https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

# Observing formation and early evolution of contrails formed by IAGOS aircraft using high-resolution LEO satellite imagery

Thymen Woldhuis[1], Zebediah Engberg[2], Susanne Rohs[3], and Vincent Meijer[1]
[1]Delft University of Technology, Faculty of Aerospace Engineering, Delft, The Netherlands
[2]Contrails.org, Kirkland, WA, United States of America
[3]Institute of Climate and Energy Systems - Troposphere (ICE-3), Forschungszentrum Jülich GmbH, Jülich, Germany.
**Correspondence:** Vincent Meijer (v.r.meijer@tudelft.nl)

**Abstract.** Persistent contrails and contrail cirrus are estimated to be a major contributor to the climate impact of aviation. The mitigation of these impacts by means of technological or operational changes benefits from the ability to skillfully model the formation, evolution, and impacts of contrails. Although these models can be evaluated and improved by use of observations of contrails obtained from remote sensing instruments, these comparisons are hindered by uncertainty in the required meteorological data (such as relative humidity) and limitations in the method of observation (such as younger contrails not being observable in geostationary satellite imagery). To address these challenges, we collocate aircraft equipped with in-situ humidity sensors from the IAGOS fleet in high-resolution (10-30 $\mathrm{m}$) satellite imagery obtained by instruments aboard the low Earth orbit Sentinel-2 and Landsat missions. The resulting dataset consists of 543 IAGOS aircraft found in satellite imagery (51% of which form contrails), which we use to evaluate predictions of contrail formation by the Schmidt-Appleman criterion (SAC) as well as predictions of contrail growth by the CoCiP model. When accounting for uncertainty in the IAGOS measurements of humidity and temperature, we find that the SAC correctly explains 98.3% of the observations. Disagreement between predictions and observation increases when using meteorological data from the ERA5 reanalysis, with only 92.1% of the observations being explained correctly. Out of the 195 annotated contrails, 48.2% of these contrails were found to persist for longer than 10 $\mathrm{s}$ (approximately the jet phase) and 8.7% longer than 120 $\mathrm{s}$ (approximately the vortex phase). The relative humidity with respect to ice is found to correlate most strongly with observed contrail lifetime, exhibiting an $\mathrm{R}^2$ value of 0.49 with the logarithm of contrail age. The observed horizontal growth during the jet and vortex phases is consistent with previous observations and contrail model results. Although the limited lifetimes of the annotated contrails prevent robust statistical conclusions for the dispersion phase, three example cases show horizontal growth rates consistent with simulations by CoCiP and that of observations in literature. Overall, this study demonstrates the potential of high-resolution LEO satellites to create observational datasets for evaluating and improving models of contrail formation and early evolution.

## 1 Introduction

Condensation trails, or contrails, are linear clouds formed behind aircraft when hot engine exhaust mixes with the colder, drier ambient air (Schumann, 1996). While most contrails dissipate within minutes, some can persist for several hours when the air is sufficiently humid with respect to ice, eventually turning into cirrus-like clouds (Kärcher, 2018). Contrails affect Earth's energy

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

balance by interacting with incoming solar and outgoing terrestrial radiation, resulting in a net warming effect, with current estimates suggesting that contrails are responsible for over half of aviation's total climate impact to date, with an effective radiative forcing of 57.4 $\mathrm{mW\ m^{-2}}$ for 2018 (Lee et al., 2021), albeit with high uncertainty (17 – 98 $\mathrm{mW\ m^{-2}}$, representing the 90% confidence interval).

Proposed options for the mitigation of contrail climate impacts include alternative fuels (Burkhardt et al., 2018; Voigt et al., 2021; Märkl et al., 2024) and re-routing aircraft to avoid the regions where persistent, warming contrails can form (Teoh et al., 2020; Martin Frias et al., 2024; Sausen et al., 2024; Sonabend-W et al., 2024). The implementation of such mitigation options is complicated by uncertainties in predictions of where contrails form, persist, and what their impact on the climate is (Petzold et al., 2025).

The uncertainties in these predictions can be assessed, and potentially reduced, by comparing model outputs to observations. These observations can come from in-situ measurements or remote sensing. In-situ observations can provide detailed information on the microphysics of contrails (Voigt et al., 2010, 2011; Jeßberger et al., 2013; Schumann et al., 2017; Voigt et al., 2021), but are often limited in their spatio-temporal coverage. On the other hand, remote sensing observations can provide larger spatio-temporal coverage but cannot provide information at the level of detail possible by in-situ measurements (Mannstein et al., 1999; Meijer et al., 2022; Geraedts et al., 2024). For example, geostationary satellite images are captured several times an hour and can be used to observe contrails (Meijer et al., 2022; Ng et al., 2023; Ortiz et al., 2025) and estimate their properties over large regions (Chevallier et al., 2023; Gryspeerdt et al., 2024; Meijer et al., 2024; Sarna et al., 2025; Sonabend-W et al., 2025).

The process of comparing these contrail observations to model outputs is complicated by limitations in the model input data and the reliability of observations. For example, numerical weather prediction (NWP) data of relative humidity, which is an important input to contrail models, are known to be unreliable in the upper troposphere (Gierens et al., 2020; Geraedts et al., 2024). Therefore, disagreement between observations and model predictions based on NWP data cannot unambiguously be attributed to errors within the contrail model. Observations themselves can also be affected by errors. For example, the relatively coarse (1-2 $\mathrm{km}$ spatial resolution at nadir) geostationary satellite images can miss thinner contrails (Driver et al., 2025; Euchenhofer et al., 2025). As a consequence, these images cannot be used to observe the first phases of the lifetime of a given contrail, and other instruments should be used instead, such as those available on the ground and low Earth orbit (LEO) satellites.

Thus, existing approaches for the evaluation of contrail models using observational data are constrained by uncertainty in relevant model input data as well as limitations of the method of observation used. Here, we address these challenges by locating in-situ measurement aircraft in high-resolution (10 to 30 $\mathrm{m}$) satellite imagery in order to evaluate models of contrail formation and early evolution while using measurement-based meteorological input data. Specifically, we collocate aircraft from the In-service Aircraft for a Global Observing System (IAGOS) program (Petzold et al., 2015) in satellite imagery from the Sentinel-2 and Landsat missions and compare this to predictions of contrail formation using the Schmidt-Appleman criterion (SAC) (Schumann, 1996) and contrail evolution from the Contrail Cirrus Prediction (CoCiP) model (Schumann, 2012).

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

The SAC has been evaluated using both in-situ and remote sensing observations of contrails, results of which have been summarised in Schumann et al. (2017). Based on 236 datapoints from 33 independent observation campaigns, that study reported that 93.6% of observed contrails were predicted by the SAC. However, because the dataset comprises only cases in which contrails were actually observed, the analysis could not evaluate situations where the SAC predicted contrail formation but none was observed. Recent ground-based observations that include aircraft both forming and not forming contrails found that SAC predictions based on NWP data correctly predict contrail formation in only 75.8% of cases (Low et al., 2025), or even fewer (Jarry et al., 2026). One reason for this might be the quality of the weather data. Gierens et al. (2020) investigated the impact of uncertainties in NWP data used for evaluating the SAC and found that NWP data correctly predicts the SAC in 94% of cases when compared to in-situ measurements. However, this study did not take any observations of contrails into account. It thus remains unclear whether the lower performance of the SAC found by Low et al. (2025) and Jarry et al. (2026) arises from limitations of the observations, uncertainties in the underlying meteorological input data, or the SAC itself.

In-situ and remote sensing observations have also been used to investigate contrail properties at different stages of their lifetimes (Iwabuchi et al., 2012; Schumann et al., 2017; Märkl et al., 2024; Low et al., 2025). These studies reported distributions of contrail properties, for example the width, depth, and ice water content, to be consistent with model simulations. Observations of young contrails (jet/vortex phase) are, however, limited, either due the difficulty of performing measurements in the aircraft wake (Jeßberger et al., 2013) or due to the limited detection capabilities of satellite instruments (Driver et al., 2025; Euchenhofer et al., 2025). LIDAR and ground-based camera studies have the ability to detect young contrails, but are constrained by their limited spatial coverage, restricting observations to the field of view of the site. High-resolution LEO satellites, however, do have the ability to detect contrails at formation, on a global-scale.

Both Sentinel-2 and Landsat have been used to observe aircraft in satellite imagery, for example for airspace management and surveillance (Zhao et al., 2018; Heiselberg, 2019; Liu et al., 2020). Observing contrails has been done in the Landsat (McCloskey et al., 2021) for the purpose of creating human-labelled datasets for evaluating contrail detection algorithms. However, this is done with a contrail-first approach, meaning contrails are annotated not linked to individual flights. Our approach differs in that we first locate the aircraft within the satellite image based on trajectory data, after which we annotate the contrail in case it is formed.

In this paper, we present a methodology for locating IAGOS aircraft in high-resolution Sentinel-2 and Landsat satellite imagery. The approach is used to collect data on contrail formation and early evolution for 543 IAGOS aircraft, 195 of which were found in Sentinel-2 imagery and thus include annotations of the entire observed contrail (in the case one was forrmed). The resulting dataset is used to evaluate the performance of the SAC in predicting contrail formation and how this is affected by input data quality (using ERA5 instead of IAGOS measurements) and satellite image resolution (by considering Landsat instead of Sentinel-2 collocations). We also compare the observations of contrail growth with corresponding model predictions, and study the relationship between observed contrail lifetimes and measurements of humidity and temperature.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

## 2 Data methodology

### 2.1 Satellite Data

### 2.1.1 Sentinel-2 Data Overview

The Sentinel-2 (Drusch et al., 2012) mission is part of Europe's Copernicus program (Berger et al., 2012). During the period relevant to this study (2013-2022), two satellite platforms were operational: the S2A (2015-2025) and the S2B (2017-present). By using the two-satellite constellation, the revisit time is reduced from 10 days to 5 days. All Sentinel-2 satellites are in a sun-synchronous orbit at an altitude of approximately 786 km. The Sentinel-2 satellites carry the Multi-Spectral Instrument (MSI) which acquires high spatial resolution (10 m to 60 m) imagery in 13 bands in the visible and near-infrared part of the electromagnetic spectrum, with a swath width of 290 km. The four MSI bands used in this study, along with their wavelengths and spatial resolution, are given in Table 1. The satellite image data is subdivided on a predefined set of 110 by 110 $\mathrm{km}$ tiles defined in the UTM/WGS84 projection. We use L1C data which is downloaded from a public Google Cloud Storage dataset (Google, 2026).

### 2.1.2 Landsat 8/9 Data Overview

The Landsat 8 (2013-present) and Landsat 9 (2021-present) satellites (Irons et al., 2012; Knight et al., 2014; Masek et al., 2020) are operated by the U.S. Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA). The Landsat 8 and 9, hereafter referred to as Landsat, have a revisit time of 16 days each, or 8 days when using both satellites. Both satellites are in a sun-synchronous orbit at an altitude of 705 $\mathrm{km}$ and have a swath width of 185 $\mathrm{km}$. The Landsat carries two instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The spectral bands used, along with their wavelengths and spatial resolution are given in Table 1. We use the Landsat 8 and 9, Collection 2, Level 1 data downloaded from the USGS EROS machine-2-machine API (U.S. Geological Survey).

