require accurate representations of global flight emissions. Such representations require both new models and large-scale high-quality data to feed them. In the spirit of open science, these data and models should be openly accessible so that they can be scrutinized and built upon. We tackle this problem by providing a new open data study on global aviation emissions in this paper.

To obtain the detailed global flight trajectory and speed information that we need, we turn to Automatic Dependent Surveillance-Broadcast (ADS-B) as a source technology. With the proliferation of crowdsourced ground receivers combined into large sensor networks, ADS-B has become a valuable source for aviation research. Established in 2013, the OpenSky Network [1] is a dedicated non-profit ADS-B network that provides such flight data to the aviation research community. With the increasing number of ground receivers, OpenSky's data has enabled many aviation research studies (over 300 to date), including previous OpenSky reports [2]–[4].

By combining OpenSky's ADS-B data with the open aircraft performance and emission model OpenAP [5], we perform a global study of aviation emissions in this paper. We study different emissions, including CO_2, H_2O, NO_X, and SO_X, within the coverage of OpenSky network receivers.

For our study, several pieces must be in place. OpenSky offers so-called state vectors, which contain the ground speed of aircraft. Wind models are obtained from Global Weather Forecast reanalysis data and combined with the ground speed to approximate the true airspeed of flights. Aircraft weight is an unknown performance parameter that affects the estimations. In this paper, we make simple assumptions based on several possible load factors. With assumed mass, calculated airspeed, and other flight conditions obtained from ADS-B, we can estimate the emissions for all trajectories observed by the OpenSky network.

Due to its extremely large quantity, we perform an analysis of all captured global flights for one year (2018), with trajectories down-sampled to one point per 30 seconds. The large volume of detailed ADS-B data in some areas and its limited availability in other regions motivates the use of complementary methods like FEAT (Fuel Estimation in Air Transportation) [6] and datasets such as OAG [7]. FEAT estimates the fuel consumption by leveraging a reduced model and exploiting flight frequencies and great circle distances between airport pairs from the historical dataset provided by OAG, a global travel data provider. An additional contribution of this paper is to assess how FEAT estimations compared with the ones performed with OpenSky data.

The remainder of this work is structured as follows. Section II provides the necessary background on our data and its collection. Section III introduces the methodology for estimating aircraft emissions at scale, followed by detailed results in Section IV. Section V analyzes the results and addresses the limitations of our study. Section VI discusses insights and takeaways from our work before Section VII concludes.

II. Background

A. The OpenSky Network

The OpenSky Network is a crowdsourced sensor network collecting surveillance data for air traffic control (ATC). Its objective is to make real-world ATC data accessible to the public and support the development and improvement of ATC technologies and processes. Since 2013, it has continuously been collecting air traffic surveillance data. Unlike commercial flight tracking networks (e.g., Flightradar24 or FlightAware), the OpenSky Network keeps the raw Mode S replies as they are received by the sensors in a large historical database, which can be accessed by researchers and analysts from different areas.

The non-profit network started with eight sensors in Switzerland and Germany and has grown to more than 5000 registered receivers at locations all around the world. At the time of writing, OpenSky's dataset contains over eight years of ATC communication data. While the network initially focused on ADS-B only, it extended its data range to the full Mode S downlink channel in March 2017 and more recently other technologies such as FLARM and VHF. The dataset currently contains more than 30 trillion Mode S replies and during peak times receives more than 20 billion messages per day.

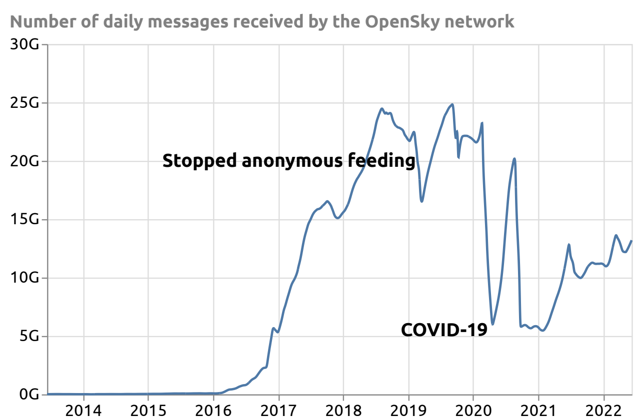

Fig. 1: The growth of OpenSky's dataset over time from 2013 to March 2022

Figure 1 shows the growth and development over the past several years, which saw the inclusion of the `dump1090` and Radarcape feeding solutions and the integration of non-registered, anonymous receivers. This practice has been discontinued in early 2019 to further ensure the quality of the feeder data. In March 2020, the number of daily messages dropped to about 30% from the previous level, reflecting the curtailment of air travel around the world due to the COVID-19 pandemic. While the air traffic numbers have returned almost to pre-pandemic levels in some areas, the number of daily messages has not. This is due to an optimized collection process that emphasizes the de-duplication of messages earlier in the collection process. This approach ensures the continued growth of sensors around the world but means that numbers are not strictly comparable over time due to breaks in the data.

