PROCEEDINGS | The 12th OpenSky Symposium

Automating the Estimation of Noise and Emissions Near Airports With ADS-B Data

Gonçalo Soares Roque^{*} and Johannes Reichmuth

Institute of Transport Science, RWTH Aachen University, Aachen, Germany *Corresponding author: goncaloroque30@gmail.com

(Received: 1 Nov 2024; Revised: 19 Jan 2025, 2 Apr 2025; Accepted: 7 Apr 2025; Published: 24 Apr 2025) (Editor: Manuel Waltert; Reviewers: Anandini Jayanthi, Céline Parzani, Saskia Drossaart van Dusseldorp)

Abstract

Aircraft arrivals and departures significantly affect nearby populations, primarily through noise pollution and the release of pollutants that degrade air quality. Estimating these environmental impacts can be a lengthy process and is typically mandated by legal regulations governing airport operations. This paper proposes a methodology to automate the estimation of environmental impacts for historical scenarios, specifically noise and pollutant emissions in the vicinity of airports, by utilizing open-source data. The automation pipeline developed retrieves the necessary databases and ADS-B data for a specified airport and time frame, and validates, pre-processes and enhances the data before estimating noise and local air quality emissions with it. The developed automation pipeline is applied to the Cologne Bonn Airport for the year of 2019. In addition to the open-source data, confidential datasets were made available containing the airport flight logs and the records from the airport noise measurement stations. This confidential dataset is used to assess the coverage of the ADS-B data and to validate the noise estimates generated with the automated process. The number of flights obtained from the ADS-B network covers ca. 82% of the flights in the airport flight logs, and the mean noise levels derived from ADS-B data deviate between 0 and 3 dB(A) from the ones recorded by the noise measurement stations, depending on the flight type and location of the noise stations. Possible reasonings for the different discrepancies observed include the assumptions made in the ADS-B data enhancement, as well as the underlying noise model and databases used. As a final step in the Cologne Bonn Airport use case, aircraft emissions reported according to the Landing & Takeoff cycle are compared with emissions estimates derived from ADS-B data. Significant discrepancies are observed between the two estimation methods which can be attributed to variations in time spent below 3000 ft AGL, average fuel flow and average EIs for each pollutant. This contribution provides an initial step toward automating the estimation of environmental impacts from arriving and departing aircraft. Further work shall focus on addressing the limitations of the methodology used to enhance the ADS-B tracks obtained and further validation of the environmental impacts estimated.

Keywords: Open-data; Noise; Emissions; Automation

Abbreviations: ANP: Aircraft Noise and Performance, EI: Emission Index, LTO: Landing and Takeoff, ICAO: International Civil Aviation Organization

1. Introduction

The direct environmental impacts of aircraft operations on the population, such as noise and air pollution, are highly concentrated within a small radius around airports. The estimation and reporting of these impacts to the public is currently driven by the legal frameworks airports must adhere

 $[\]otimes$ 2024 by the authors. This is an Open Access article, distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/)

to, such as the Environmental Noise Directive (END) [1] for noise and the EU Directive 2008/50/EC [2] regarding the assessment and management of pollutant concentrations. The studies conducted within this legal framework take place at irregular intervals, involve extensive data collection and preprocessing steps and are usually performed separately for each environmental impact.

The use of detailed trajectory data to estimate airport environmental impacts, particularly aircraft noise, has been an area of research even since before the widespread availability of flight tracking data, despite not being a standard practice within the regulatory frameworks followed by airports [3] [4]. Since crowdsourced flight tracking data, such as the Opensky Network [5], have become available, this topic has been further researched. For noise, Pretto et al. in [6], [7] have focused on reconstructing ground tracks and flight profiles with the models specified in Doc 29 [8] to adapt the trajectories produced to the ones observed in the data obtained from the Opensky Network. For aircraft emissions, the use of flight tracking data has seen more use with at-altitude emissions or complete trajectories from origin to destination airport, aiming at comparing different modelling approaches for the influence of flight characteristics on emissions [9] [10].

While significant progress has been made in leveraging detailed trajectory data for environmental impact analysis, developing noise and local air quality studies for each airport remains a complex and resource-intensive task. These studies are often hindered by inconsistent data availability, workload-intensive preprocessing requirements, and the need for tailored approaches for each airport and scenario. The lack of standardized methodologies further complicates efforts to compare results across different airports or studies, limiting transparency and the ability to derive broader insights.

This work addresses these challenges by proposing an automated and standardized process for estimating noise and emissions affecting local air quality in the immediate vicinity of airports, using open-source datasets. By automating the workflow and utilizing publicly available data, the proposed approach enhances reproducibility, transparency, and enables direct comparability of results across different airports. For a specified airport and time-frame, arrival and departure trajectories are obtained from ADS-B data, then filtered, processed, and enhanced before being associated with relevant noise and emissions databases. The airport environmental impact calculation engine GRAPE is subsequently used [11]. Noise is estimated with the noise model specified in Doc 29 [8] and local air quality emissions estimated by first calculating fuel flow as specified in ICAO Doc 9889 [12] and thereafter applying the Boeing Fuel Flow Method 2 (BFFM2) [13]. The pollutants in the scope of this work are hydrocarbonates (HC), carbon monoxide (CO), nitrogen oxides (NOx) and non-volatile Particulate Matter (nvPM), from here on referred to as local air quality pollutants. The scope of this work is restricted to emission inventories, without addressing the dispersion of pollutants through the atmosphere, which is heavily influenced by meteorological conditions.

The proposed automated process is demonstrated using a case study at Cologne Bonn Airport, where confidential datasets containing flight logs and noise measurement records were used for validation. This dataset is used to assess the coverage of the trajectory data obtained and validate the noise estimation performed.

This paper is organized as follows. Section 2 describes the data sources used, their access, and how they are associated. The specifics of the automated process built upon open source data are detailed in Section 3. Section 4 presents the application of this process to the Cologne Bonn Airport use case. Finally, Section 5 outlines the main conclusions and suggests directions for further research.