### 2.2 IAGOS fleet and data

The In-service Aircraft for a Global Observing System (IAGOS) (Petzold et al., 2015) operates measuring equipment on commercial aircraft to provide in-situ atmospheric measurements on a global scale. The IAGOS dataset, which starts in 1994, consists of the former research project MOZAIC, CARABIC (now IAGOS-CARABIC), and the currently active CORE program. In this study, we only considered the aircraft that are part of the CORE program, as shown in Table 4, as these were predominantly active throughout the operational periods of Sentinel-2 and Landsat 8 and 9.

We use the air temperature ($T$) and the relative humidity w.r.t. liquid ($\mathrm{RH}_\ell$) from the IAGOS-CORE dataset, both of which are measured by the IAGOS capacitive hygrometer (ICH). The ICH uses a capacitive humidity sensor (a modified Vaisala HUMICAP® sensor) and a platinum resistance temperature sensor (PT100) (Helten et al., 1998; Neis et al., 2015a). The IAGOS aircraft trajectory data and ICH measurements are provided every 4 $\mathrm{s}$. During this study we observed a time shift in the IAGOS data in the order of 0-15 $\mathrm{s}$ when compared to ADS-B data or satellite imagery. To resolve this, we linearly interpolate

**Table 1.** Sentinel-2 MSI and Landsat 8/9 OLI and TIRS bands used in this study, along with their specifications.

| Satellite | Instrument | Band number | Name | Central wavelength (nm) | Bandwidth (nm) | Resolution (m) |
|---|---|---|---|---|---|---|
| Sentinel-2 | MSI | 2 | Blue | 490 | 65 | 10 |
| Sentinel-2 | MSI | 3 | Green | 560 | 35 | 10 |
| Sentinel-2 | MSI | 4 | Red | 665 | 30 | 10 |
| Sentinel-2 | MSI | 10 | Cirrus | 1375 | 30 | 60 |
| Landsat | OLI | 2 | Blue | 482 | 60 | 30 |
| Landsat | OLI | 3 | Green | 561 | 60 | 30 |
| Landsat | OLI | 4 | Red | 655 | 30 | 30 |
| Landsat | OLI | 9 | Cirrus | 1373 | 20 | 30 |
| Landsat | TIRS | 10 | TIRS-1 | 10895 | 590 | 100 |
| Landsat | TIRS | 11 | TIRS-2 | 12005 | 1010 | 100 |

the IAGOS measurement and position data to a 0.1 s resolution and select the time at which the aircraft position is closest to the annotated aircraft position, when corrected for parallax (see subsection C4).

IAGOS data undergoes quality control, the outcome of which is communicated by means of a validity flag. These flags are either good (0), limited (2), erroneous (3), not validated (4), or missing value (7) (IAGOS). For this study, only measurements flagged as good (0), for both $T$ and RH$_\ell$, were used. This entails that the measurements passed all required quality control tests. In addition to these validity flags, we utilise the estimates of measurement uncertainty that come with IAGOS data for both $T$ and RH$_\ell$ to construct 95% confidence intervals. The uncertainty of $T$ and RH$_\ell$ measurements are estimated to be 0.5 K (Berkes et al., 2017) and 5%-6% (Neis et al., 2015a, b) respectively, with the latter determined by comparing the ICH measurements with the FISH (Meyer et al., 2015) and SEALDH-II (Buchholz et al., 2013) instruments during the AIRTOSS-ICE field campaign (Neis et al., 2015a). However, instead of taking this constant $T$ and RH$_\ell$ uncertainty, we rely on the uncertainty estimates provided directly by the IAGOS data processing. In general, these uncertainties are comparable to the 0.5 K for air temperature and 5%-6% for RH$_\ell$, but they offer a more accurate representation of the uncertainty associated with individual measurements.

Humidity measurements by the ICH in the upper troposphere and lower stratosphere (i.e. the aircraft cruise altitude region) may also be affected by reduced sensor sensitivity and long sensor response times. Firstly, relative humidity measurements by the ICH have been reported to suffer from reduced sensitivity to water vapour concentrations in dry stratospheric air masses (Kunz et al., 2008; Rolf et al., 2024), with a lower detection limit between 10 ppmv (Engel et al., 2006) and 30 ppmv (Rolf et al., 2024). This loss of sensitivity is primarily caused by adiabatic heating of the air flow, as it is brought to rest within the ICH housing. Because of this temperature increase, the measured RH$_\ell$ is lower than the ambient RH$_\ell$. The effect of this lower detection limit on the analysis conducted in this study will be further discussed in subsection 2.5. Secondly, the humidity sensor has a temperature-dependent response time, which varies from 1 s at 300 K to 120 s at 210 K. Several studies

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

have attempted to correct this time delay or mitigate its effects on downstream analysis. For example, (Gierens et al., 2007) aggregates measurements to a coarser 1 $\mathrm{min}$ resolution, whereas (Borella et al., 2024) tries to reconstruct the measured signal by use of a temperature-dependent exponential moving average. We find that reconstructing $\mathrm{RH}_\ell$ for the IAGOS flights at 4 $\mathrm{s}$ timestamps results in a noisy signal (on average increasing the first-difference standard deviation by 43.3%), resulting in more outlier measurements. Resampling to 1 $\mathrm{min}$ timestamps mitigates this noise, but does so at the expense of temporal resolution. Therefore, we chose not to apply any of the aforementioned correction methods.

### 2.3 Numerical Weather Prediction Data

We use numerical weather prediction (NWP) data from the ECMWF ERA5 reanalysis dataset (Hersbach et al., 2020). Air temperature ($T$), specific humidity ($q$) and pressure ($p$) are used to compute $\mathrm{RH}_\ell$ and $\mathrm{RH}_i$ with the expressions for the saturation pressure w.r.t. liquid from Buck (1981) and ice from Alduchov and Eskridge (1996). Other variables required to use the CoCiP contrail model (Schumann, 2012) also come from the ERA5 dataset. The reanalysis data is downloaded from the Analysis-Ready, Cloud Optimized (ARCO) ERA5 Dataset (Carver and Merose, 2023) hosted on Google Cloud, which has a regular grid of $0.25° \times 0.25°$ horizontal resolution and includes 137 vertical model levels extending from the surface to approximately 80 $\mathrm{km}$. The air temperature and relative humidity are linearly interpolated in the horizontal and vertical dimensions to match the aircraft's position and time.

### 2.4 Aircraft Performance Modeling

We use two aircraft performance models (APMs) to estimate the engine fuel mass flow and the overall engine efficiency which are required as inputs for modeling contrail formation and evolution: the Base of Aircraft Data (BADA) Family 4 Release 4.2 (EUROCONTROL, 2016) and the Poll-Schumann (PS) (Poll and Schumann, 2021a, b, 2025) model. The aircraft type and engine combinations for the IAGOS fleet are given in Table 4. All computations were performed using the open-source *pycontrails* library (Shapiro et al., 2026). Flight trajectories provided by IAGOS at 4 $\mathrm{s}$ resolution were resampled to 1 $\mathrm{min}$, in order to reduce the influence of small positional uncertainties and obtain smoother overall engine efficiency outputs. Both APMs require data on air temperature and horizontal winds as input, which can either come from ERA5 or the IAGOS measurements. We have quantified the changes in overall engine efficiency estimated by both APMs when using these different meteorological data sources: the resulting root-mean square "error" was found to be close to 0.005 for the PS model, and 0.009 for the BADA 4.2 model, which is almost 2 orders of magnitude smaller than typical engine efficiency values of 0.3-0.4. Due to the relatively small magnitude of these meteorology-dependent efficiency differences, we have chosen to perform aircraft performance calculations with the data source that is used for the prediction of contrail formation, which we discuss in more detail in subsection 2.5.

Comparison of the two APMs shows that BADA 4.2 engine efficiencies show larger intra-flight variability than those obtained from the PS model. As a result, the efficiency values calculated for the satellite observation times also show increased variability, typically ranging between 0.2 and 0.4 during cruise, as opposed to 0.25 to 0.35 with the PS model. Moreover, during descent, BADA 4.2 calculates an overall engine efficiency of zero, whereas the PS model calculates values between 0.1 and

0.3. For the subsequent analyses, we therefore only use the PS engine efficiency estimates. Additional details and a comparison of the two performance models are provided in subsection A3.

In the PS model, the overall engine efficiency at the optimum condition is expected to have an uncertainty of no more than ±10% (Poll and Schumann, 2021b). Since this uncertainty is related to the optimum condition, an extra 2.5% buffer is added to account for suboptimal conditions. In the updated version (Poll and Schumann, 2025), an extra uncertainty of ±2.5% is added due to the effect of in-service engine deterioration, based on the reasoning that an airline only allows for an increase of 5% in fuel consumption before a maintenance intervention is planned. Thus, we take an uncertainty related to the overall engine efficiency of 15% into consideration for the subsequent analyses. For example, for an engine efficiency of 0.3, we take the confidence interval as [0.255, 0.345].

## 2.5 Contrail Models

### 2.5.1 Schmidt-Appleman criterion

Contrails form when water condenses onto particles in the engine exhaust plume. For this condensation to occur, the engine exhaust plume is required to be supersaturated with respect to liquid for some period of time during the mixing process with the ambient air. The Schmidt-Appleman criterion (SAC) (Schumann, 1996) can be used to ascertain whether this supersaturation occurs, by modeling the mixing of the engine exhaust with the ambient air as an isobaric process. As a consequence, the mixing process can be represented by a straight line in a chart of temperature versus water vapor partial pressure ($e$) which connects the ambient air state with the state of the engine exhaust. When this "mixing line" is plotted in conjunction with curves representing the saturation pressure w.r.t. liquid ($e_\ell$) and ice ($e_i$), the SAC can be visually confirmed by checking whether the mixing line ever lies above the curve corresponding to $e_\ell$. Schumann (1996) has derived an expression for the slope of the mixing line ($G$, in Pa K$^{-1}$) as a function of engine, fuel and atmospheric parameters

$$G = \frac{c_p p}{\varepsilon} \frac{\mathrm{EI}_{\mathrm{H_2O}}}{\mathrm{LHV}(1-\eta)}. \tag{1}$$

Here, $\mathrm{EI}_{\mathrm{H_2O}}$ is the emission index of water vapor associated with combustion of the jet A-1 fuel, $c_p$ is the heat capacity of air at constant pressure, $p$ is the ambient pressure, $\varepsilon$ is the ratio of the molar mass of water vapor to that of dry air, LHV is the lower heating value of the fuel, and $\eta$ is the engine efficiency. The values used for all these parameters, with exception of $p$ and $\eta$ (which vary), are given in Table 2.