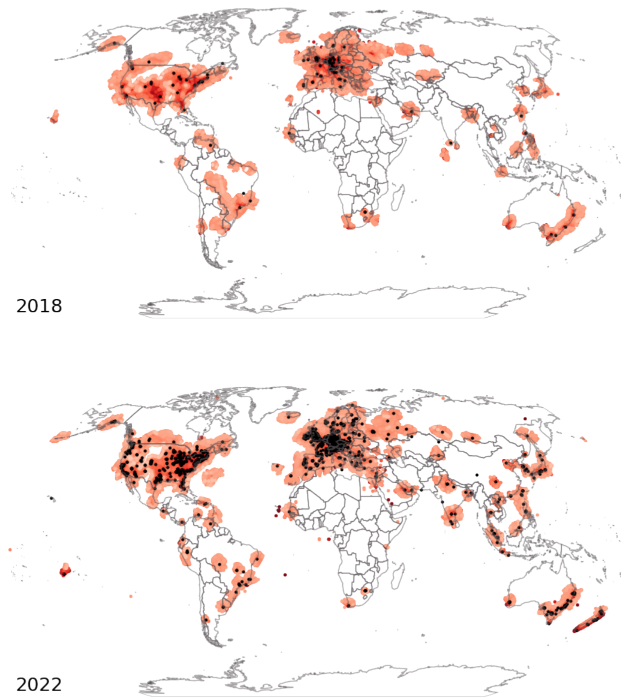

Fig. 2: OpenSky's global coverage in 2018 and 2022

The global data reception of the OpenSky Network fully depends on its crowdsourced network of receivers, comprised mostly of enthusiasts, academics, and some companies. The coverage of any single sensor is limited by the line-of-sight range of the antennas (about 400-500 km for the best-performing ones reaching the radio horizon). This means that such a crowdsourced network's organic growth is effectively a proxy for densely-populated wealthier areas around the world. Between 2018 and 2022 (Figure 2), the global footprint of the coverage thus reached a certain saturation point, new sensors mostly increased lower altitude reception in already covered areas in Europe, the US, and other industrialized countries. Notable coverage extensions can still be seen in the Middle East, East Asia, and New Zealand. Desert areas and oceans are naturally lacking ground-based coverage due to their physical limitations. Commercial ADS-B providers partly address this shortcoming with space-based ADS-B [8].

Besides the payload of each Mode S downlink transmission, OpenSky stores additional metadata. Depending on the receiver hardware, this metadata includes precise timestamps (suitable for multilateration), receiver location, and signal strength. For more information on OpenSky's history, architecture, and use cases refer to [1], [9] or visit the website https://opensky-network.org.

B. Description of the OAG dataset

The Official Aviation Guide (OAG) [7] sells curated aviation data about scheduled flight movements and some airport ground transportation links covering several years of global city-pair operations.

The OAG data for 2018 used for this study contains around 47.2 million lines. Each line corresponds to a flight (or an airport ground connection movement) and includes the carrier, airport origin and destination IATA codes, aircraft type IATA code, dates-times, and great circle distances.

The dataset contains all commercial flights that were scheduled in 2018, including some military and helicopter operations. However, it does not contain general aviation, and freight traffic is limited.

Preprocessing of the dataset is necessary to align the provided IATA codes with the ICAO codes used by our models, as well as to remove duplicated and wrong flight information (see Section V-F for further details).

III. Methodology

A. Data collection and processing

For this study, we collect all flights recorded by the OpenSky network during the year 2018, resulting in approximately 2.5 terabytes of data containing global state vectors sampled every 30 seconds. The data is then cleaned with unused columns removed and converted to a compact *parquet* format and organized per day.

All flight trajectories are processed using the `traffic` Python library [10], where sampling and filtering are applied. Then, we estimate the fuel consumption and emissions using OpenAP [5] based on the state vector data of each trajectory. One of the challenges raised from the data processing is to properly handle flights spanning two consecutive days. The other challenge is the computation resources and time required for estimating the emissions, since the fuel flow estimation for each flight has to be computed sequentially, updating the estimation of the aircraft mass at every timestamp.

To cope with such an immense computational task, we leveraged the *DelftBlue* supercomputer provided by TU Delft [11], with can reduce the computation time from approximately two months to one week. Once the emissions are calculated, we generate the aggregated statistics per day and use that for the analysis of the paper. Figure 3 details the process of handling the OpenSky data and estimating the emissions from flight trajectories.