2. Data Sources and Association

2.1 Data Sources

To estimate the environmental impacts of each arrival and departure at a given airport within a specific time frame, the first challenge is obtaining complete trajectory data for each operation within a specified distance from the airport. For this, we rely on ADS-B data provided by the Opensky Network. The task of identifying the departure or arrival airport for a given trajectory has been addressed in previous research [14], and the solution has been integrated into the available API. This allows direct access to arrival and departure trajectories. Note that a trajectory is defined as a series of ADS-B data points associated with the same transponder code and callsign, with no more than a 10-minute gap between consecutive points. Each trajectory obtained must be associated with the databases used to estimate noise and emissions. To do so, the aircraft database available from the Opensky Network is used. The airport identification, coordinates and runways are obtained from airport data in the public domain. Airport METAR reports are used to obtain weather data for each ADS-B trajectory point and obtained from the archive maintained by the Iowa Environmental Mesonet (IEM), available in the public domain. Finally, the noise and emissions specific datasets are obtained from EASA and are accessible without restrictions. The Aircraft Noise and Performance (ANP) database and the respective substitution table are used to support the Doc 29 noise model and the Engine Emissions Databank (EEDB) in the estimation of emissions. This emissions database is complemented with a database containing data for turboprop engines from the Swedish Defence Research Agency (FOI). Table 1 summarizes the data used in the automated process and where it is sourced from.

Data	jource Remarks	
ADS-B data	Opensky Network	Accessible after obtaining credentials.
Aircraft data	Opensky Network	Accessible after obtaining credentials.
Airports data	https://ourairports.com	Airport data in the public domain.
METAR reports	IEM	Airport weather observations in the public domain.
ANP Database	EASA	Published online, accessible without restrictions.
ANP Substitution Table	EASA	Published online, accessible without restrictions.
EEDB	EASA	Published online, accessible without restrictions.
FOI Database	FOI	Available upon request.

2.2 Aircraft Abstraction

Each trajectory obtained from the Opensky Network is associated with an individual physical aircraft, of a certain model and with a specific engine type. However, the noise and emissions databases used do not contain data for each specific physical aircraft. Instead, each group aircraft according to specific criteria, mostly justified by the context in which they are developed. The ANP database provides data for its self-defined aircraft types, but is accompanied by a substitution table which offers two groupings:

- aircraft grouped according to ICAO code, engine type, maximum takeoff weight (MTOW) and maximum landing weight (MLW).
- · simplified grouping according to just aircraft ICAO code and engine type.

The emissions databases are developed in conjunction with the aircraft engine certification process

and therefore are split into each specific engine. The aircraft data available from the Opensky Network, which contains data for each individual physical aircraft, contains among others the aircraft ICAO code and a description of the aircraft type. Based on this data availability, the ANP substitution table grouping according to aircraft ICAO code and engine type is used as the overarching aircraft abstraction in this work. Aircraft with ICAO codes not found in this dataset are not considered. For each trajectory, the ICAO code of the aircraft is used to retrieve the respective ANP data entry. For the cases in which there is a differentiation in the ANP substitution table by engine type, the entry with the most similar engine description to the one obtained from the Opensky Network data is used. Unfortunately, the association with the emissions databases cannot be automated, as the engine descriptions available in these databases are significantly different than the ones obtained from the Opensky Network. For this reason, for each aircraft and engine combination defined in the ANP grouping, a suitable engine ID was manually selected. The principal selection criterion is the commonality of the engine for a given aircraft ICAO code and variant. For the few cases for which more than one unique engine in the emissions databases fit the criteria, the engine with the highest rated thrust is used.

3. Automated Environmental Impact Estimation

The automated process developed in this work involves all the steps required for the estimation of environmental impacts. A simplified view of the steps involved is presented in Figure 1. The automation is implemented in python in two separate libraries. The first is a fork of *traffic* [15], an open source library that provides functionality for working with and processing air traffic data, including the Opensky Network. The fork adds functionality to *traffic* required in the processing steps described below, such as accessing the ANP and IEM databases and the computation of calibrated airspeed and acceleration for each trajectory point. The second library was developed specifically for this automated process, and connects the air traffic data obtained to the functionality provided by GRAPE. The behaviour of the process can be controlled via configuration files, e.g. to specify which versions of the environmental impact databases to use or to override access to other datasets. The only mandatory configurations are the access credentials to the Opensky Network historical data and the file path to the GRAPE executable.



Figure 1. Data flow and automation steps

3.1 Processing Steps

The trajectory processing steps are performed on a per operation basis and can be summarized into the following categories:

• Filtering stage: perform both data integrity checks on each operation as well as outlier detection on trajectory points.

- Trajectory modification: enforce homogeneity in the trajectories obtained and discard parts of the trajectory not in scope.
- Feature enhancement: add features to each trajectory point to enable the estimation of noise and local air quality emissions.

These steps are described in detail below.

Filtering Stage

The automated process starts by accessing trajectory data for all arrivals and departures for the selected airport within the selected time frame, limited by a configurable bounding box. Initial processing discards invalid data as well as outliers in positional and kinematic variables with a median filter pass (e.g. latitude, longitude or ground speed). Thereafter, the following integrity checks are performed:

- runway: a departure or arrival runway must be found for departure and arrival operations respectively. An operation is associated with a runway if it is aligned with it for more than one minute.
- ANP aircraft: the ICAO code of the aircraft must be found in the ANP substitution table. The corresponding ANP aircraft thrust type must be given in either pounds of force or percentage of maximum static thrust, as other thrust types are not yet supported by GRAPE.
- go around and runway change (arrivals): based on the same approach as for the runway integrity check, arrivals which perform a go around are discarded. Furthermore, if the arrival operation changes runway on final approach, it is also discarded (only applicable to airports with parallel runways). This filtering step was introduced to avoid non-standard aircraft states in the data, which would most likely provoke outliers in the environmental impacts estimated.
- total ground distance: at least 10 nautical miles must be traversed. This value is chosen rather arbitrarily and can be adjusted in order to account for specificities at different airports.

Trajectory Modification

Following filtering, a smoothing step is performed in order to avoid strong variations that would not occur in reality. A rolling mean over a five second window is performed on altitude, ground speed and climb rate. This is followed by an enforcement of commonality between all operations. While the filtering step ensures the data has a certain quality, for example at least 10 nautical miles traversed, there is no matching in the data at a certain geographical location or feature value. However, this commonality is required especially to compare different approaches in the estimation of local air quality emissions. To achieve this, points which are after the last point aligned with the runway for arrivals (i.e. after the aircraft has crossed the landing threshold), or before the first point aligned with a runway for departures are discarded. Thereafter, for arrivals, a point at the landing threshold is added to the end of the trajectory. The last minute aligned on the runway is used to estimate the values at the threshold. Ground speed and climb rate are set as the mean value, and altitude estimated through linear regression. For departures, the departure threshold is added to the beginning of each departure. Both ground speed and climb rate are set to zero. The trajectory processing is finalised with a resampling step that interpolates points at every second.