Given the slope $G$ and the ambient temperature and partial pressure of water vapor, we are able to plot the mixing line and ascertain whether a contrail can form. An alternative approach is to, given the slope $G$, determine the range of ambient conditions for which contrail formation is possible or not. This involves drawing a "limiting mixing line" which is tangent to the curve of saturation pressure w.r.t. liquid. The point of tangency, corresponding to temperature $T_{LM}$, is defined by

$$\frac{de_\ell(T_{LM})}{dT} = G. \tag{2}$$

**Table 2.** Values and sources of the fixed parameters used in the mixing slope.

| Parameter | Value | Unit | Source |
|---|---|---|---|
| $\mathrm{EI}_{\mathrm{H_2O}}$ | 1.23 | $\mathrm{kg\,kg^{-1}}$ | Lee et al. (2021) |
| $c_p$ | 1004.66 | $\mathrm{J\,kg^{-1}\,K^{-1}}$ | - |
| $\varepsilon$ | 0.622 | - | - |
| LHV | 43.2 | $\mathrm{MJ\,kg^{-1}}$ | Voigt et al. (2021) |

Generally, the functional form of $e_\ell(T)$ prohibits an analytical solution for $T_{LM}$. Instead, we use the approximation for $T_{LM}$ given by Schumann (1996)

$$T_{LM} = -46.46 + 9.43\ln(G - 0.053) + 0.720[\ln(G - 0.053)]^2, \tag{3}$$

where $T_{LM}$ is in K and G is in $\mathrm{Pa\,K^{-1}}$. Given $G$ and $T_{LM}$, it is possible to determine whether certain ambient conditions allow for contrails to form or not. In order to quantify how "close" particular ambient conditions (in this case $T$ and $\mathrm{RH}_\ell$) are to the threshold of contrail formation, we also quantify the temperature at which the same $\mathrm{RH}_\ell$ would lead to such threshold conditions. This temperature, denoted as $T_{LC}$, is found by solving

$$T_{LC} = T_{LM} - \frac{e_L(T_{LM}) - \mathrm{RH}_\ell e_L(T_{LC})}{G}, \tag{4}$$

which is done numerically by use of Newton's method. Finally, the difference between this temperature and the ambient one is computed

$$\Delta T_{LC} = T - T_{LC}. \tag{5}$$

When $\Delta T_{LC} < 0$, the SAC is satisfied and contrails form.

As mentioned in subsection 2.2, the IAGOS ICH sensor has a lower detection limit (in terms of water vapor mixing ratio) between 10 $\mathrm{ppmv}$ and 30 $\mathrm{ppmv}$. At temperatures below 215 K, ice supersaturation ($\mathrm{RH}_i > 1$) can occur at these water vapor mixing ratios. Although the SAC depends on aircraft/engine parameters (e.g. $\eta$) and ambient conditions ($p$, $T$, and $\mathrm{RH}_\ell$), it is usually already satisfied in completely dry air for temperatures below 215 K, meaning that inaccurate $\mathrm{RH}_\ell$ measurements will not affect the outcome of the SAC. In the temperature range 215 K to 230 K, the slow time response can cause inaccurate measurements of $\mathrm{RH}_\ell$. Given that the air temperature can be still measured with an uncertainty of 0.5 K and the fact that the satisfaction of the SAC is predominantly influenced by the temperature (Gierens et al., 2020), the slow time response of the humidity sensor is not expected to strongly affect the results. Nevertheless, we will take into account these and other uncertainties in our evaluation.

### 2.5.2 Contrail Cirrus Prediction Tool (CoCiP)

The Contrail Cirrus Prediction Tool (CoCiP) (Schumann, 2012) is a Lagrangian Gaussian plume model that simulates the full contrail life-cycle. CoCiP does not explicitly model the evolution of the contrail dimensions in the jet and vortex phase. It uses

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

a parametric model to estimate the initial width, depth, and downward displacement that are used to initialise the dispersion phase. This initial width of the contrail is thus the width at the end of the wake vortex phase, which takes place approximately 120 s after formation (Hoshizaki et al., 1972), and is given by (Schumann, 2012)

$$B_1 = \frac{N_{\text{dil}}(t_0)\dot{m}_F}{(\pi/4)\rho D_1}, \tag{6}$$

where $N_{\text{dil}}(t_0)$ is the dilution factor at the time $t = t_0$, with $t_0$ being the effective time scale; $\dot{m}_F$ is the mass fuel flow in $\text{kg m}^{-1}$; $\rho$ the ambient air pressure in $\text{kg m}^{-3}$; and $D_1$ the initial depth in m, which is a function of the maximum downward displacement $z_w$. This maximum downward displacement is a function of the wake vortex separation, $b_0$, initial circulation, $\Gamma_0$, and the Brunt-Väisälä frequency $N_{\text{BV}}$.

The lifetime of a contrail in CoCiP is determined by the time at which the contrail segment has fully sublimated. In CoCiP, sublimation is assumed to occur when either the ice particle number concentration drops below $10^3\ \text{m}^{-3}$ or the optical depth falls below $10^{-6}$.

We use the CoCiP implementation from *pycontrails* (Shapiro et al., 2026). Flight trajectory data were obtained from IAGOS and resampled to a temporal resolution of 1 s, instead of the default 4 s. Because LEO satellites typically observe only a short segment of an individual contrail (because it is limited by the swath width of the satellite), simulations were run for a maximum contrail age of 20 min with 1 s time steps. The scanning time of the satellite was rounded to the nearest second, such that the contrail dimensions at the scanning time of the satellite could be extracted for each waypoint.

The meteorological input data used to run CoCiP are obtained from ERA5. However, the relative humidity w.r.t. ice ($\text{RH}_i$) in ERA5 shows disagreement with in-situ measurements (Gierens et al., 2020). Several approaches exist to correct ERA5 $\text{RH}_i$ for contrail modelling, typically by applying correction factors derived from aggregate statistics of IAGOS observations (Teoh et al., 2022, 2024), or using machine learning methods to improve the NWP data (Wang et al., 2025). In this study, in-situ IAGOS $\text{RH}_i$ measurements are available, allowing for a more accurate, local correction. The ERA5 $\text{RH}_i$ values used by CoCiP are adjusted by scaling each value with a factor equal to the ratio of the IAGOS $\text{RH}_i$ and the ERA5 $\text{RH}_i$ at the scanning time of the satellite.

### 2.6 Collocation of aircraft in LEO satellite imagery

We use aircraft position data (either from IAGOS data or by use of ADS-B data) in combination with satellite ephemeris and imaging sensor specifications to estimate if and where a given aircraft should be located in a particular satellite image: this process is referred to as “collocation”. We find that interpolation of aircraft position data to the image capture time and converting this to the local UTM coordinate system used within the image, leads to differences between the predicted and observed aircraft position on the order of several kilometers. Given that two or more aircraft are occasionally found within a distance of a few kilometers in these images, we apply several corrections to the collocation process in order to resolve such ambiguities. Specifically, we improve the accuracy of the collocation process by accounting for parallax displacement of the aircraft and by using a more precise image capture or “scan” time corresponding to the specific area of the image where the aircraft is expected to be located.

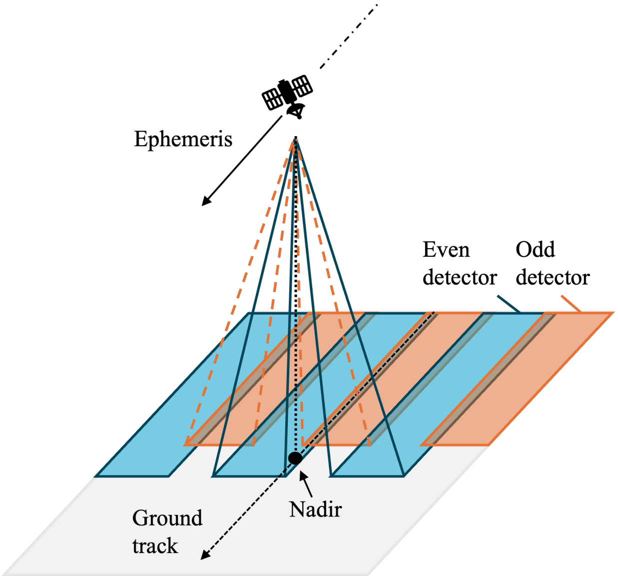

**Figure 1.** Diagram of scanning time discontinuity due to detector architecture in sensing instrument (note that only 6 detectors are shown of the total 14 detectors in Landsat and 12 detectors in Sentinel-2). Odd detectors (orange) are rear-facing, even detectors (blue) are forward-facing. There is a small overlap between the detectors.

The images are orthorectified to the Earth's surface during Landsat and Sentinel-2 processing, meaning that pixels are geometrically corrected assuming they lie at ground level. While this correction ensures accurate positions for surface features, it introduces a parallax displacement for objects located above the surface of the Earth when they are captured by the instrument at an off-nadir angle. Given the viewing geometry (represented by zenith and azimuth angles) of the satellite instrument and the altitude of the object, this parallax displacement, typically on the order of 1 to 2 $\mathrm{km}$, can be estimated. These viewing angles are provided for each image pixel in the Landsat image metadata, whereas the Sentinel-2 imagery only provides a coarse resolution ($25 \times 25$ km) estimate. We have created an algorithm to correctly upsample the Sentinel-2 viewing angle data which takes into account the staggered architecture of the sensors in the Sentinel-2 MSI (see Figure 1). This algorithm is described in Appendix C. The geometric (GNSS) altitude of the aircraft, with respect to sea-level (WGS-84), is obtained from the aircraft position data.

The second correction involves the estimation of the time at which individual image pixels were scanned. Although each satellite image is given a single "sensing time", this corresponds to the pixel scan time averaged over the entire scene. In practice, scanning such a scene (110 $\mathrm{km}$ by 110 $\mathrm{km}$ for Sentinel-2 and 170 $\mathrm{km}$ by 185 $\mathrm{km}$ for Landsat) takes approximately 16.8 $\mathrm{s}$ for the Sentinel-2 instrument and 32 $\mathrm{s}$ for the Landsat instrument, during which an aircraft cruising at 250 $\mathrm{ms}^{-1}$ can move 4.2 and 7.6 $\mathrm{km}$, respectively. For this reason, it is important to obtain a pixel-by-pixel estimate of the scan time. Both the Landsat and Sentinel-2 imagers are push-broom scanners, but their staggered sensor architecture implies that the pixel scan time is not only a function of the along-track position (as one would expect for a push-broom scanner) but also a function of the cross-track position (determining whether the pixel is scanned by an odd or an even detector). To account for this, we first find the satellite ephemeris position (and the corresponding time) closest to a given pixel, and then apply a constant time offset that depends on the detector used to capture the pixel. Further details on this process can be found in Appendix C.

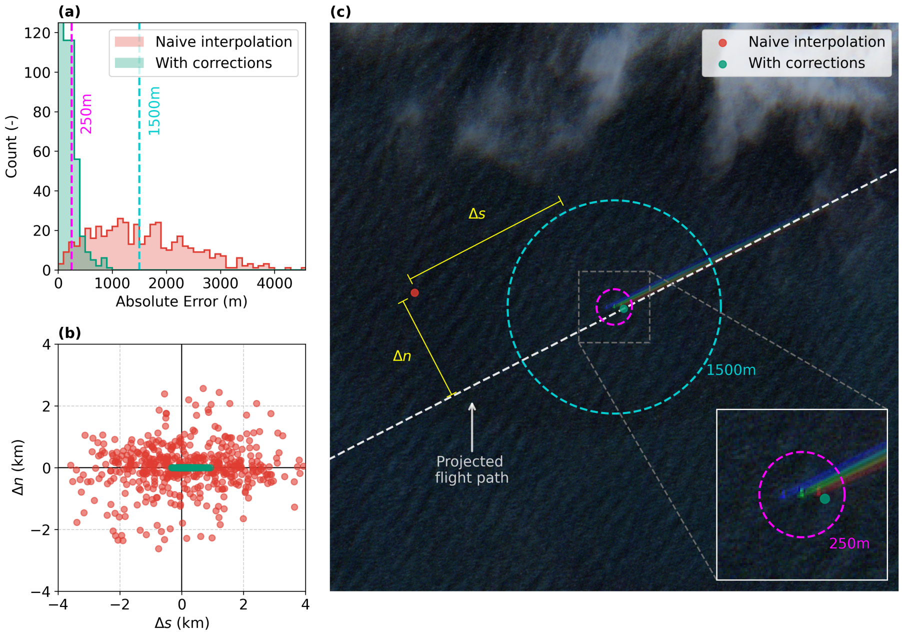

**Figure 2.** (a) Histogram of collocation errors for naive interpolation (red) and after scanning time and parallax corrections (green). The two vertical dotted lines indicate reference errors for 250 $\mathrm{m}$ (magenta) and 1500 $\mathrm{m}$ (cyan). (b) Scatterplot showing the errors normal, $\Delta n$, and tangential, $\Delta s$, to the flight path from the predicted positions to the manually annotated locations. (c) Sentinel-2 true colour image with the PH-TFM (Boeing 787-8 Dreamliner) aircraft, with the 250 $\mathrm{m}$ and 1500 $\mathrm{m}$ reference errors. Collocation for naive interpolation (red) and after corrections (green) shown with the $\Delta n$ and $\Delta s$ errors (yellow).