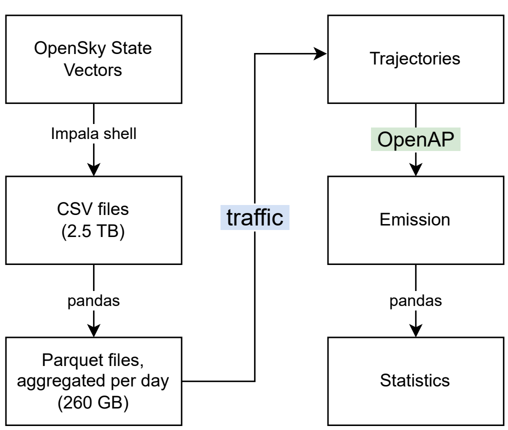

Fig. 3: Data processing and emissions calculation

B. Estimation of fuel consumption based on ADS-B tracks

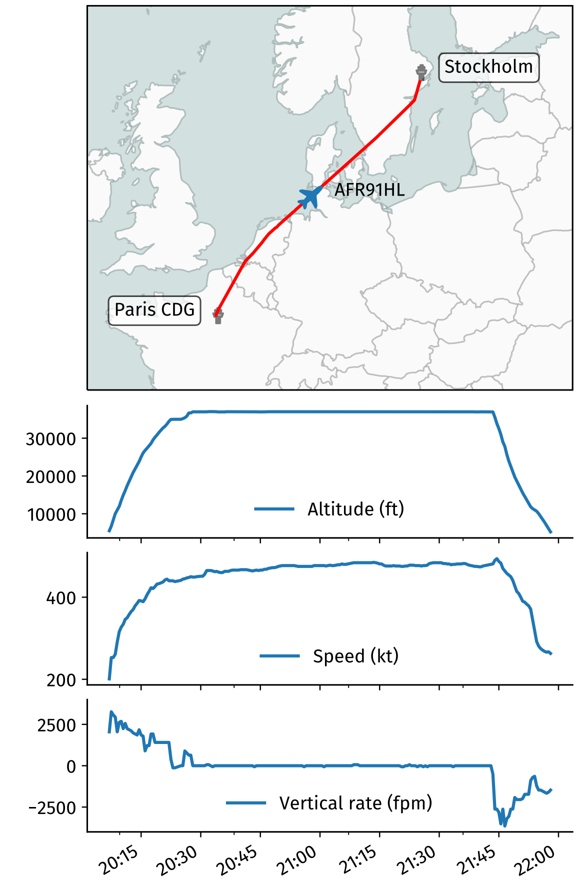

Fig. 4: Flight trajectory of an example flight

Figure 4 shows an example flight trajectory and the main states we use for the emissions estimation. We calculate the flight emissions using the OpenAP library, which provides the necessary aircraft performance and emissions models to estimate fuel consumption and emissions with real trajectory data. Different gas emission types for common aircraft types are considered, including CO_2, H_2O, NO_X, SO_X, CO, and HC.

Currently, OpenAP defines the aircraft performance and emissions models for more than 20 common aircraft type codes. For less common aircraft types, we choose a similar type based on a synonym database defined in OpenAP. In total, we identify more than 22 000 aircraft (tail numbers) across more than 30 aircraft types (53 types with synonyms).

In addition to the flight trajectory data, the calculation of fuel flow and emissions also requires information on aircraft mass. Since aircraft mass information is not publicly available, mass is often inferred based on flight trajectories. However, the process requires expensive computations [12], [13]. In our case, when dealing with global flights, simple assumptions have to be made. To this extent, we estimated four different emission values based on four different assumptions of mass for each flight. Similar to a previous study [4], these masses are defined at 60%, 70%, 80%, and 90% of the aircraft's maximum take-off weight.

The fuel flow is linearly correlated to CO_2, H_2O and SO_X. For other types of emissions, OpenAP uses and extends upon the ICAO Aircraft Engine Emissions Databank [14] and Fuel Flow Method [15]. The first model provides the base emissions of engines at the static test and the second model provides corrections of emissions at different flight conditions related to altitudes. With these models, NO_X, CO, and HC emissions for trajectories can be calculated.

C. FEAT reduced model with OAG

As an alternative to estimating fuel and CO_2 consumption from ADS-B tracks, the FEAT method [6] is a reduced model approach. This method is particularly adapted to speeding up computation when estimations are needed at the global scale every year involving millions of flights, as well as when different traffic scenarios need to be compared. In FEAT, payload, engine degradation, and flight route inefficiencies are approximated to derive a reduced-order distance-fuel relationship based on a quadratic regression function.

For each aircraft type, FEAT generates a set of flight profiles with the BADA model [16] covering the operational range of the aircraft. Then, it estimates the fuel consumption for each profile with BADA. This results in a series of point estimates on which a quadratic curve can be fitted. Once the regression coefficients per aircraft type are calculated, yearly global fuel consumption and CO_2 can be computed in a matter of milliseconds by exploiting city-pair traffic frequencies available in datasets like OAG.

In addition to computational efficiency, the combination of FEAT and OAG is useful to overcome the lack of ADS-B coverage in some regions of the planet. It should be noted however that OAG contains only scheduled flights with a very limited volume of freighter operations.

D. OpenAP, FEAT reduced model, and OpenSky data

FEAT can be used independently of the OAG dataset. We extract from the OpenSky data the aircraft ICAO transponder codes and link them to aircraft type codes, city pairs, and flight distances for each type of aircraft. We perform a fuel consumption analysis between OpenAP and FEAT using these aggregated flight distances and total great circle distances.

IV. Results

In this section, we compare the global fuel estimation and CO_2 for 2018 obtained by FEAT with the OAG dataset with the one from OpenAP with full OpenSky ADS-B trajectories.

The 2018 OAG dataset contains around 47.2 million in lines corresponding to city-pair movements including airport ground connections with trains or buses, which we removed to focus only on air traffic. Helicopters and a few military and old aircraft were also deleted, as they could not be matched with a FEAT reduced model.