Feature Enhancement

The final processing step is to enrich each trajectory point with the features required to estimate noise and local air quality emissions which are not available in the ADS-B data. For the Doc 29 noise model, these features are bank angle, true airspeed, corrected net thrust per engine, and an identification of the points belonging to the takeoff roll and the landing roll. For this automated

process, bank angle is not considered and set to zero for all trajectory points. The true airspeed is estimated based on ground speed and the wind vector obtained from the METAR reports. For each operation, the respective airport METAR report which is closest in time to each trajectory point is used, both for true airspeed as well as for all other variables where weather data is required. A significant hurdle is the estimation of net thrust for every trajectory point. The existing models to estimate net thrust can be split into physics models and regression models. For physics models, the thrust is estimated through either the force balance or conservation of energy equations. In order for them to be applied, the other forces actuating on the aircraft, namely weight, lift and drag also need to be estimated. Regression models on the other hand apply empirical correlations for each aircraft between net thrust and the variables on which it depends, such as altitude, speed and atmospheric conditions. The downside of regression models is that they provide regression coefficients for a specific aircraft state (high lift devices configuration and thrust lever setting), when in reality different configurations are used (e.g. the use of thrust reduction procedures). The choice of thrust model for this automated process needs to guarantee its applicability to all aircraft considered, namely the ones available in the ANP substitution table. For this reason, the approach used by this automated process follows the performance model within Doc 29, which uses a mixture of both model types described above.

For arrivals, corrected net thrust per engine is calculated with the following force balance equation:

$$\frac{F_n}{\delta} = \frac{W}{N \times \delta} \times (R \times \cos \gamma + \sin \gamma + a/g)$$
(1)

Acceleration *a* and descent angle γ (negative by convention) can be directly calculated from the ADS-B data. The pressure ration δ is obtained with the ISA atmospheric model and the closest METAR report as described above. To estimate weight *W* and drag to lift coefficient *R*, certain assumptions are required. In this automated process, the weight is set to 90% of the MLW as provided by the ANP database, aligning with the Doc 29 methodology for default arrival profiles [8]. The R coefficient is also obtained from the ANP database where different values are specified for different flap settings. The flap setting for each trajectory point is determined using a fixed flap deployment schedule specific to each ANP aircraft, derived from the default arrival profiles in the ANP database. Since the flap schedule is defined based on calibrated airspeed, this parameter is estimated from true airspeed, accounting for both air density and compressibility effects.

For departure operations, net thrust calculations are split into jet engine and turboprop engine powered aircraft. Both are based on regression models, derived separately for the two common thrust ratings used during a departure operation, maximum takeoff followed by maximum climb. The following formulas are used:

$$\frac{F_{n,jet}}{\delta} = E + F \times V_c + G_A \times h + G_B \times h^2 + H \times T$$
(2)

$$\frac{F_{n,prop}}{\delta} = \frac{\eta \times P_P}{\delta \times V_T} \tag{3}$$

Calibrated airspeed V_c , true airspeed V_T , temperature T and pressure ratio γ are all calculated as described above for arrival thrust. The regression coefficients E, F, G_A , G_B and H, propeller efficiency η and propulsive power P_P are aircraft and thrust setting specific and obtained from the ANP database. They are available for two thrust settings, maximum takeoff and maximum climb. This automated process does not consider the use of reduced thrust procedures. The thrust cutback point (change from maximum takeoff to maximum climb thrust setting) is estimated by finding the first local maximum of the variation of climb rate with time for points between 500 and 3500 ft AGL. This method assumes that thrust reduction takes place between this limiting altitudes, which is in par with the ICAO noise abatement departure procedures recommendations.

Finally, the Doc 29 noise model requires an identification of points belonging to the takeoff and landing roll in order to apply corrections to the noise estimation specific to these phases of flight. As described above, arrival trajectories are clipped at the landing threshold, for which the landing roll is not considered. We further simplify the noise estimation in this work by not attributing the takeoff roll phase to any trajectory point, effectively disregarding the noise corrections specific to this flight phase.

The trajectory enhancements applied in order to use the Doc 29 noise model cover most of the requirements to use the ICAO Doc 9889 fuel flow model and the BFFM2 emissions model, namely true airspeed and thrust estimation. The only missing information required for each trajectory point is the Landing and Takeoff (LTO) phase as defined in the LTO cycle. For arrivals, all points are attributed the *approach* phase as the trajectory is clipped at the landing threshold. For departures, the thrust cutback point obtained in the thrust estimation process above is used to split the trajectory into the *takeoff* and *climb out* phases.

3.2 Environmental Impact Calculation

The environmental impacts are estimated with GRAPE, also developed by the authors. GRAPE is open-source and provides an implementation of the required noise, fuel flow and emissions models, with configuration parameters which control the exact behaviour of the implemented models. Furthermore, it provides full automation capabilities and is database independent, as it requires the user to provide all the necessary data to perform the calculations. These characteristics are extensively used by the automated process developed in this work. First, the command line tool provided by GRAPE is used to create an empty study and import the ANP database selected by the user (in this study the ANP version 2.3 is used), which transforms it to the GRAPE internal format for Doc 29 data. Note that a GRAPE study is simply a *sqlite* database which follows a predefined schema. This database can be further edited to automatically import data into the study. This functionality is used to import all further input data, namely the EEDB and FOI databases, the aircraft abstraction defined and the trajectories obtained from the ADS-B data. Before importing, the ICAO Doc 9889 approach to deal with missing smoke number (SN) values is implemented and applied for values obtained from the EEDB.