The results of these corrections to the collocation process are evaluated by quantifying the distance between the predicted aircraft position and the manually annotated position of the aircraft in the satellite image, and are shown in Figure 2 for the Sentinel-2. The collocation errors for Landsat are shown in Appendix C. We decompose the error into components normal and tangential to the flight path. As shown in Figure 2b, the residual error after applying the corrections is predominantly tangential to the flight path. To help interpret the magnitude of the error statistics shown in Figure 2a-b, we show an example of an aircraft found in a Sentinel-2 image in Figure 2c along with the predicted aircraft positions with and without corrections. The corrected collocation process estimates the position of 95% of the aircraft to within a distance of 487 $\mathrm{m}$, whereas the process without the discussed corrections can only narrow down the aircraft position to 3.06 $\mathrm{km}$ (for 95% of the data). This constitutes an order of magnitude improvement in the accuracy of the collocation process.

We hypothesise that the residual errors in the collocated positions are due to: ADS-B timestamp and positional uncertainties (typically in the order of several tens of meters according to Schäfer and Jonáš (2025)), the omission of digital elevation models

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

(which are used in Sentinel-2 (Copernicus, 2024) and Landsat (Franks et al., 2020) processing) in the collocation algorithm, and optical distortions due to the lens composition (Binet et al., 2022).

### 2.7 Annotation of contrails formed by collocated aircraft

The collocation process described in subsection 2.6 is used to find the location of IAGOS aircraft within both Sentinel-2 and Landsat imagery, after which we manually identify whether the aircraft in question is forming a contrail. If a contrail is found, it is annotated either by use of a (multi)polygon (for Sentinel-2 imagery) or a linestring along the longitudinal axis of the contrail (for Landsat imagery).

The first step in this annotation process is to determine the exact location of the aircraft, which is typically found close to but not exactly at the collocated position. In 12% of cases, out of the total number of cases where an intersection is found between the IAGOS trajectory and the satellite tiles, the aircraft was found to be obscured by clouds or the trajectory data was incomplete. These cases are excluded from the analysis. The second step is to determine whether the aircraft - in case it is visible in the imagery - is forming a contrail. If this is indeed the case, we also annotate the outline (centerline) of the contrail within Sentinel-2 (Landsat) imagery using a polygon (linestring).

For annotating the Sentinel-2 imagery we use the Copernicus Browser (Copernicus), an ESA platform for visualizing satellite data. This tool automatically mosaics adjacent images, enabling contrails to be annotated continuously across the full Sentinel-2 swath width of approximately 290 $\mathrm{km}$. The contrails were manually outlined by the authors as polygon features following their visible extent in the green band, which offers 10 $\mathrm{m}$ spatial resolution and generally the highest contrast against the background; the blue band appears less bright over clouds or ocean background, while the red band is less bright over land. Additional bands, including the red, blue, and cirrus bands, were used as aid during the annotation process. The cirrus band was especially useful for long-lived contrails or contrails present above white or “noisy” surfaces (e.g. snowy mountains, salt planes, or cities). In Appendix D, we show four example contrails over clouds, an ocean, a forest, and an urban area (see Figure D1). In addition, we include two cases in which the contrail is visible either only in the true-colour image or only in the cirrus band (see Figure D2).

For annotating the contrails found within Landsat imagery: we download the relevant scenes (a scene covers an area of 170 by 185 $\mathrm{km}$) and process these to obtain zoomed images centered on the collocated aircraft position (covering an area of 8 by 8 $\mathrm{km}$). In each case, we computed three images: a true colour image, a monochromatic image from the cirrus band, and a false colour image from the thermal bands (B10 and B11), following the colour scheme from McCloskey et al. (2021). The cirrus band was used for annotating the contrail, with help of other true colour and false colour images where necessary. Figure D4 shows an example of a contrail as seen in the true colour image, cirrus band, and false colour image.

In both Landsat and Sentinel-2 imagery, a fraction of annotated contrails extends beyond the swath: implying that we cannot see the full extent of the contrail. Such annotations are flagged accordingly, to avoid potential misinterpretation of the annotation.

The Sentinel-2 contrail annotations are further processed to obtain estimates of the contrail width as a function of its age. To estimate the centerline of a polygon, we first compute its minimum rotated rectangle (MRR). When multiple polygons are present, as is the case when the contrail crosses the detector borders, each polygon is processed individually. Then, we compute the longest edge of the MRR and its centroid to construct a centerline along the longest axis. We place points along the polygon centerline spaced by 100 $\mathrm{m}$ and estimate the contrail width at these points by computing the width of the polygon orthogonal to the centerline. The age of a contrail at each point (along the centerline) is determined by first calculating the distance from that point to the location of the aircraft. This distance is then divided by the aircraft's groundspeed to estimate the contrail age. The main limitation of this method is that it does not account for horizontal advection of the contrail by the wind. For instance, if the horizontal wind vector is aligned with the direction of flight, a point located 1000 $\mathrm{m}$ from the aircraft may have been advected further downstream or upstream. Assuming a wind speed of 50 $\mathrm{ms}^{-1}$ and an aircraft groundspeed of 250 $\mathrm{ms}^{-1}$, this can lead to an error in contrail age of approximately 20%. Alternative methods, such as generating centerline points by advecting the flight path according to wind, are limited by the accumulation of wind speed errors: over time, these errors can cause the advected path to diverge from the actual contrail. As a result, measurements of contrail width may be taken at an angle that is not perpendicular to the contrail, leading to an overestimation of its width. Thus, the choice of method is a trade-off between uncertainties in contrail age or in contrail width.

An example of a Sentinel-2 annotation and the corresponding estimated contrail width is given in Figure 3. As mentioned before, while the green visible band is used for the annotation, we visualise the cirrus band in Figure 3d for clarity. The estimated contrail width (see Figure 3f) is not found to be a smooth function of contrail age: several "jumps" are noticeable for contrail ages larger than 100 s. These jumps are in part due to the contrail crossing detector boundaries: which introduces discontinuities in the contrail annotation (see Figure 3b). However, the remainder of the variability in Figure 3 corresponds to the breakdown of the wake vortex pair. In other annotations such as Figure D3, we can directly see the inception of the Crow instability (Crow, 1970) which ultimately leads to the breakdown of the vortex pair.

### 2.8 Evaluation of the Schmidt-Appleman criterion and comparison to literature

We assess the agreement between SAC predictions and satellite observations using a contingency table, given in Table 3, which classifies each prediction as a true positive (TP), false positive (FP), false negative (FN), or true negative (TN).

**Table 3.** Contingency table for the validation of the SAC against satellite observations.

| | **Observed (Yes)** | **Observed (No)** |
|---|---|---|
| **Predicted (Yes)** | True Positive (TP) | False Positive (FP) |
| **Predicted (No)** | False Negative (FN) | True Negative (TN) |

The entries of the contingency table are combined into metrics that are used to assess the performance of the SAC. Specifically, we make us the precision Pr and recall Re which are defined by

$$\mathrm{Pr} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \qquad \mathrm{Re} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}. \tag{7}$$

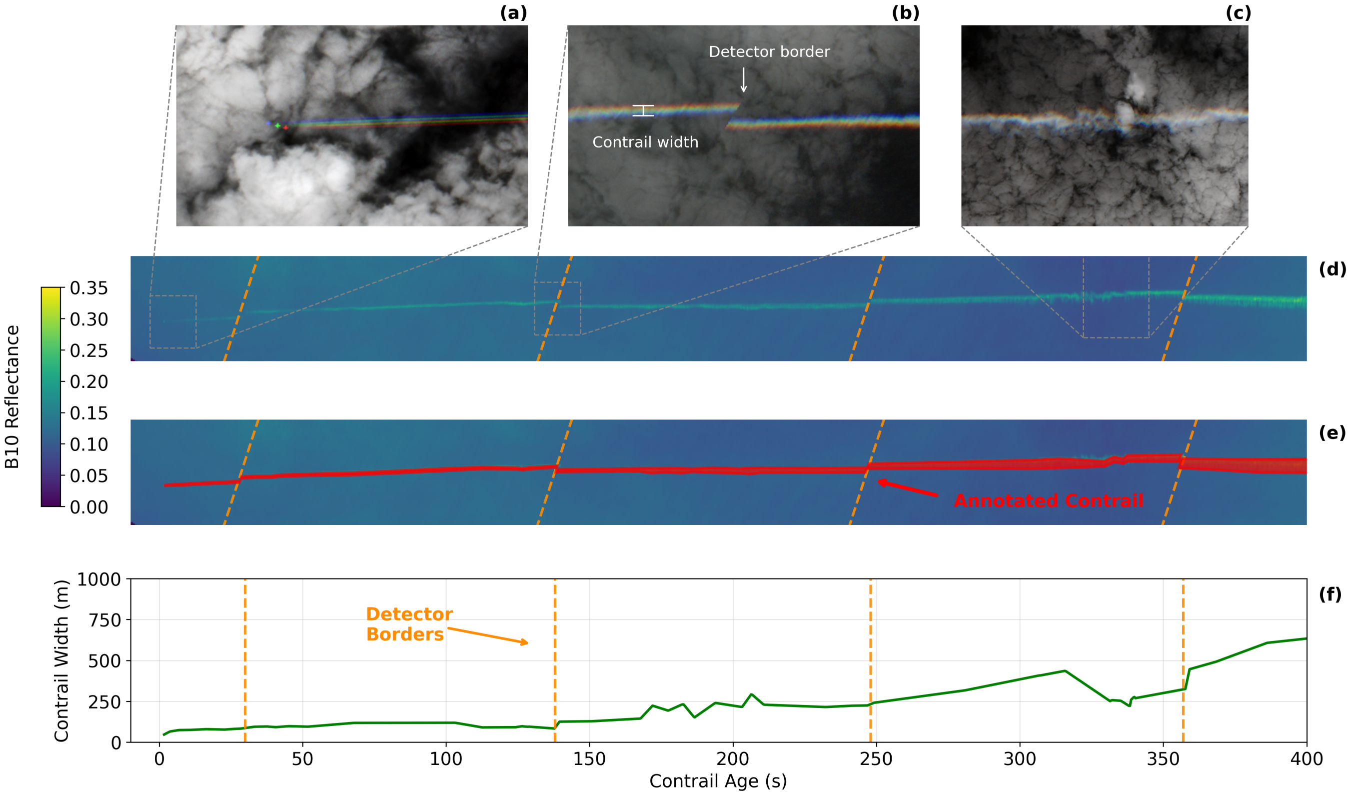

**Figure 3.** (a) Sentinel-2 true colour image showing the aircraft and the contrail in the first 10 s. (b) Sentinel-2 true colour image of the contrail after 140 s showing discontinuity caused by the detector border. (c) Sentinel-2 true colour image showing how the contrail loses its linearity after the wake vortex pair has broken down. (d) Sentinel-2 B10 top-of-atmosphere reflectance of the contrail. The vertical dimension of the image has been increased by a factor of 2 as compared to the horizontal dimension to improve the visibility of the contrail. (e) Same as (d), with manually annotated contrail polygon. (f) Contrail width as a function of contrail age. Dashed lines indicate the locations of the detector borders.