Pre-processing also included the aggregation of flights by origin, destination, and aircraft type code, with the number of flights and great circle distance associated. In a few cases, it was also necessary to remove duplicated information and to correct invalid distances for certain aircraft types and city pairs.

We also aligned the IATA aircraft type codes used in the OAG dataset with the ICAO ones used by the FEAT models. In some cases, one IATA type code corresponds with several ICAO codes (e.g. IATA code `32S` → `A318/A319/A320/A321`), so we estimated a traffic distribution of the ICAO codes per origin-destination pair based on either historic ADS-B data or by assuming a uniform distribution when ADS-B data was not available for a specific origin-destination pair.

After pre-processing, over 38.4 millions of flights were left. The total fuel consumption and CO_2 emissions calculated for 2018 are 261 Mt and 818 Mt respectively, which is close to the 257 Mt and 812 Mt published in the original FEAT paper [6].

TABLE I: Total fuel consumption and emissions estimated with FEAT reduced model (based on OAG schedules) and with OpenAP (based on OpenSky Network data). Flights with negeative HC and CO values were removed from the CO and HC emission statistics.

	FEAT OAG data	**OpenAP OpenSky data**
Fuel consumption (in Mt)	261	141
CO_2 emissions (in Mt)	818	444
H_2O emissions (in Mt)	·	173
NO_X emissions (in Mt)	·	2.54
SO_X emissions (in Mt)	·	0.18
CO emissions (in Mt)	·	0.322
HC emissions (in Mt)	·	0.021

V. Analysis

A. Limitations regarding coverage

The coverage of the OpenSky Network (Figure 2) relies on the positions of feeding receivers from the community, which roughly correlates with the most inhabited and wealthy areas of the planet. Oceans and deserts, as well as most of Africa lack coverage. Space ADS-B, exploited by commercial ADS-B data providers, has been a recent approach to address the lack of coverage in those areas, but it was not available for our study due to the high cost. It should also be noted that space-based ADS-B largely works for the en-route airspace and not in terminal areas.

In this analysis, we addressed lacking coverage by filling gaps over oceans by interpolating trajectories along the great circle. This approach should be sufficient for transatlantic flights but may be lacking realism in the Pacific Ocean where coverage is sparser.

The OAG dataset could be helpful to estimate how much we miss. However a systematic link between this dataset—based on scheduled commercial flights, listed as IATA flight numbers—and the ADS-B data— utilizing ICAO callsigns, a low sampling rate (a feasibility constraint coming with this study), and missing coverage—is difficult. Only a few city pairs were sampled for a non-representative comparison in Section V-E.

As an upside, we note that the OpenSky Network coverage is good exactly in those places around the world where air traffic is most dense (except for mainland China). Hence, we are confident that it provides a representative sample of the most relevant routes and airspaces.

B. Limitations around general aviation

General aviation (GA) is poorly covered by this study for several reasons:

- the OAG dataset only references commercial flights;
- many GA aircraft are still not ADS-B compliant [4];
- GA aircraft models of consumption and emissions are still immature.

As general aviation accounts for a large share of global aviation emissions in terms of CO_2 per passenger per kilometer, it would be worth investigating this area further and developing a performance model for general aviation as well. The progression of ADS-B Out mandates should aid future studies and the integration of other data sources such as UAT and FLARM.

C. Uncertainty due to wind forces

ADS-B messages contain positional information (latitude, longitude, barometric, and GPS altitude) and their derivative (true track angle, ground speed, and vertical rate). In some areas of the world, true airspeed, indicated airspeed, and Mach number can be transmitted, upon request by a secondary surveillance radar, as part of the Enhanced Surveillance (EHS) standard [4].

For this study, we made all fuel flow estimation computations based on the ground speed values which are part of the ADS-B. To validate the soundness of this assumption, we compared the total fuel consumption for 5000 flights covering various distances (short, medium, and long haul flights), once using the ground speed, and once using a true airspeed value interpolated from wind field values found in ERA5 historical data [17].

Figure 5 plots one point per flight and compares the fuel consumption estimation based on ground speed and true airspeed values. As expected, these are mostly centered on the $x = y$ line. Figure 6 considers the error ratio between both values. The positive part of the density plot (in orange) means the assumption *overestimates* the fuel consumption, the negative part (in blue) means it *underestimates* it, by up to $\pm 15\,\%$. The assumption that ground speed can be sufficient may be hard to justify for single flights but should compensate itself for the whole dataset: on this subset, the average value of the difference ratio is $+1.5\,\%$. This remaining positive value may reflect the effect of airlines optimizing their routes to benefit from tailwinds or to limit the negative impact of strong headwinds.

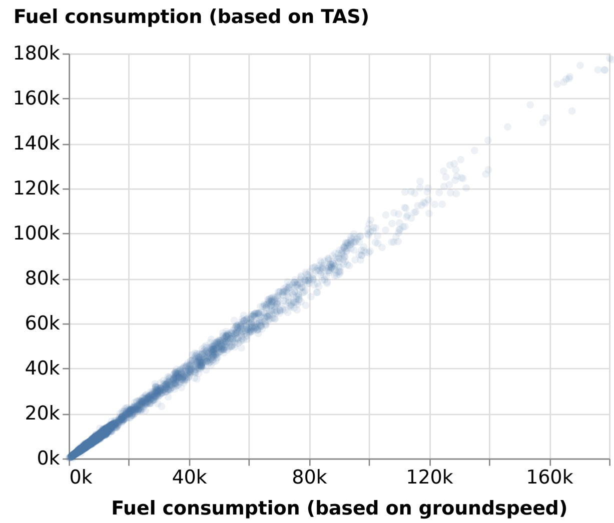

Fig. 5: One point corresponds to one of 5000 flights on January 1st, 2018. Fuel consumption (in kg) using the ground speed matches the x-coordinate, using the true air speed matches the y-coordinate.