The next step is to define the calculation runs and their parameters. GRAPE uses a parent/child structure, where a performance run is the parent of noise and local air quality emissions runs. The parameters of the parent performance run (e.g. weather data, coordinate system) are used by all children runs. For this automated process, a performance run is created which uses the WGS 84 coordinate system and the automatically obtained METAR reports containing temperature, pressure, wind conditions and relative humidity, typically at 30 minutes intervals. The METAR report closest in time to each operation is used where weather data is required. As trajectory data is already available as described above, the only further action performed by the performance run is to estimate fuel flow. This is estimated for each trajectory point according to the ICAO Doc 9889 model which interpolates from the LTO fuel flow values based on aircraft thrust. The option to correct fuel flow at MSL to the conditions observed at the aircraft with the BFFM2 is used. A noise run is created which uses the Doc 29 noise model and the atmospheric absorption defined in the SAE-ARP-5534. The list of receptors for which to estimate noise is obtained from the automated process configuration file. Finalising, two different emissions runs are defined. Both runs calculate gas and particle emissions, and use the FOA 4 method to estimate nvPM emission indices (EIs) in case they are not available in the EEDB [16]. The first run uses the segments obtained with the ADS-B trajectories, and the BFFM2 method to obtain gas pollutant EIs based on fuel flow and correct to the atmospheric conditions observed at the aircraft. The second emissions run serves as a baseline and calculates emissions according to the LTO cycle, disregarding any trajectory data or weather conditions. As the trajectories are clipped at the runway thresholds, the time attributed to the taxi/idle mode is set to zero seconds.

After inserting all the input data and defining all the necessary parameters in the GRAPE study, the command line tool capabilities of GRAPE are once again used to automatically launch the program, open the defined study, and run the noise and emissions calculation runs defined. Upon finish, the results are directly available in the same *sqlite* file for further analysis.

3.3 Limitations

The main limitations of this automated process to report airport environmental impacts lie in the assumptions and approximations made to enhance the trajectory data obtained from ADS-B. First, relying on METAR data for atmospheric parameters introduces inaccuracies in the thrust estimation, as well as in the methods used to estimate noise, fuel flow, and local air quality emissions. Utilising atmospheric models with high-resolution weather data would enhance accuracy and is a crucial step toward integrating emissions dispersion into the automated process. Furthermore, estimating net thrust according to the procedure described above makes significant assumptions. For arrival thrust estimation with Equation 1, the aircraft weight is set to 90% of MLW and the flap retraction schedule is static. These variables are not readily available, and estimating them from ADS-B data is only possible to a certain degree. Recent studies have focused on training machine learning models with data sets where weight and flap settings are known, to then predict these variables based solely on trajectory data [17] [18]. However, such models are not yet available for a significant amount of aircraft. For departures, using the regression models described assumes that full takeoff thrust is used for every flight. In reality, thrust reduction is a common procedure used to improve the life cycle of engines and reduce maintenance costs. The thrust reduction percentage used for each departure is dependent on a multitude of variables such as weight, runway length, inclination and condition as well as weather. The availability of runway and weather information for this automated process covers some of the data requirements to estimate the thrust reduction parameter for each departure. However, one major influencing parameter, takeoff weight, is not readily available. Similar to the missing parameters for arrivals, recent studies have focused on developing predicting algorithms with machine learning techniques [18] [19] [20]. The unavailability of such models for a significant number of aircraft also applies for departures. When such models become available, the automated process defined in this work can be improved to account for those variations.

A further improvement to the automated process described lies in improving the modelling of the landing and takeoff rolls. The approach used may be enhanced by using landing or takeoff roll data from ADS-B data if available. As commonality between all flights must be guaranteed, this approach must be complemented with an estimation method for the takeoff and landing roll (e.g. the performance model within Doc 29) for missing data. Finally, the approach described in Section 2 introduces limitations in the accuracy of environmental impact estimations. Currently, each flight is associated with noise and emissions databases based solely on the aircraft ICAO code and, for a limited number of aircraft, the engine type. This simplification does not fully capture the actual fleet mix operating at a given airport. Given sufficient data availability, the automated process could be enhanced by incorporating a more granular aircraft classification, enabling a more precise linkage between flight trajectories and environmental impact databases.

4. Use Case: Cologne Bonn Airport

The automated process described in Section 3 is applied to the Cologne Bonn Airport (airport ICAO code EDDK) for the year of 2019, the last year of operations without the influence of the COVID-19 pandemic. This use case is a demonstration of different analysis that the automated process enables as well as its capabilities.

4.1 Metrics and Validation Dataset

The automated process is evaluated in this use case with a confidential historical dataset obtained from the Cologne Bonn Airport. The airport provided the flight logs as well as the records from the noise measurement system for the year of 2019. This allows to evaluate the coverage of the ADS-B data retrieved from the Opensky Network as well as to compare the noise estimates obtained with the automated process against the records from the noise measurement system. In this analysis, the values recorded at the noise stations are viewed as absolute truth, and the objective of the noise estimation with the ADS-B data is to be as close to the measured values as possible. Noise can be estimated by the automated process at the locations of the noise measurement stations, and directly compared to the recorded values. However, there is no direct association between each individual flight recorded by the airport, and the flights obtained from ADS-B data. Therefore, aggregated metrics and the noise level distributions are analysed. For emissions, per flight data is impracticable to record and not available. The legal framework airports follows usually mandates the report of emissions according to the LTO cycle, a static method which is independent of the actual aircraft trajectory, weather conditions or any other variable. We evaluate the emissions estimated with ADS-B data against this reporting method, recognizing that none of the two methods can be considered as absolute truth. Nonetheless, we expect the results obtained with flight specific trajectories obtained from ADS-B data to be closer to the truth than the static values found in the LTO cycle.

Note: from here on, the automated process data and its results are referred to as the ADS-B data, and the Cologne Bonn Airport data as the validation dataset.

The flight logs provided contain a list of all arrivals and departures which occurred at the airport in 2019 for aircraft with MTOW above 10 000 kg. There were a total of 134 788 flights, as expected evenly distributed between arrivals and departures. Flights with aircraft not considered in the aircraft abstraction defined for the automated process (i.e. for which entries in noise and emissions databases are not available) were discarded (approximately 1% of all flights). The fleet mix operating at Cologne Bonn Airport is relatively homogeneous, dominated by narrow body aircraft. The A320 and B737 aircraft families account for approximately 72% of all flights. In terms of runway distribution, departure operations overwhelmingly use the main runways 13L/31R, which account for more than 95% of all operations. For arrivals, the share of flights which take place in the main runways is slightly lower, as the crosswind runway 24 accounts for approximately 19% of all arrivals.