The precision corresponds to the fraction of predicted contrails that are observed, and the recall corresponds to the fraction of observed contrails that were also predicted to form. They have been used previously to quantify the performance of contrail forecasts by Geraedts et al. (2024) and Meijer (2024).

In addition to using the dataset created in this work, we also use two existing observational datasets of contrails for quantifying the performance of the SAC. These are the dataset presented by Low et al. (2025) and the dataset described by Jarry et al. (2026), both of which consist of contrails found in ground-based camera images. We also compare the differences in SAC performance found when switching meteorological data sources to the results by Gierens et al. (2020).

Both Gierens et al. (2020) and Low et al. (2025) directly provide entries of the contingency table, which we use to compute precision and recall according to the definitions above. In contrast, Jarry et al. (2026) do not provide these numbers and we

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

therefore process their dataset to obtain the entries of the contingency table. Their dataset consists of ground-based camera images with contrail annotations that are matched to flights passing through the camera field of view (FOV) based on ADS-B data, not limited to IAGOS aircraft. In order to be included in the contingency table, we require that 1) the camera zenith angle is limited to a maximum of 60° to avoid geometric distortion near the image borders; 2) only flights above an altitude of 7 km are considered, as lower-altitude aircraft are not expected to form contrails; 3) flights without a valid flight_id are discarded to avoid erroneous matches caused by unreliable ADS-B data. For each remaining flight we determine whether it formed a contrail by checking for annotations with the relevant flight_id in the images where the flight is within the FOV and images after the flight exited the FOV. The SAC has been computed for each flight with an overall engine efficiency of 0.30 (0.35 and 0.40 were also tested but resulted in lower precision values) using temperature, humidity and pressure from the the ECMWF ERA5 reanalysis dataset (Hersbach et al., 2020).

## 3 Results and Discussions

### 3.1 Contrail Dataset Overview

The application of the collocation process to IAGOS data for the years 2013 to 2022 has resulted in a total of 543 annotated collocations (366 in Sentinel-2 imagery and 177 in Landsat 8/9 imagery). For 195 of these Sentinel-2 collocations, we have annotated the contrail that is being formed by the IAGOS aircraft. The spatial and temporal distribution of the collocations are shown in Figure 4. Most collocations are found over (or close to) land masses, since the Sentinel-2 and Landsat satellites are primarily used for land observations. IAGOS aircraft have been collocated in both climb (54 times), cruise (455 times) and descent (34 times) phases. The mean geometric (GNSS) altitude of the collocations is 10.5 $\mathrm{km}$, with 90% of the collocations being above 9.0 $\mathrm{km}$. Further details on the vertical distribution of the collocations are discussed in subsection 3.2. Figure 4b shows that almost all collocations occur from 2017 onward, with a single exception in 2013. The limited availability of data between 2013 and 2018 can be attributed to an error in the IAGOS data acquisition until 2018, which has since been fixed, but results in validity flags of 2 for all repaired data and are therefore excluded from the analysis. Additionally, the decrease in 2020–2021 can be attributed to a reduced number of flights during the COVID-19 pandemic. In the period 2017 to 2021, the Sentinel-2 program had two satellites operational, compared to only one Landsat satellite (Landsat 9 was only commissioned at the end of 2021), resulting in more Sentinel-2 collocations.

The number of collocations found for each aircraft in the IAGOS fleet as well as the total number of flights during the years 2013 to 2022 are given in Table 4, along with details on the aircraft type and engine. Since both Sentinel-2 and Landsat satellites operate in sun-synchronous orbits, each location on Earth is observed consistently at the same local time. Consequently, the number of collocations associated with each aircraft is strongly influenced by typical flight routes and departure schedules. For example, F-GZCO recorded 474 flights in 2022, yet only 10 collocations were identified. During this period, the aircraft predominantly operated routes across Europe and North-Africa, with 86% of departures occurring after the local satellite revisit time. In contrast, aircraft operating inter-continental routes, such as transatlantic flights, have a higher probability of being collocated, as their trajectories can be captured by multiple satellite overpasses.

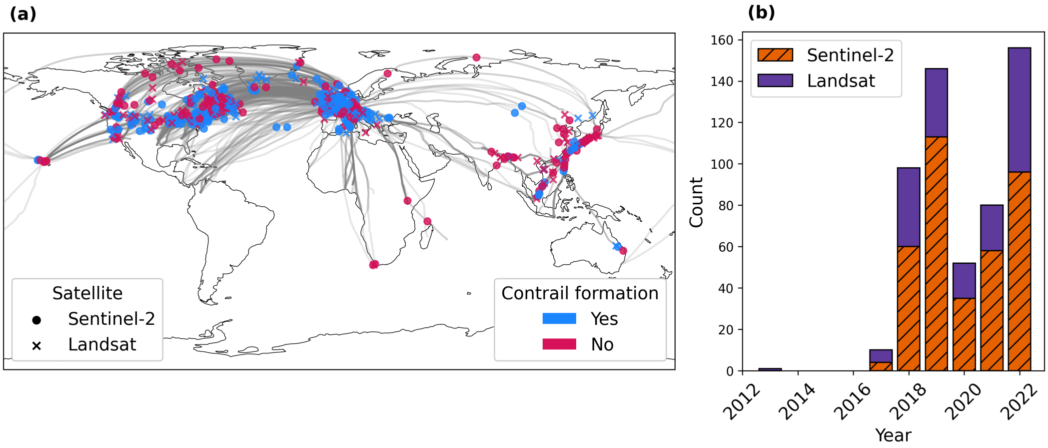

**Figure 4.** (a) Location of the 543 annotated collocations coloured by satellite. Grey lines indicate the flight paths of the collocations. (b) Histogram of scanning time of the collocations summed by year.

**Table 4.** Number of collocations for each aircraft within the IAGOS fleet, along with information on the aircraft and engine type. Aircraft without collocations and/or flights during the period of this study are included for completeness.

| IAGOS | Registration | Aircraft Type | Engine | BPR | Exhaust Type | Collocations | Total flights 2013-2022 |
|---|---|---|---|---|---|---|---|
| 1 | D-AIGT | A340-313 | 4x CFMI CFM56-5C4 | 6.4 | mixed | 137 | 3572 |
| 2 | B-18806 | A340-313 | 4x CFMI CFM56-5C4 | 6.4 | mixed | 1 | 2624 |
| 3 | F-GLZU | A340-313 | 4x CFMI CFM56-5C4 | 6.4 | mixed | 13 | 2645 |
| 4 | B-HLR | A330-343 | 2x RR Trent 772B-60 | 5.0 | mixed | 4 | 3643 |
| 5 | EC-GUQ | A340-313 | 4x CFMI CFM56-5C4 | 6.4 | mixed | 0 | 474 |
| 6 | D-AIKO | A330-343 | 2x RR Trent 772B-60 | 5.0 | mixed | 218 | 4851 |
| 7 | B-18317 | A330-302 | 2x GE CF6-80E1A4 | 5.1 | unmixed | 67 | 3494 |
| 8 | N384HA | A330-243 | 2x RR Trent 772B-60 | 5.0 | mixed | 55 | 1735 |
| 9 | F-GZCO | A330-203 | 2x GE CF6-80E1A3 | 5.1 | unmixed | 19 | 1212 |
| 10 | B-18316 | A330-302 | 2x GE CF6-80E1A4 | 5.1 | unmixed | 9 | 329 |
| 11 | D-AIKE | A330-343 | 2x RR Trent 772B-60 | 5.0 | mixed | 20 | 321 |
| 12 | C-GEFA | A330-343 | 2x RR Trent 772B-60 | 5.0 | mixed | 0 | 0 |
| 13 | EC-MSY | A330-202 | 2x GE CF6-80E1A4 | 5.1 | unmixed | 0 | 0 |

### 3.2 Evaluation of contrail formation predictions by the Schmidt-Appleman criterion

For each collocation within the dataset, we evaluate the Schmidt-Appleman criterion (SAC) using both IAGOS and ERA5 data and compare this to the corresponding annotation. One such comparison is given in Figure 5, for a case where no contrail

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

formation is observed. The engine efficiency used within the evaluation of the SAC is derived from the PS model using IAGOS data as input: this is used to construct the threshold mixing line in Figure 5. The shaded area represents the envelope of threshold mixing lines resulting from varying the engine efficiency by $\pm 15\%$. The ambient conditions obtained from ERA5 and IAGOS data are both shown in Figure 5. In this particular case, the SAC evaluated using ERA5 data incorrectly predicts that contrail formation should occur. When using the IAGOS data however, the SAC is not satisfied in agreement with the observation. We give three more of these examples in subsection A1, using both Sentinel-2 and Landsat imagery. The example provided in Figure 5 shows a difference in ERA5 and IAGOS-based air temperatures of 1.9 K, which is on the higher end of the examples in the dataset. Over the whole dataset, we found of a mean difference of -0.583 K ($T_{\mathrm{ERA5}} - T_{\mathrm{IAGOS}}$, with a standard deviation of 0.757 K.

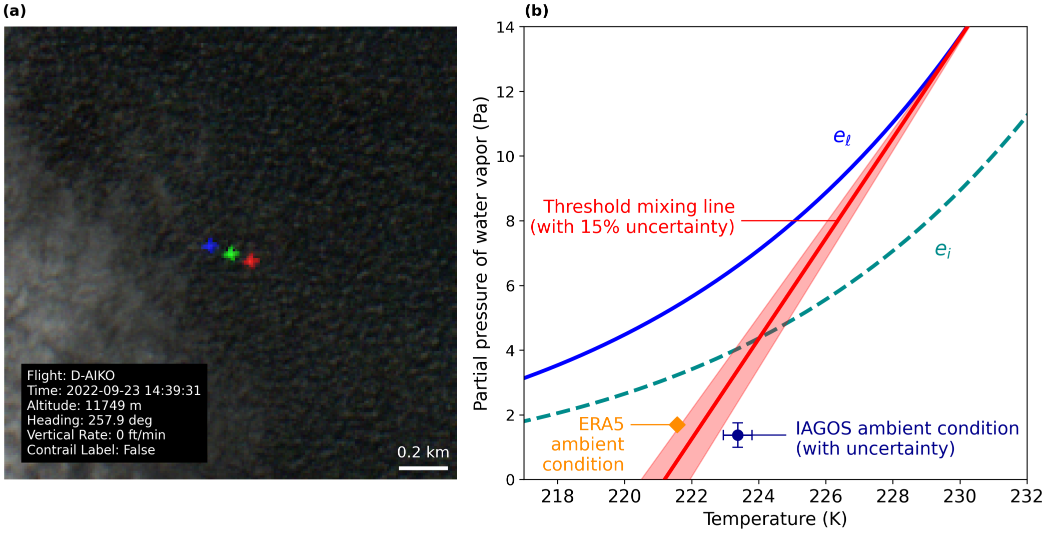

**Figure 5.** (a) Sentinel-2 true colour image showing the D-AIKO aircraft without a contrail. The rainbow effect (i.e. three differently coloured aircraft being visible) arises from the staggered configuration of the sensing instrument, which introduces a small time offset between the individual bands. This time delay results in a parallax difference between the bands. (b) Corresponding SAC diagram including saturation pressure w.r.t. liquid ($e_\ell$) and ice ($e_i$); the threshold mixing line resulting from the PS model engine efficiency; the ambient conditions from ERA5; and the IAGOS measurements with uncertainties.