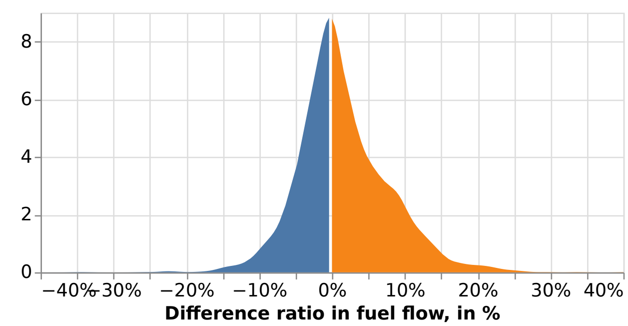

Fig. 6: Distribution of difference ratio between fuel consumption per flight, calculated with ground speed and with true air speed. A positive ratio means that ground speed overestimates the fuel flow.

D. Uncertainty around the aircraft mass

Since aircraft mass is unknown in the OpenSky dataset, we have to make simple assumptions for emissions estimations. We propose to calculate the interval of emissions based on aircraft mass between 60% and 90% of the maximum takeoff mass. For example, the emissions from the example flight (Figure 4) and their uncertainties are shown in Figure 7. We can observe the uncertainties caused by unknown aircraft mass in these estimations.

When all flights are considered, the uncertainties in total emissions can also be calculated by aggregating emission status over the entire dataset.

E. Limitations due to the great circle approach with FEAT

The FEAT approach uses a reduced model to estimate fuel consumption based on the total flown distance. Although FEAT considers trajectories between city pairs as great circles, its quadratic coefficients integrate the effect of a distance correction factor estimated with historical ADS-B data to account for ATM inefficiencies (see Appendix I [6] for further details).

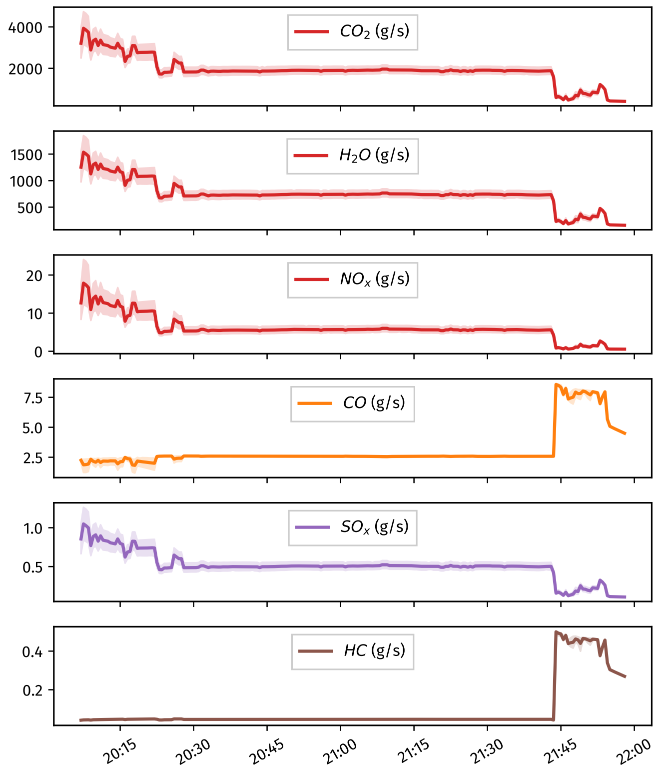

Fig. 7: Emissions estimated based on the example flight

This gives a fairly good approximation when aggregating such a volume of data over a year. Conversely, the OpenAP approach uses the full trajectory to compute an instantaneous fuel flow along time and yields a total fuel consumption different for every day on the same route.

Figure 8 plots a comparison between the FEAT estimation (the circle marks on top) and the distribution of OpenAP estimation as density plots (below) for a few routes. Most frequent aircraft types are represented in different colors.

On most city pairs, FEAT estimations are in general higher than OpenAP estimations. Overall, values seem consistent, but there is no systematic way to relate the value to the ground truth consumption. On the other hand, Figure 9 shows peculiarities that the great circle approximation fails to capture the fact that jet streams' effects on transatlantic flights. Eastbound flights tend to stick to the great circle route, while westbound flights fly more to the North to limit the effects of headwinds. The distributions, despite both being purely based on the ground speed, show different fuel consumption values in both directions.

Fuel consumption in the distributions of Figures 8 and 9 show very different values according to aircraft types. In Figure 10, we compare fuel consumption per passenger for different aircraft types and show that the most recent aircraft tend to burn less fuel and emit fewer pollutants. Values for Airbus A380 are very optimistic, as it has proved difficult to fly with full aircraft of more than 800 passengers.