The noise station records were obtained for all 17 fixed noise stations around the airport. In total, there were 137 450 noise events for arrivals, and 191 096 for departures, corresponding to approximately two noise events per arrival and three per departure. Figure 2 depicts the placement of the noise stations around the airport. The flight track heatmap displayed in Figure 2 was obtained with a sample of 5000 flights from the Opensky Network.

The number of noise events per flight type and per noise station is not homogeneous across noise stations. In order to focus the analysis and validation on the most significant noise stations (i.e. the ones with most events), an arbitrary threshold was set to 5% of total arrivals or departures, for the station to be considered for the respective flight type. After applying this criteria, the stations M01, M02, M05, M06, M07 and M08 are considered for arrivals, and M06, M08, M11, M14, M17 and M18 for departures. By definition, the noise stations register a noise event only when a certain



Figure 2. Layout of noise measurement stations around the airport

maximum level is exceeded. This threshold level was extracted from the noise events by analysing noise maximum value (LAMAX) distribution at each station. The LAMAX distributions of each noise station for each hour of the day were analysed, as available airport information suggested that different thresholds may be in use for different times of day. The analysis revealed that for most noise stations, two different thresholds are defined. During the day, between 06:00 and 22:00, the LAMAX minimum value observed is 65 dB. At night, the LAMAX minimum value is 63 dB. Two noise stations, M01 and M05, have 65 dB as minimum LAMAX throughout day and night.

4.2 Flight Results

The automated process identified 108 715 operations from the Opensky Network associated with the Cologne Bonn Airport for the year of 2019, also equally split between arrivals and departures. This corresponds to approximately 81.5% of the number of flights in the validation dataset. The missing flights are split evenly across operation type, runway, time and aircraft type. The discrepancy between the number of flights in the validation dataset and the ADS-B data lies therefore most likely in the coverage of the Opensky Network, specifically at low altitudes, as this is required for the identification of an airport (in our case the Cologne Bonn Airport) as the arrival or departure airport.

After applying the flight filtering steps described in Section 3, 16 712 flights were removed. There were 7880 flights which trajectory did not align with any runway, 6856 flights with an aircraft ICAO code not found in the ANP substitution table and 3468 flights which total ground distance covered was less than 10 nautical miles. Two of these criteria, runway association and minimum cumulative ground distance, are directly associated with the quality and completeness of the ADS-B data. An increase in the number of ADS-B receivers in the Opensky Network around the airport will most likely increase the quality of the data around the airport and reduce the number of operations not meeting these integrity criteria. The significant number of flights removed due to a missing

ANP substitution table entry is mostly likely due to inaccuracies or incompletenesses in the aircraft database maintained by the Opensky Network, and may be fixed by correcting them. Note that in the validation data, only 1235 flights out of 134 788 could not be associated with a valid ANP substitution aircraft. The number of arrivals which were identified as either performing a go around or a runway change on final approach is relatively low in comparison to the other three criteria (304 and 440 flights respectively).

4.3 Noise Results

After pushing the processed data through GRAPE as described in Section 3, the LAMAX and Sound Equivalent Level (SEL) single event noise metrics are available at each receptor for each flight obtained with ADS-B data and can be compared against the validation dataset. From the relevant noise stations selected above, the results for the noise stations M02 for arrivals and M11 for departures are widely different from the results for all other noise stations. These significant discrepancies are most likely due to a data error regarding the location of the noise stations, and their results are therefore not further discussed. The comparison between number of noise events for each station between ADS-B data and validation dataset is provided in Table 2. For the ADS-B data, it is simulated that each flight produces a noise event at each noise station if the LAMAX value is equal or higher to the noise station threshold (i.e. the minimum LAMAX value recorded at the noise station). The different thresholds for different times of day obtained as described above are considered. For noise stations where the noise event distribution range includes the threshold, this approach could distort the results as under-predicted LAMAX values below the threshold would be excluded. A correlation between flight tracks (including timestamps) and measured noise events would facilitate a direct, one-to-one comparison between the ADS-B data and the validation dataset, eliminating the need for defining thresholds. In comparison to the flight coverage rate, the coverage rate for noise events is even smaller for arrival operations, and approximately the same for departure operations. The lowest coverage rate occurs for arrival operations at station M06, likely due to the low elevation angle between the standard arrival glide path and this noise station. Notably, despite being closer to the runway, station M06 recorded only about 77% of the arrival noise events captured by noise station M08.

	Noise Station	Validation Data	ADS-B Data	Percentage
Arrivals	M01	28034	21069	75%
	M05	11870	6765	57%
	M06	18779	9019	48%
	M07	24244	14327	59%
	M08	24251	17573	72%
Departures	M06	35703	24875	70%
	M08	34856	24891	71%
	M14	15216	11989	79%
	M17	21744	15969	73%
	M18	12115	7916	65%

Table 2. Noise Event Count

The SEL histogram comparison for arrival flights for two selected noise stations, M05 and M08, is presented in Figure 3. The noise stations are selected in order to demonstrate two different trends, which are differentiated by the proximity of the noise station to the airport. The two noise stations closest to the respective landing threshold, M05 and M06, show practically equivalent noise level distributions for validation and ADS-B data. At this proximity to the airport, aircraft are generally

already fully aligned with the runway and on the correct glide path. A certain amount of thrust is required to maintain this alignment and fly at or close to the landing speed. However, at the noise stations M01, M07 and M08, which are further away from the respective landing threshold, the ADS-B data distribution is approximately 2 to 3 dB(A) lower than the validation data. These discrepancies may be attributed to either the assumptions made in estimating the thrust parameter for arrivals (90% of MLW and fixed flap deployment schedule), the simplified thrust estimation method itself, or due to the noise modelling methodology specified in Doc 29. For the latter in particular, the Noise-Power-Distance (NPD) tables on which it is based determine noise primarily based on thrust setting and distance between aircraft and receptor. While aerodynamic noise is indirectly considered when developing the NPD tables, this is based on the final stages of the arrival operation, where noise levels are relatively high and aircraft are usually already fully configured for landing. Regarding the LAMAX metric, the same overall trends are observed for all noise stations and not further discussed.