The results from the evaluation of the SAC predictions using the entire dataset are given in Table 5 as a function of the satellite platform and meteorological data used. For the “IAGOS (within uncertainty)” meteorological data source, we consider the SAC prediction to be correct if any combination of temperature and relative humidity values within the 95% confidence intervals leads to a correct prediction (see the examples in Figure A1 and Figure A2). We now discuss the results in Table 5 by focusing on the the effect of observation quality, the effect of the meteorological data source and by comparing the results to the literature. Firstly, for identical weather data, observations based on Landsat exhibit a consistently lower precision than those based on Sentinel-2, with differences ranging from 0.048 to 0.084, while recall remains comparable between the two

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

**Table 5.** Summary of performance metrics for the SAC compared against the satellite observations for IAGOS and ERA5 weather data. Bold row indicates best performing dataset.

| Satellite | Meteorological Data | TP | FP | FN | TN | Precision | Recall |
|---|---|---|---|---|---|---|---|
| Landsat | ERA5 | 80 | 15 | 2 | 80 | 0.842 | 0.976 |
| Landsat | IAGOS | 78 | 11 | 4 | 84 | 0.876 | 0.951 |
| Landsat | IAGOS (within uncertainty) | 81 | 5 | 1 | 90 | 0.942 | 0.988 |
| Sentinel-2 | ERA5 | 192 | 23 | 3 | 148 | 0.893 | 0.985 |
| Sentinel-2 | IAGOS | 192 | 8 | 3 | 163 | 0.960 | 0.985 |
| **Sentinel-2** | **IAGOS (within uncertainty)** | 194 | 2 | 1 | 169 | **0.990** | **0.995** |

(differences $< 0.034$). The lower SAC precision for Landsat observations is driven by a relative increase in false positives, where the SAC is satisfied but no contrail is observed. This reduction in precision when using Landsat instead of Sentinel-2 observations is found to be statistically significant ($p$-value of 5%), and we hypothesise that the reduced resolution of the Landsat RGB imagery compared to the Sentinel-2 RGB imagery (30 $\mathrm{m}$ instead of 10 $\mathrm{m}$) leads to a larger amount of contrails that are unobservable. Examples of how a reduced image resolution complicates the contrail observation process can be found in Appendix A and Appendix D. To assess the possibility that the observed differences in precision and recall between the Landsat and the Sentinel-2 observations are due to sampling effects, we have performed Monte Carlo simulations (further explained in subsection A2). The results indicate that under the null hypothesis that the precision and recall have the same value for the two satellites, the probability of finding the observed differences in precision and recall are 0.005 and 0.06 respectively. In other words, the observed differences in precision and recall for the two different satellites are unlikely to be explainable by random sampling effects alone.

Secondly, we see that the use of IAGOS instead of ERA5 data leads to better agreement between the SAC predictions and observations for both satellite instruments. The precision increases by 0.067 for the Sentinel-2 observations and by 0.034 for those from Landsat. In the analysis of Gierens et al. (2020), the SAC was calculated using ERA5 and compared against SAC calculated using IAGOS, which was considered the ground truth. Their results indicate a precision of 0.923 and a recall of 0.992. When we consider the SAC computed using IAGOS data as the ground truth - rather than the satellite observations - we find a precision of 0.932 which is within 1% of the value found by Gierens et al. (2020).

Table 5 shows that the use of ERA5 data results in a larger number of false positives than when using IAGOS data as input for the SAC. One possible explanation for this is the reported mean cold bias of approximately 0.5 $\mathrm{K}$ in ERA5 temperatures (Wolf et al., 2025), which lowers $\Delta T_{\mathrm{LC}}$ and can increase the number of false positive predictions by the SAC. The increase in false positives leads to a decrease in precision, without affecting recall.

In addition, we evaluated the effect of using the relative humidity of the clear-sky portion of the ERA5 grid point, calculated from specific humidity and cloud cover fraction (Tompkins et al., 2007). While this did change the relative humidity in 19% of cases, it had no impact on the results of any of the collocations and thus no effect on the evaluation of the SAC.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

Lastly, we discuss the evaluation results where the SAC prediction is considered to be correct if any combination of overall engine efficiency, air temperature or relative humidity values within the 95% confidence intervals of IAGOS data leads to a correct prediction. This increases the precision to 0.990 (0.942) and the recall to 0.995 (0.988) for the Sentinel-2 (Landsat) observations. Thus, the SAC correctly predicts 99.2% of Sentinel-2 observations and 96.6% of Landsat observations when IAGOS measurement uncertainty is taken into account.

In other observational datasets, the level of agreement between observations and SAC predictions is found to be lower than that obtained using the dataset introduced here. Using the data from Low et al. (2025), we obtain a precision of 0.717 and an average recall of 0.982 across all waypoints. Similarly, using the dataset from Jarry et al. (2026), we obtain a precision of 0.444 and a recall of 0.973 when comparing annotated contrails with corresponding flight trajectories. In both ground-based studies, the reported precision is lower than that observed in the Landsat and Sentinel-2 results presented here. One contributing factor may be the spatial resolution of ground-based methodologies. At an altitude of 10 $\mathrm{km}$, the all-sky camera used in Jarry et al. (2026), has an effective pixel resolution of approximately 50–100 $\mathrm{m}$, which likely contributes to an increased number of false positives. However, the wide-angle camera used in Low et al. (2025), has a resolution at 10 $\mathrm{km}$ of approximately 5-10 $\mathrm{m}$, higher than both satellites. We thus further attribute the remaining discrepancy to reduced observation quality inherent to ground-based camera methods, including the presence of low-level clouds, or sun glare effects. Overall, the higher precision reported by Low et al. (2025) compared to Jarry et al. (2026) is consistent with the hypothesis that spatial resolution is the dominant factor in contrail formation detection, which would also explain the higher precision found for Sentinel-2 observations relative to Landsat.

### 3.2.1 Effects of uncertainty in ambient conditions and engine efficiency

In Figure 6a we show the collocations as a function of $\Delta T_{\mathrm{LC}}$ and altitude, coloured according to the outcome of comparing the SAC prediction with the corresponding observation. Figure 6b focuses on incorrect predictions within the cruise altitude region, and shows the 95% confidence interval of $T_{\mathrm{LC}}$ by using the IAGOS uncertainty for $T$ and RH$_\ell$ and the 15% overall engine efficiency uncertainty leading to the upper and lower bounds of $\Delta T_{\mathrm{LC}}$. We see that even when accounting for uncertainty, 2 (5) false positives and 1 (1) false negatives for the Sentinel-2 (Landsat) are still present (as also presented in Table 5). Interestingly, several false positives identified using Landsat occur at $\Delta T_{\mathrm{LC}}$ values between -3 and -4 $\mathrm{K}$, which are more negative than the false positives found using Sentinel-2. This could be explained by the lower spatial resolution of the Landsat imagery relative to Sentinel-2, complicating the observation of contrail formation.

Figure 6b indicates that the uncertainty in the $\Delta T_{\mathrm{LC}}$ is driven primarily by the overall engine efficiency uncertainty: which accounts for a change in $\Delta T_{\mathrm{LC}}$ of $\pm 0.66$ $\mathrm{K}$ on average for all false positives and false negatives. The second largest driver of uncertainty for these false positives and false negatives is the air temperature, with an average change in $\Delta T_{\mathrm{LC}}$ of $\pm 0.41$ $\mathrm{K}$ followed by RH$_\ell$ with $\pm 0.25$ $\mathrm{K}$.

Figure 6b shows that incorrect predictions are all located between GPS altitudes of 9 $\mathrm{km}$ and 13 $\mathrm{km}$, three of which are during the climb phase and the remainder during cruise. The spatial distribution of these erroneous predictions is consistent

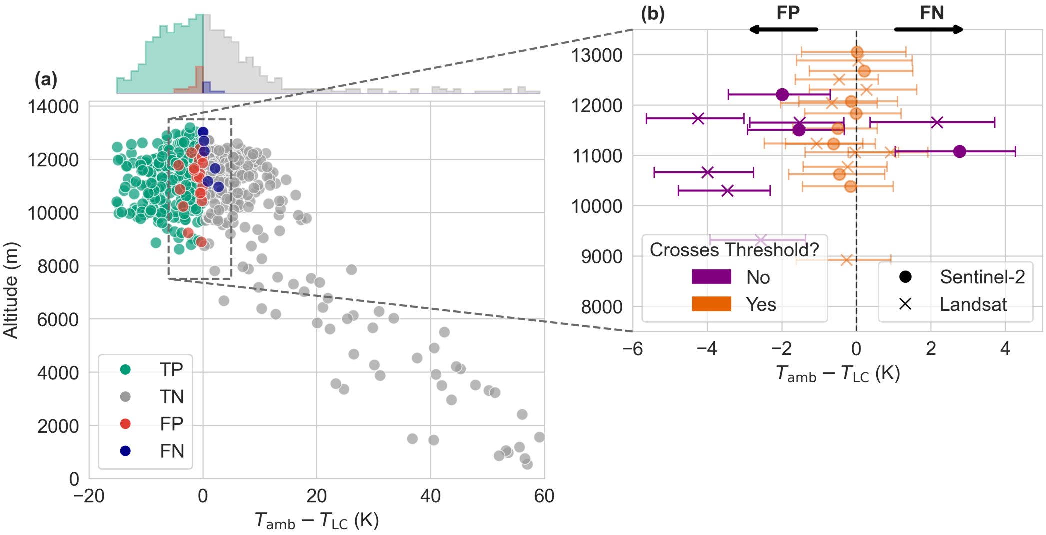

**Figure 6.** (a) All 543 collocations shown as a function of the temperature difference $\Delta T_{\mathrm{LC}}$ (derived from the IAGOS measurements) and altitude, coloured according to the outcome of comparing the SAC prediction with the corresponding observation. The histogram above the plot displays the marginal distribution of $\Delta T_{\mathrm{LC}}$ for each evaluation outcome. (b) Same as (a) but focusing on incorrect predictions in the cruise altitude region. Horizontal error bars represent 95% uncertainty from overall engine efficiency and IAGOS measurements. The altitude of each point is randomly changed slightly to avoid overlap and for better clarity. Points crossing the contrail formation threshold are shown in orange whereas those that do not cross the threshold, even when accounting for uncertainty are shown in purple. with lower opacity. Cross (circle) markers indicate Landsat (Sentinel-2) satellites.

with the overall spatial distribution of the dataset (not shown). The false positives (false negatives) are all found between 217.0 K (220.5 K) and 224.6 K (228.7 K), with an average temperature of 220.5 K (223.4 K).