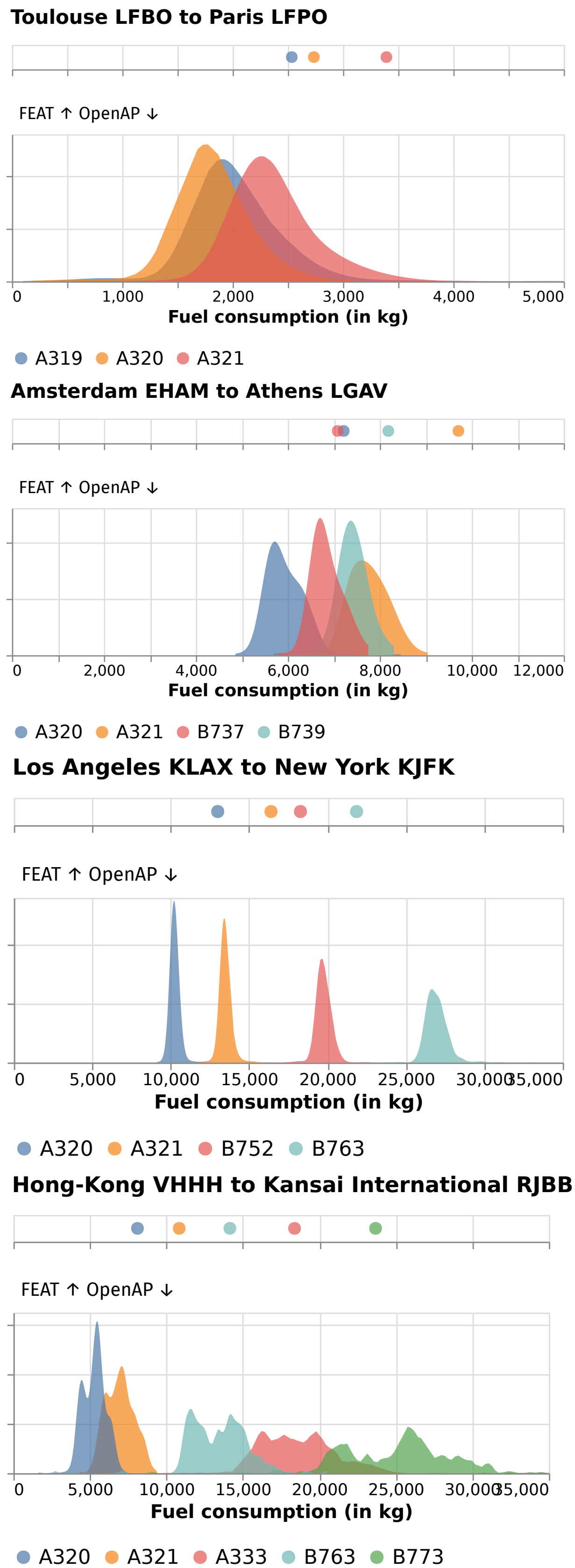

Fig. 8: Comparison of fuel consumption estimations between FEAT (great circle distance based) and OpenAP with ADS-B trajectories, interpolated along great circles when data is missing.

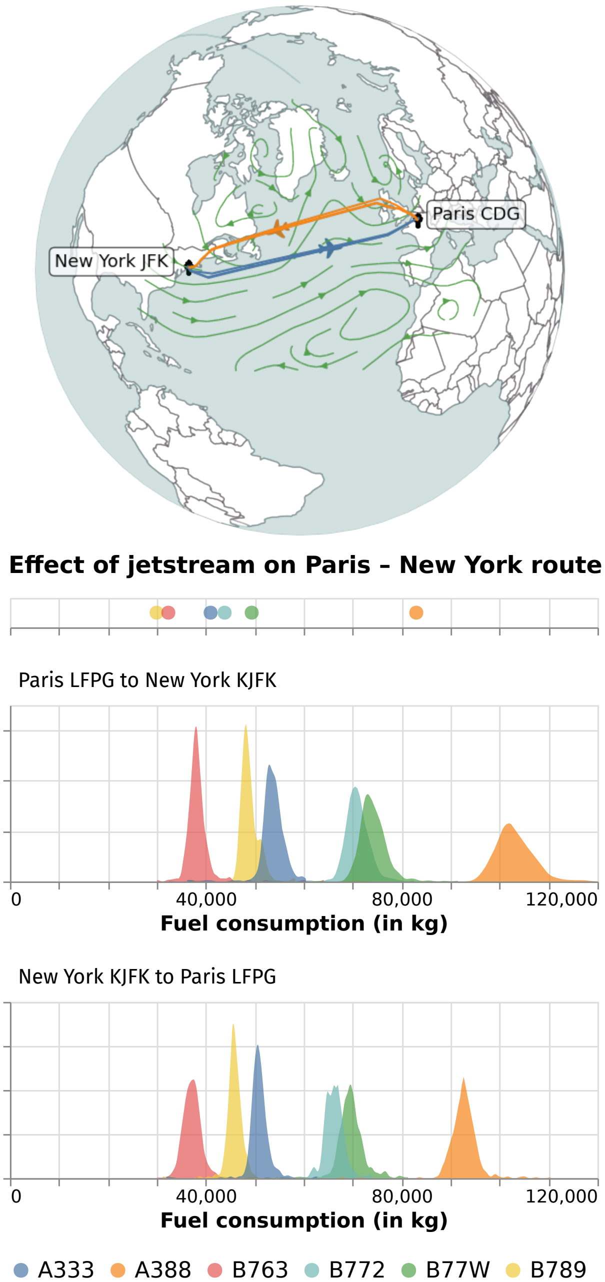

Fig. 9: The great circle approach with FEAT misses the impact of the jet stream effects on trajectory optimisation in its estimation.