Figure 3. Arrival flights SEL comparison

Figure 4 presents the SEL histogram comparison for departures also for two selected noise stations, M08 and M18. The noise level distributions in the ADS-B data exhibit two distinct curves, one at lower noise values and another at higher noise values, unlike the validation dataset, which follows a single Gaussian distribution. Since the distributions are presented per noise station, they incorporate noise values from a diverse range of aircraft types. The likely explanation for the discrepancy between the ADS-B and validation datasets is the inherently lower standard deviation in estimated noise, which does not fully account for the complex variables present in real-world conditions. The homogeneity of the aircraft fleet at Cologne Bonn Airport further reinforces this conclusion, as for all noise stations the first curve is primarily shaped by the most frequently observed narrow-body aircraft, while the second curve is driven by the most common wide-body aircraft. The same differentiation as for arrival flights, regarding the proximity of the stations to the airport, is observed for departures. The results for noise station M08 are equivalent to noise station M06, both stations being relatively close to the airport. For noise station M08, the ADS-B data tends to have a higher amount of noise events in the higher values area, approximately above the 85 dB(A) threshold. An explanation for this could be the assumption in the automated process that every departure always uses the maximum thrust available. Considering the use of reduced takeoff thrust would result in lower estimated noise values (note that for the automated process, thrust reduction has no impact in the climb profile, as the ADS-B data is used). For noise stations M14, M17, and M18, the discrepancy is more accentuated. The ADS-B data produces approximately 2 to 3 dB(A) higher values than the validation dataset. As shown in Figure 5 for noise station M18, the same effect is not observed for the LAMAX metric. The reasoning behind the discrepancy for the SEL distribution is unlikely to be thrust reduction, as it can be reasonably expected that its effect on noise level will be higher at locations closer to the airport. Further away from the airport, at the locations of M14, M17 and M18, aircraft have in general already performed both thrust cutback and flap retraction, for which the assumption of using the full thrust available should be less significant than at lower altitudes. The most likely reasoning for the discrepancy in SEL distributions for noise stations farther away from the airport lies in the noise modelling, specifically in the quality of the SEL NPD tables found in the ANP database for the most frequent aircraft operating at Cologne Bonn Airport. This possibility is further corroborated by the findings by Giladi et al. in [21], and would also explain why the difference occurs only for the SEL metric and not for LAMAX.



Figure 5. Departure flights LAMAX comparison

Finally, the year noise equivalent continuous sound levels L_{eq} are provided in Table 3. Noise levels are shown for arrivals and departures for the respective relevant noise stations, as well as for all flights for the two noise stations which are relevant for both arrivals and departures. The interpretation of the results needs to take into account the discrepancy in number of noise events between validation dataset and ADS-B data. Both this discrepancy and the discrepancies observed for arrival and departure noise as described above have an influence in the cumulative noise metric. Despite them, the automated process is able to estimate year L_{eq} with an accuracy between 0 and 1 dB(A) for departure operations. This is in part due to the assumption of always using maximum available thrust for all departures, which counteracts the lower number of noise events for the ADS-B data. For arrivals the results are less promising. The year L_{eq} discrepancy between validation and ADS-B data ranges between approximately 2 and 4 dB(A), apart from the noise station M06. Due to the lower number of noise events for ADS-B data and the short comings of arrival noise modelling as described above, the cumulative values tend to be lower for the ADS-B data.

Overall, the results obtained are in agreement with the ones obtained by previous studies. Pretto et al. in [6] report an underestimation of 0 to 5 dB(A) in their trajectory reconstruction approach when compared to noise measurements. This comparison was made for the average daily equivalent sound levels across 23 selected days at Zurich Airport. In [3], Strümpfel et. al report an underestimation of less than 1 dB(A) for 13 departure flights at Berlin Airport. Their trajectory reconstruction approach shares similarities with the method used in this study, as both retain the trajectory points obtained from ADS-B data (and radar data in the Strümpfel et. al) and only thrust is estimated (Strümpfel et. al estimate it via a physics model for conservation of energy).

	Noise Station	Validation Data	ADS-B Data	Difference
Arrivals	M01	52.8	50.0	-2.8
	M05	55.2	52.9	-2.3
	M06	46.6	47.4	0.8
	M07	50.2	46.4	-3.8
	M08	55.0	52.5	-2.5
Departures	M06	54.5	53.9	-0.6
	M08	54.1	53.2	-0.9
	M14	46.3	46.6	0.3
	M17	48.9	48.6	-0.3
	M18	45.0	45.0	0.0
All	M06	55.1	54.8	-0.3
	M08	57.6	55.8	-1.8

Table 3. Year noise equivalent sound level L_{eq} comparison

4.4 Emissions Results

After finalizing the automated process, local air quality emission results are available for each segment of each flight trajectory. Additionally, the baseline LTO cycle values are also obtained for each flight, based solely on aircraft type. In par with the LTO cycle and the trajectory processing described in Section 3, only trajectory segments below 3000 ft AGL are considered. For the LTO cycle, arrival values are obtained solely from the *approach* mode and departure values are the sum of the *takeoff* and *climb out* modes. The analysis below is split by aircraft type and focuses on the 5 most frequent aircraft in the ADS-B data for the Cologne Bonn Airport.

A comparison of the time spent below 3000 ft AGL between ADS-B data and the LTO cycle is presented in Figures 6a and 6b for arrivals and departures respectively. The LTO cycle time is a constant value for each, independent of any other variable. For arrivals, the ADS-B data results align well with the LTO cycle value. Heavier aircraft have a higher approach speed and spend therefore on average less time below 3000 ft. The strong alignment with the LTO cycle observed at Cologne Bonn Airport may differ significantly at other airports where arrival procedures require aircraft to level off or maneuver at or below 3000 feet. For departures, the observed discrepancies are larger. Across all aircraft, the 50% confidence interval (representing the middle half of flights around the median) falls below the LTO cycle value of 2.9 minutes. The higher discrepancy observed may be partially attributed to operational conditions at Cologne Bonn Airport. However, the departure trajectory reconstruction approach outlined in Section 3 could also contribute to these discrepancies. Specifically, the practice of ensuring trajectory homogeneity by adding the departure threshold as the first point may result in the initial segment being modelled at a higher speed than what occurs in reality, depending on the location of the first recorded ADS-B trajectory point. Enhancing the automated process to either include the takeoff roll in the ADS-B data or estimate it more accurately would help eliminate this source of error.