## 3.3 Contrail evolution

### 3.3.1 Lifetimes of observed contrails

The estimated lifetimes (corresponding to the estimated age of the end of the centerline of the polygon annotation) for the 195 contrails observed in Sentinel-2 imagery are shown in Figure 7. We find that 48.2% of contrail lifetimes are longer than 10 s (approximately the duration of the jet phase) and 8.7% are longer than 120 s (approximately the end of the vortex phase). Contrails for which the annotation polygon ends at the satellite image swath edge, implying that the estimated lifetime is a lower bound on its actual lifetime, are shown using red markers in Figure 7.

Figure 7a shows contrail age as a function of the IAGOS $\mathrm{RH}_i$ value at the oldest point of the contrail, indicating that mean lifetimes increase with increasing $\mathrm{RH}_i$. A linear regression between $\mathrm{RH}_i$ and the logarithm of contrail lifetime yields

$\mathrm{R}^2 = 0.49$. A similar analysis using ERA5 $\mathrm{RH}_i$ produces $\mathrm{R}^2 = 0.51$, suggesting a comparable relationship between humidity and contrail persistence in both datasets.

Contrails with lifetimes shorter than 10 s have a mean $\mathrm{RH}_i$ of $0.40 \pm 0.22 (1\sigma)$ $(0.40 \pm 0.24)$ for IAGOS (ERA5) data. Those lasting between 10–120 s exhibit a mean $\mathrm{RH}_i$ of $0.76 \pm 0.32$ $(0.74 \pm 0.27)$, and contrails persisting longer than 120 s have a mean $\mathrm{RH}_i$ of $0.96 \pm 0.20$ $(0.99 \pm 0.14)$.

Contrails have been shown to persist in slightly ice-subsaturated air (Li et al., 2023; Petzold et al., 2025), which is consistent with our findings that contrails surviving past the vortex phase have a mean $\mathrm{RH}_i$ of $0.96 \pm 0.20$. Four contrails have been found at $\mathrm{RH}_i$ below 90%, which could also be attributable to the fact that the IAGOS temperature and relative humidity measurements are taken only at the moment the aircraft passes a location, whereas the observed contrail will have aged and been advected by the wind, leading to a spatial and temporal mismatch between the IAGOS measurements and the actual atmospheric conditions of the older points of the observed contrail.

Figure 7b shows contrail age against the $\Delta T_{\mathrm{LC}}$, where more negative values indicate a higher maximum supersaturation w.r.t. liquid achieved during the mixing process of the engine exhaust and the ambient air. We see from Figure 7b that generally, more negative values of $\Delta T_{\mathrm{LC}}$ indicate longer-lived contrails, however the trend is weaker than that of $\mathrm{RH}_i$. This is also visible in a lower $\mathrm{R}^2$ value 0.23 between the $\Delta T_{\mathrm{LC}}$ and the logarithm of the contrail age. In four cases, a contrail was observed while having a $\Delta T_{\mathrm{LC}}$ above 0. We have already established in Table 5, that with Sentinel-2 observations and IAGOS data, there are three observed contrails not predicted by the SAC. An additional reason in this case can be the spatial and temporal mismatch between the IAGOS measurements and the actual atmospheric conditions.

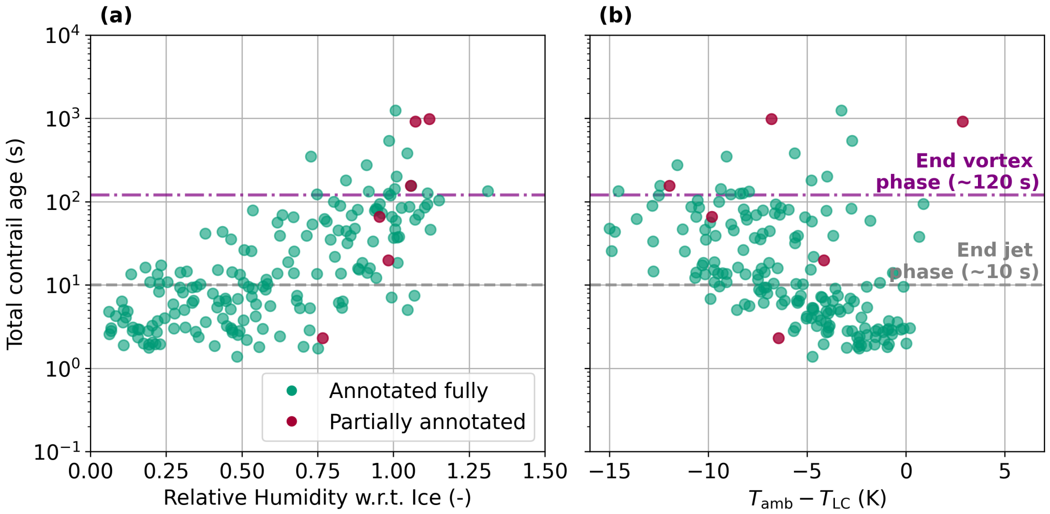

**Figure 7.** (a) Total contrail age in seconds against $\mathrm{RH}_i$ from IAGOS at the ”oldest” point of the contrail. Contrails that leave the satellite image are coloured "red", contrails that are fully annotated in the image are "green". (b) Total contrail age in seconds against $\Delta T_{\mathrm{LC}}$ from IAGOS at the ”oldest” point of the contrail.

### 3.3.2 Contrail width in the jet and vortex phases

We compute the contrail age and width using the methods described in subsection 2.7 and interpolate the contrail width to 1 s intervals such that we can determine the mean width and 95% confidence interval for every timestep. The resulting time series is shown in Figure 8. An overview of the evolution of the width individual contrails can be found in Appendix B. 505

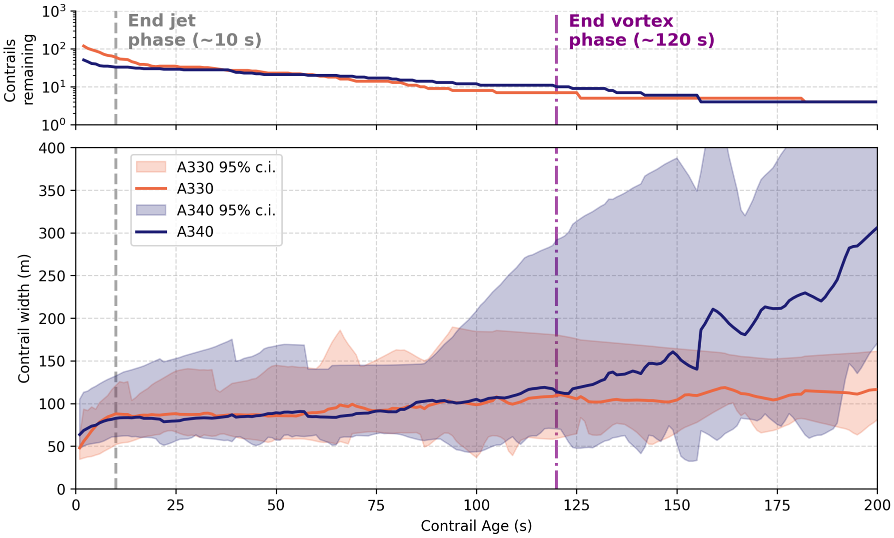

**Figure 8.** Statistics on the evolution of contrail width as a function of age. The number of surviving contrails as a function of age is given at the top. The approximate end of the jet phase (10 s) and end of the vortex phase (120 s) are indicated using dashed lines.

Figure 8 shows a jet and vortex width evolution pattern that is qualitatively consistent with previous observations (Hoshizaki et al., 1972) and with model simulations of the jet and vortex phases (Paoli and Shariff, 2016). The jet phase is characterised by continuous growth of the exhaust jets by turbulent mixing, followed by entrainment of the jets in the wake vortex pair after which mixing is inhibited and horizontal growth stagnates. For both aircraft types, the initial width in the jet phase is approximately 50 $\mathrm{m}$, comparable to the initial wake vortex separation $b_0$ of 47.3 $\mathrm{m}$ for both aircraft types . At this point in the contrail lifetime, the two vortex pairs are still visible in the satellite image as is shown in Figure 3a. Subsequently, the contrail enters the vortex phase, during which the width remains constant at values between 80 $\mathrm{m}$ and 100 $\mathrm{m}$. Although the lateral extent of the contrail does not change substantially, the initially separated vortex pair gradually merges into a single uniform trail (as viewed from the satellite perspective) which is shown in Figure 3b. Only 17 contrails survive for longer than 120 seconds, leading to an increasing spread in the observed width values beyond this point. Although the width of the A330 contrails is found to remain relatively constant after 120 seconds, the A340 contrails appear to resume growing beyond this point. However, the latter behaviour of the A340 contrails is dominated by a single contrail with significant horizontal growth: indicating the limited statistical significance of the width data beyond the vortex phase.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

The value for the contrail width $B_1$ at the end of the vortex phase predicted by CoCiP using Equation 6 is found to be approximately 20 $\mathrm{m}$ for both the A330 and A340 aircraft, assuming maximum take-off weight and $V_a = 250\ \mathrm{ms}^{-1}$. This initial width is in a similar range to the widths presented in Figure 4 from Schumann (2012), where the initial widths are reported to be between 10 and 30 $\mathrm{m}$. Comparing this to Figure 8, we conclude that the observed width is therefore a factor of 2-4 higher than that calculated in CoCiP. It should be noted however that the initial width has a relatively small impact on the evolution of the contrail within the dispersion phase modeled by CoCiP: the initial depth is considerably more important as it affects the rate of horizontal growth by vertical wind shear. The CoCiP implementation in *pycontrails* (Shapiro et al., 2026) uses the wake vortex separation as a starting point for the subsequent diffusion phase, which is closer to the values found here.

Remote-sensing based observations of contrail width have been reported previously (Freudenthaler et al., 1995; Sassen, 1997; Immler et al., 2008; Mannstein et al., 2010; Iwabuchi et al., 2012; Schumann et al., 2013; Low et al., 2025; Jarry et al., 2026), but primarily focus on contrails observed after the vortex phase. Alternatively, computational fluid dynamics (CFD) simulations, most commonly large-eddy simulations (LES), can also be used to derive the contrail width at various times in the jet and vortex phases. In Paoli et al. (2013), the jet phase of a contrail was simulated using LES for both two-engine and four-engine aircraft scenarios. While optical depths or contrail width are not specifically derived, the width can be derived from the spatial distribution of the simulated particles from the 2D cross-sections through the vortex center at 10 s and 20 s after formation. Based on this approach, the contrail width is approximately 70 $\mathrm{m}$ (90 $\mathrm{m}$) for the two-engine (four-engine) case at 10 $\mathrm{s}$, increasing to about 80 $\mathrm{m}$ (100 $\mathrm{m}$) at 20 $\mathrm{s}$. Similarly, Paugam et al. (2010) presents estimates for contrail width in the early dispersion phase. In particular, they find a contrail width between 100 and 150 $\mathrm{m}$ at 160s after formation. Both these results are comparable to the observed widths shown in Figure 8. Lewellen and Lewellen (2001) presents LES results for a B747 contrail during the first 300 $\mathrm{s}$ of its evolution. A top-down view reveals a nearly constant contrail width of roughly 100 $\mathrm{m}$ between 10 $\mathrm{s}$ and 100 $\mathrm{s}$, after which the contrail begins to expand horizontally, albeit with signs of wake vortex breakup.