F. Comparison between OpenAP and FEAT model

After all flight data from 2018 was processed, we were about to provide a further in-depth comparison between the OpenAP and FEAT model.

For each flight in the dataset, we calculate its fuel consumption with OpenAP considering the uncertainty of takeoff mass, and we also estimated its fuel consumption using the FEAT model. We then aggregated the results for all aircraft model types based on the total distances per aircraft type code.

Table II presents the difference in fuel flow estimations between OpenAP and FEAT for the top 20 aircraft types based on the same flights in the OpenSky dataset. The red color indicates FEAT over estimates fuel consumption compared to the maximum possible fuel from OpenAP, while the blue color indicates FEAT underestimates the fuel consumption compared to the minimum possible fuel from OpenAP. While the final results for all aircraft types combined for OpenAP and FEAT align with each other, there are quite some differences for specific aircraft types.

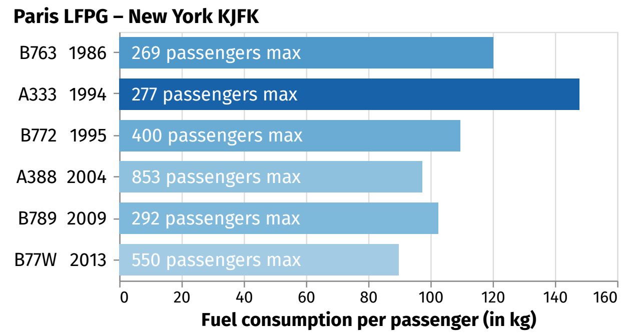

Fig. 10: FEAT estimation of the fuel consumption per passenger on a Paris–New York flight. Most recent aircraft tend to burn less fuel per passenger, but the analysis is only valid if the aircraft is full.

TABLE II: Comparison of fuel estimation of top 20 aircraft types. OpenAP uses total flight distance, and FEAT uses total great circle distances.

type code	distance (Mkm)	OpenAP (Mt)	great circle (Mkm)	FEAT (Mt)
B738	2.78	13.14 - 16.93	2.55	19.70
A320	2.15	8.75 - 11.09	1.98	14.77
A321	1.18	6.05 - 7.85	1.08	9.67
B77W	0.88	16.20 - 19.78	0.82	14.38
A319	0.69	3.13 - 3.94	0.61	4.25
A333	0.58	8.10 - 9.95	0.55	7.88
B739	0.50	2.57 - 3.24	0.47	3.51
B763	0.45	4.73 - 5.67	0.43	5.02
A332	0.43	5.81 - 7.09	0.41	5.59
B772	0.40	7.26 - 8.64	0.38	5.79
B752	0.33	2.51 - 3.16	0.31	2.91
E75L	0.32	0.84 - 1.04	0.28	1.39
B744	0.30	6.35 - 8.03	0.28	6.11
B737	0.30	1.53 - 1.89	0.27	1.96
A388	0.29	6.85 - 8.92	0.28	7.75
B789	0.27	3.15 - 3.83	0.25	2.60
B788	0.25	2.22 - 2.76	0.23	2.38
B77L	0.22	4.21 - 5.21	0.21	3.11
A359	0.21	3.17 - 3.81	0.20	2.72
A20N	0.21	0.73 - 0.95	0.19	1.24
Total	12.738	107.30 - 133.78	11.793	122.720

VI. Discussions

In this paper, we carry out an unprecedented analysis of fuel consumption and pollutant emissions estimations based on:

- an open-source aircraft performance model (OpenAP) run on a global crowdsourced network of trajectory data

(OpenSky Network). Trajectories are interpolated along with great circles in areas with partial coverage. A range of mass is used, and the effects of wind are ignored;
- a reduced model (FEAT) run on global aircraft movements schedules (OAG). This approach approximates trajectories with a simplified mass model and great circle distances, while ignoring the effects of altitude and wind.

Fuel and CO_2 estimations are consistent between the two approaches and broadly validate the approach of using aircraft schedules to substitute for missing coverage. For pollutants other than CO_2, a comparative analysis between the two approaches is not possible, as there is no way to estimate these pollutants with FEAT so far. In addition, the final results do agree with another recent study [18].

The resulting order of magnitude is 261 megatons of fuel, i.e. 818 megatons of CO_2 for the impact of commercial aviation on the environment in one pre-covid year. This is roughly 100 kilograms of CO_2 emissions per capita (worldwide). In comparison, the CO_2 emissions per capita in 2018 was 4.80 tons worldwide, but few countries (e.g. Chad, Rwanda, Burundi, Somalia) emitted less than 100 kilograms per capita that year.