While differences in time significantly contribute to discrepancies in emissions calculations between the LTO cycle approach and the use of ADS-B data, the developed automation also accounts for the impact of conditions at the aircraft (e.g. thrust, altitude, and temperature) on both the estimated fuel flow and the pollutants EIs for each trajectory segment. Figure 7 illustrates the impact of each discrepancy factor on the absolute differences in fuel and emission values calculated using ADS-B data compared to the LTO cycle. The results for HC and nvPM number are omitted, as they are equivalent to CO and nvPM mass respectively.



Figure 6. Flight time below 3000 ft. The box represents the interquartile range (50% confidence interval), and the median is shown as a line inside the box. The outside lines extend to 1.5 times the interquartile range and points outside this range are not shown.



Figure 7. Total fuel and pollutant emissions difference between ADS-B data and LTO cycle and respective causes. Data labels show the total percentual difference between ADS-B data and LTO cycle.

For arrivals, differences in time are the primary contributor to discrepancies across all aircraft and pollutants. Accounting for trajectory segment specific conditions, such as thrust and altitude, slightly

increases the estimated fuel consumption for all aircraft. The influence of trajectory segment conditions on the estimated EIs in comparison to the LTO cycle for arrivals is observable for CO and HC. Jet engine aircraft exhibit both higher EIs and steeper EI gradients at lower thrust settings for these two pollutants, making these results expected. The exact magnitude and direction (positive or negative) of the deviation from the LTO cycle EIs depends on the fuel flow estimated for each segment, which is influenced by the specific conditions of the trajectory segments. Accounting for trajectory segment specific conditions has a significantly greater impact on estimated fuel consumption and pollutant emissions for departures. For all three narrow-body aircraft, the influence of these conditions on fuel flow, and the resulting discrepancies compared to the LTO cycle values, is substantial, making it the primary cause of differences in total fuel consumption and pollutant emissions for all three aircraft. For the two B767 aircraft, where the time difference below 3000 ft between ADS-B data and the LTO cycle is higher (see Figure 6b), the influence of trajectory segment specific conditions on fuel flow is less pronounced. The impact of trajectory segment conditions on EIs is also stronger for departures, especially for the A319, B738 and B763 aircraft. The difference in pollutants emitted by these three aircraft between ADS-B data and LTO cycle has an observable component due to the differences in EIs. Although this component is smaller than the impact caused by difference in fuel flow, it is significant and equal to or higher than the differences caused by differences in time below 3000 ft.

Using ADS-B data to estimate time below 3000 ft, fuel consumption, and pollutant emissions reveals significant discrepancies compared to the LTO cycle approach. The absolute differences range from 2% to 14% for arrivals and from 7% to 36% for departures. Discrepancies in time below 3000 ft are the primary cause of the differences observed for arrivals. However, for departures, the influence of trajectory segment conditions on fuel flow and EIs also contributes significantly to the observed discrepancies. A key limitation of this analysis is the inability to assess how the use of detailed ADS-B data affects the accuracy of local air quality emissions estimates. While it is assumed that this approach yields more precise results than simpler methods based solely on the LTO cycle, this assumption remains unverified. Integrating emissions dispersion modelling into the automated process and comparing the results with measured or modelled pollutant concentrations at various locations would address this limitation. Such an approach would provide deeper insights into the advantages of using high-resolution trajectory data for assessing the environmental impact of aircraft operations near airports.

5. Conclusion and Further Work

The major contribution of this work lies in the provision of an open source framework to estimate noise and local air quality emissions for historical flights with ADS-B data, for any given airport and time frame. Besides trajectory data obtained from the Opensky Network, airport, METAR reports and environmental impact databases are automatically accessed in order to automate the estimation. Each ADS-B trajectory is filtered, smoothed, associated with each of the environmental impact databases and enhanced with additional features before estimating noise and emissions. The open source environmental impact calculation engine GRAPE is used, developed by the same authors. Noise is estimated with the Doc 29 noise model at any given number of user specified locations and emissions estimated for each flight segment by first estimating fuel flow with the ICAO Doc 9889 model and subsequently the BFFM2. We applied this automation to an year of operations at the Cologne Bonn Airport, for which an additional dataset was made available containing the flight logs and records from the noise stations provided mixed results, depending on type of operation and location of the noise station relative to the airport. For noise stations close to the airport, the noise level distributions estimated with the ADS-B data closely matched the ones obtained from the noise

measurement system. For the ones further away from the airport, arrival noise was under-predicted and departure noise over-predicted by 2 to 3 dB(A). The comparison of the fuel and emissions estimated with the ADS-B data against the LTO cycle values showed absolute percentual discrepancies in the low double digits (10 to 20%) for the majority of aircraft and quantities analysed (time, fuel and pollutants). While differences in time below 3000 ft AGL are the primary cause of discrepancies for arrivals, for departures, the influence of trajectory segment-specific conditions on fuel flow is the main driver of discrepancies when compared to the LTO cycle. Additionally, differences in estimated EIs were observed to also contribute to differences in total pollutants emitted, especially for departures and for CO and HC in the case of arrivals.

In future work, the trajectory processing within the automated system can be enhanced, as outlined in Section 3. Key areas for improvement include more accurate estimation of the thrust parameter, a more comprehensive approach to modelling the takeoff and landing roll phases of each trajectory and incorporating more detailed data, such as high resolution weather data. Furthermore, the automated process currently estimates noise and emissions using the most accurate trajectory data available, obtained from ADS-B. Future work could expand on this by incorporating different trajectory reconstruction methods, such as performance models outlined in Doc 29 or BADA, or alternative ground track reconstruction methods, into the automated process and analysing their impact on the resulting noise and emissions estimates. Finally, applying the use case discussed in Section 4 to other airports and time periods with available validation datasets could further substantiate the findings of this study regarding the estimation of environmental impacts using ADS-B data.

Author contributions

- Gonçalo Soares Roque: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization and Writing (Original Draft)
- Johannes Reichmuth: Supervision, Writing-Review and Editing

Open data statement

The main contribution of this paper is in the form of an environmental impact estimation tool available under https://github.com/goncaloroque30/GRAPE-traffic. As discussed in the paper, all data sources used by this tool are either in the public domain or openly accessible. The validation dataset used to evaluate the results obtained with the tool for the use case presented is not publicly available.

Reproducibility statement

The code developed within this work and all the analysis conducted in this paper are available under https://github.com/goncaloroque30/GRAPE-traffic.

References

- [1] European Union Parliament and Council. Directive 2002/49/EC of the European Parliament and of the Council Relating to the Assessment and Management of Environmental Noise. 2002.
- [2] European Union Parliament and Council. *Directive 2008/50/EC of the European Parliament and of the Council on Ambient Air Quality and Cleaner Air for Europe.* 2008.
- [3] Josephin Hübner and Christoph Struempfel. "Aircraft Noise Modeling of Departure Flights based on Flight Track Data and Actual Aircraft Performance Parameters". In: *Lärmbekämpfung* 15 (Jan. 2020), pp. 188–193. DOI: 10.37544/1863-4672-2020-06-18.

- [4] B Schäffer, S Plüss, G Thomann, R Bütikofer, and MM Boone. "Aircraft noise calculations for relevant periods of day using a complete set of radar data". In: Proceedings of the NAG/DAGA 2009, International Conference on Acoustics, including the 35th German Annual Conference on Acoustics (DAGA). 2009, pp. 1414–1417.
- [5] Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, and Matthias Wilhelm. "Bringing up OpenSky: A large-scale ADS-B sensor network for research". In: *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*. 2014, pp. 83–94. DOI: 10.1109/IPSN.2014.6846743.
- [6] Marco Pretto, Lorenzo Dorbolò, Pietro Giannattasio, and Alessandro Zanon. "Aircraft operation reconstruction and airport noise prediction from high-resolution flight tracking data". In: *Transportation Research Part D: Transport and Environment* 135 (2024), p. 104397. ISSN: 1361-9209. DOI: https://doi.org/10.1016/j.trd.2024.104397. URL: https://www.sciencedirect.com/ science/article/pii/S1361920924003547.
- [7] Marco Pretto, Pietro Giannattasio, Michele De Gennaro, Alessandro Zanon, and Helmut Kühnelt. "Web data for computing real-world noise from civil aviation". In: *Transportation Research Part D: Transport and Environment* 69 (2019), pp. 224–249. ISSN: 1361-9209. DOI: https: //doi.org/10.1016/j.trd.2019.01.022. URL: https://www.sciencedirect.com/science/article/pii/ S1361920918306254.
- [8] ECAC. ECAC.CEAC Doc 29 Report on Standard Method of Computing Noise Contours around Civil Airport. 2016.
- [9] Junzi Sun, Luis Basora, Xavier Olive, Martin Strohmeier, Matthias Schäfer, Ivan Martinovic, and Vincent Lenders. "OpenSky Report 2022: Evaluating Aviation Emissions Using Crowdsourced Open Flight Data". In: Sept. 2022. DOI: 10.1109/DASC55683.2022.9925852.
- [10] Antonio Filippone, Nicholas Bojdo, Shreya Mehta, and Ben Parkes. "Using the OpenSky ADS-B Data to Estimate Aircraft Emissions". In: *Engineering Proceedings* 13.1 (2021). ISSN: 2673-4591. DOI: 10.3390/engproc2021013011. URL: https://www.mdpi.com/2673-4591/13/1/11.
- [11] Gonçalo Soares Roque. GRAPE. Version 1.0. 2024. URL: https://github.com/goncaloroque30/ GRAPE.
- [12] ICAO. Doc 9889 Airport Air Quality Manual. 2020.
- [13] Doug DuBois and Gerald C. Paynter. ""Fuel Flow Method 2" for Estimating Aircraft Emissions". In: *SAE Transactions* 115 (2006), pp. 1–14.
- [14] Martin Strohmeier, Xavier Olive, Jannis Lübbe, Martin Schäfer, and Vincent Lenders. "Crowd-sourced 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/.
- [15] Xavier Olive. "Traffic, a Toolbox for Processing and Analysing Air Traffic Data". In: *Journal of Open Source Software* 4 (2019), p. 1518. DOI: 10.21105/joss.01518.
- [16] Gordon E. Andrews, Ian D. Andrews, Darren W. Dixon-Hardy, Bernard M. Gibbs, Hu Li, and Stephen Wright. "Airport PM10 Emissions: Development of a First Order Approximation (FOA) Methodology for Aircraft and Airport Particulate Mass Emissions". In: *Volume 2: Combustion, Fuels and Emissions, Parts A and B.* ASME Turbo Expo 2010: Power for Land, Sea, and Air. 2010, pp. 363–375. DOI: 10.1115/GT2010-22437.
- [17] Gabriel Jarry, Daniel Delahaye, and Eric Feron. "Approach and Landing Aircraft On-Board Parameters Estimation with LSTM Networks". In: 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT). 2020, pp. 1–6. DOI: 10.1109/ AIDA-AT48540.2020.9049199.
- [18] Aida Sharif Rohani, Tejas G. Puranik, and Krishna M. Kalyanam. "Machine Learning Approach for Aircraft Performance Model Parameter Estimation for Trajectory Prediction Applications". In: 2023 IEEE/AIAA 42nd Digital Avionics Systems Conference (DASC). 2023, pp. 1–9. DOI: 10. 1109/DASC58513.2023.10311271.

- [19] Jonas Meister. "Thrust Estimation for Departing Jet Aircraft Using Position Data Based on a Machine Learning Approach". In: 10th Convention of the European Acoustics Association. 2023. DOI: 10.61782/fa.2023.1131.
- [20] Jonas Meister. "Power Setting Estimation of Departing Civil Jet Aircraft Based on Machine Learning". In: *Journal of Aircraft* 61.4 (2024), pp. 1195–1204. DOI: 10.2514/1.C037619. URL: https://doi.org/10.2514/1.C037619.
- [21] Ran Giladi and Eliav Menachi. "Validating aircraft noise models: Aviation environmental design tool at Heathrow". In: *Journal of Air Transport Management* 116 (2024), p. 102557. ISSN: 0969-6997. DOI: https://doi.org/10.1016/j.jairtraman.2024.102557. URL: https://www. sciencedirect.com/science/article/pii/S096969972400022X.