Without surprise, wealthier countries are responsible for the carbon burden of commercial aviation. The same can be seen for emissions produced by private or business jets. In addition, general aviation aircraft remained out of this study because of lacking performance models, which can be a future area of research focus.

In this paper, we use precomputed coefficients provided with the original FEAT [6] implementation, based on simulated flight profiles and fuel estimations with BADA models. A novel approach could be to replace them with data from actual flight profiles and the OpenAP emission model.

VII. Conclusion

In this paper, we analyzed several years of real-world and high-resolution aircraft trajectory data collected by the OpenSky Network. Using the openly accessible flight emission models of OpenAP, we proposed this new approach to assess global aviation estimations based on this high-resolution data.

We concluded that reduced models are sufficient for large-scale aggregated emissions statistics, but precise models remain more accurate and provide better assessments to measure the impact of particular flight procedures.

By comparing the estimations produced by the reduced-order model, FEAT, with OAG data, we find large discrepancies in fuel consumption and CO_2emissions due to the lack of coverage from OpenSky. Such lack of coverage may account for approximately 45% of missing emissions using only the OpenSky data.

Despite some natural limitations, our analysis showed again the power and utility of large-scale crowdsourced aviation data to assess the air traffic environmental impact around the world. We believe our work can help policymakers, who can build on our model and data to make informed decisions impacting aviation and its emissions in the future.

References

[1] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. 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)*, Apr. 2014.

[2] M. Schafer, X. Olive, M. Strohmeier, M. Smith, I. Martinovic, and V. Lenders, "OpenSky Report 2019: Analysing TCAS in the Real World using Big Data," in *Proceedings of the 38th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, Sep. 2019, p. 9.

[3] X. Olive, A. Tanner, M. Strohmeier, M. Schafer, M. Feridun, A. Tart, I. Martinovic, and V. Lenders, "OpenSky Report 2020: Analysing in-flight emergencies using big data," in *Proceedings of the 39th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, 2020, p. 10.

[4] J. Sun, X. Olive, M. Strohmeier, M. Schafer, I. Martinovic, and V. Lenders, "OpenSky Report 2021: Insights on ADS-B Mandate and Fleet Deployment in Times of Crisis," in *Proceedings of the 40th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, 2021, p. 10.

[5] J. Sun, J. Hoekstra, and J. Ellerbroek, "OpenAP: An open-source aircraft performance model for air transportation studies and simulations," *Aerospace*, vol. 7, no. 8, p. 104, 2020. [Online]. Available: https://doi.org/10.3390/aerospace7080104

[6] K. Seymour, M. Held, G. Georges, and K. Boulouchos, "Fuel Estimation in Air Transportation: Modeling global fuel consumption for commercial aviation," *Transportation Research Part D: Transport and Environment*, vol. 88, p. 102528, Nov. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S136192092030715X

[7] OAG Aviation Worldwide Limited 2019. Historical schedules data for 2012 and 2018. [Online]. Available: https://www.oag.com/airline-schedules-data

[8] M. Strohmeier, D. Moser, M. Schafer, V. Lenders, and I. Martinovic, "On the applicability of satellite-based air traffic control communication for security," *IEEE Communications Magazine*, vol. 57, no. 9, pp. 79–85, 2019.

[9] M. Strohmeier, "Research usage and social impact of crowdsourced air traffic data," *The 8th OpenSky Symposium*, vol. 59, no. 1, p. 1, 2020.

[10] X. Olive, "traffic, a toolbox for processing and analysing air traffic data," *Journal of Open Source Software*, vol. 4, no. 39, p. 1518, Jul. 2019. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.01518

[11] Delft High Performance Computing Centre (DHPC), "DelftBlue Supercomputer (Phase 1)," https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase1, 2022.

[12] R. Alligier and D. Gianazza, "Learning aircraft operational factors to improve aircraft climb prediction: A large scale multi-airport study," *Transportation research part C: emerging technologies*, vol. 96, pp. 72–95, 2018.

[13] J. Sun, H. A. Blom, J. Ellerbroek, and J. M. Hoekstra, "Particle filter for aircraft mass estimation and uncertainty modeling," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 145–162, 2019.

[14] ICAO, "Aircraft engine emissions databank," *International Civil Aviation Organization*, 2020.

[15] D. DuBois and G. C. Paynter, ""fuel flow method2" for estimating aircraft emissions," *SAE Transactions*, pp. 1–14, 2006.

[16] A. Nuic, D. Poles, and V. Mouillet, "BADA: An advanced aircraft performance model for present and future ATM systems," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 10, pp. 850–866, Aug. 2010. [Online]. Available: http://doi.wiley.com/10.1002/acs.1176

[17] H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut. (2018) ERA5 hourly data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). [Online]. Available: https://dx.doi.org/10.24381/cds.bd0915c6

[18] F. Quadros, M. Snellen, J. Sun, and I. Dedoussi, "Global civil aviation emissions estimates for 2017–2020 using ads-b data," *Journal of Aircraft*, 2022. [Online]. Available: https://doi.org/10.2514/1.C036763