### 3.3.3 Comparing the observed evolution of contrail width in the dispersion phase with CoCiP predictions

The small amount of contrails that survive the vortex phase and enter the dispersion regime in our dataset does not allow for robust statistical conclusions on the horizontal growth rates of contrails. However, we illustrate here how the observed evolution of a contrail can be compared to a model such as CoCiP for the three observed contrails that have lifetimes longer than 600 s. Future work can extend our analysis of observed contrails to aircraft outside of the IAGOS fleet, in order to apply this type of analysis to a much larger dataset.

The horizontal growth of the three oldest contrails in our dataset is shown in Figure 9 together with the simulations from CoCiP. Note that all the three contrails have a mean IAGOS $\mathrm{RH}_i$ above 100%. Despite the limited sample size, the simulated contrail widths exhibit good overall agreement with the observations, as the average dispersion rate in CoCiP differs from the observed values by at most 38%. Nevertheless, the small number of available cases limits the statistical robustness of this comparison.

Our contrail dispersion results are consistent with observations from the literature. For example, Freudenthaler et al. (1995) report ground-based LIDAR observations of contrails having width growth rates between 18 and 140 $\mathrm{m\ min^{-1}}$, which aligns well with the three contrail cases shown in Figure 9, which exhibit growth rates of approximately 102, 56, and 49 $\mathrm{m\ min^{-1}}$.

Additional Sentinel-2 and Landsat observations covering a wider range of atmospheric conditions, aircraft and engine types 555 as well as contrail lifetimes, would therefore be valuable to enable a more quantitative validation of dispersion rates in CoCiP and other contrail evolution models.

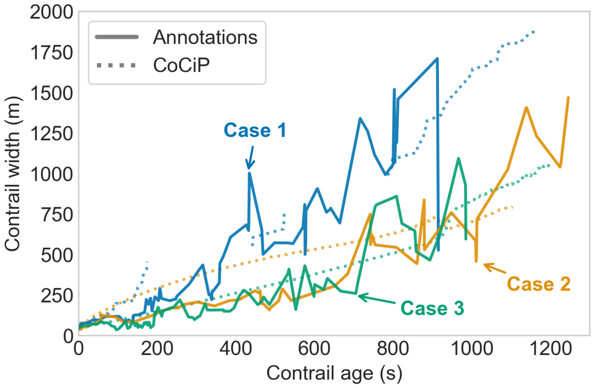

**Figure 9.** Evolution of contrail width (m) over contrail age (s) for the three cases that persist longer than 10 min. Different colours indicate the individual cases. Each CoCiP simulation used 1 s timesteps and a maximum contrail age of 20 min. Relative humidities ($\mathrm{RH}_\ell$ and $\mathrm{RH}_i$) from ERA5 were scaled with IAGOS measurements. For Case 1, the simulated contrail is interrupted by regions of sublimation, visualised by breaks in the line.

## 4 Conclusions

We have presented a method for locating flying aircraft in high-resolution low Earth orbit (LEO) satellite imagery from Sentinel-2 and Landsat. We applied this method to the IAGOS aircraft fleet in order to create a dataset of observations of 560 contrail formation and early evolution combined with in-situ measurements of temperature and humidity. The resulting dataset has been used for the evaluation of model predictions of contrail formation by the SAC, and horizontal growth by CoCiP.

The final dataset consists of 543 collocated IAGOS aircraft, in which 277 cases of contrail formation were observed. Out of these 277 contrails, of which 195 were annotated, 48.2% were found to persist for longer than 10 s (approximately the jet phase) and 8.7% longer than 120 s (approximately the vortex phase). The maximum contrail age identified in the dataset is 565 1250 s. Relative humidity with respect to ice was found to correlate strongly with contrail lifetime, exhibiting a $\mathrm{R}^2$ value of 0.49 with the logarithm of contrail age. Contrails with lifetimes longer than 120 s showed a mean $\mathrm{RH}_i$ of $0.96 \pm 0.20$, with four contrails found at $\mathrm{RH}_i$ below 90% suggesting that contrails can survive the vortex phase in ice-subsaturated air.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

A comparison of observed contrail formation with predictions indicates that the SAC can explain 98.3% of the observations when using temperature and relative humidity measured by IAGOS, when accounting for measurement uncertainty. When using meteorological data from the ERA5 reanalysis as input, we find that only 92.1% of observations agree with predictions by the SAC. This reduced agreement between observation and predictions illustrates the importance of accurate input data when comparing models and observations, in order to isolate model errors from those present in the inputs. Additionally, we find that for the same meteorological data source (e.g. IAGOS) the agreement between SAC predictions and Landsat observations is consistently lower than when comparing against Sentinel-2 data. Specifically, the precision of the SAC is found to be lower for Landsat observations, whereas recall remains approximately equal. This is caused by more frequent predictions of contrail formation by the SAC where Landsat imagery shows no identifiable contrail. We attribute this difference primarily to the coarser spatial resolution of Landsat RGB imagery ($30\ \mathrm{m}$) leading to a larger fraction of contrails that are not detectable compared to the Sentinel-2 imagery ($10\ \mathrm{m}$). All together, these results emphasise the important role played by limitations in meteorological input data and observations when evaluating contrail models. We therefore hypothesise that lower levels of agreement between SAC predictions and observations reported in previous studies may have been caused by inaccurate NWP data and/or observations.

Furthermore, the dataset presented here contains 195 annotated contrails which can be used to study the early evolution of contrails. We used these annotations to analyze the evolution in contrail width as a function of age, and compared these to previous observations as well as CoCiP model predictions. During the first $10\ \mathrm{s}$ (approximately jet phase) we observe rapid increases in contrail width as the turbulent engine exhaust jets mix with ambient air. This growth levels off near $10\ \mathrm{s}$ as the exhaust jets become entrained in the wake vortex pair, and contrail width is observed to remain approximately constant until about $120\ \mathrm{s}$ (approximately end of the vortex phase) after which horizontal growth is resumed, presumably driven by turbulence and vertical wind shear. These observed patterns are consistent with previous observational and modeling studies. The limited number of contrails found to persist for longer than 10 minutes does not allow for robust statistical conclusion on the rate of contrail growth in the dispersion phase. However, a qualitative comparison of the growth of the three oldest contrails in the dataset with simulations by CoCiP shows that the dispersion rate in the latter model differs from the observations by at most 38%.

Although the results of this study suggest that the SAC can predict contrail formation with a high level of skill (that increases with the quality of the meteorological input data used), we note that the dataset used covers only a limited number of aircraft-engine combinations. Specifically, the dataset consists of only 10 different aircraft, which are all part of two wide-body aircraft families and are equipped with three different engine families. Future work can utilise the collocation methodology presented here, possibly augmented by automated contrail detection techniques, to increase both the size and coverage of the dataset. This will enable the evaluation of the SAC for a broader set of aircraft and engine combinations, as well as study the early evolution of contrails formed by different aircraft. Another avenue of future work is to infer additional information on contrail microphysical properties as well as wake vortex pair instabilities and decay from the satellite imagery, which we have not attempted here.

*Code and data availability.* The dataset of IAGOS measurements, observations and metadata used in this study are publicly available at Woldhuis et al. (2026a). The IAGOS data can be downloaded from the IAGOS data portal at https://doi.org/10.25326/06 (Boulanger et al., 2020). The code to process this data and reproduce the figures are available at Woldhuis et al. (2026b). Some of the results in this paper were produced with the publicly available *pycontrails v0.60.2* library (Shapiro et al., 2026), to which the collocation algorithm has been added.

## Appendix A: Additional results for the Schmidt-Appleman criterion analysis

### A1 Additional collocations with the Schmidt–Appleman criterion

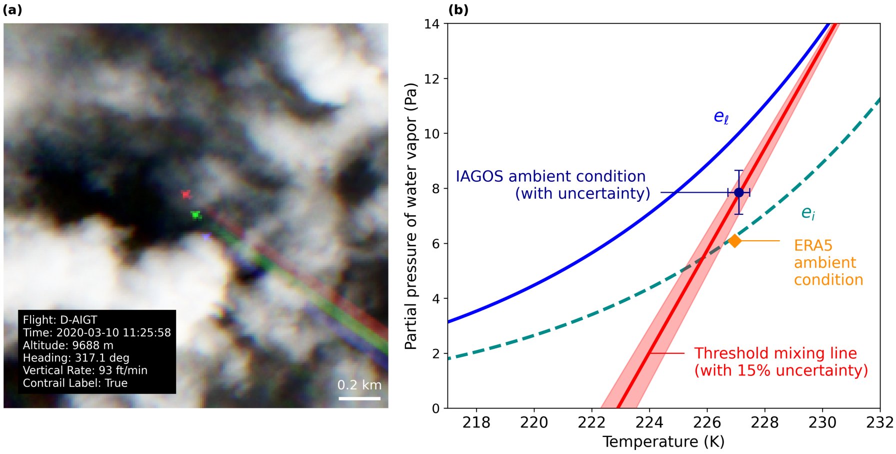

**Figure A1.** (a) Sentinel-2 true colour image showing the D-AIGT aircraft forming a contrail. The rainbow effect (i.e. three differently coloured aircraft being visible) arises from the staggered configuration of the sensing instrument, which introduces a small time offset between the individual bands. This time delay results in a parallax difference between the bands. (b) Corresponding SAC diagram including the threshold mixing line resulting from the PS model engine efficiency, the ambient conditions from ERA5, and the IAGOS measurements with uncertainties.

### A2 Statistical robustness of SAC precision and recall differences

We use a Monte Carlo simulation to assess the possibility that the observed differences in precision and recall between the Landsat and Sentinel-2 observations are a sampling artifact resulting from a small dataset size. We combine the contingency table entries obtained for the Landsat and Sentinel-2 observations (see Table 5) to compute the base rate $s$, recall Re and false alarm rate $F$. The expressions for the base and false alarm rate are

$$s = \frac{\mathrm{TP} + \mathrm{FN}}{n_{\mathrm{landsat}} + n_{\mathrm{sentinel}}}, \qquad F = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{A1}$$

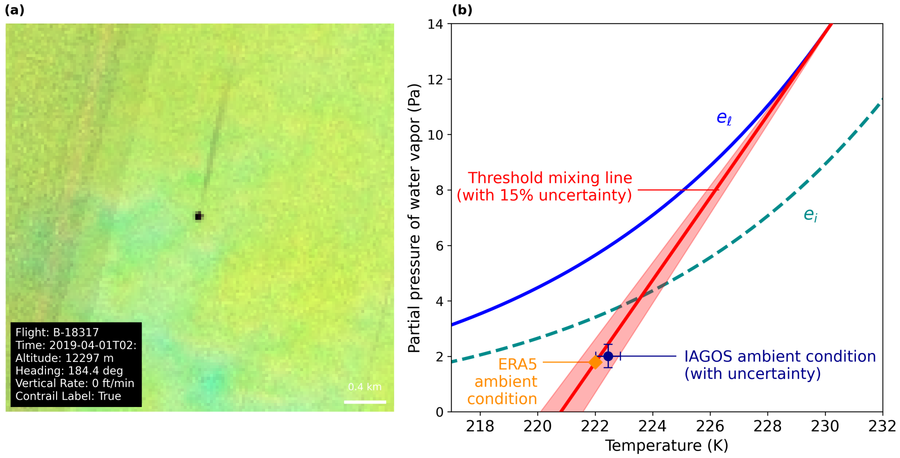

**Figure A2.** Same as Figure A1, but for a Landsat false colour composite.

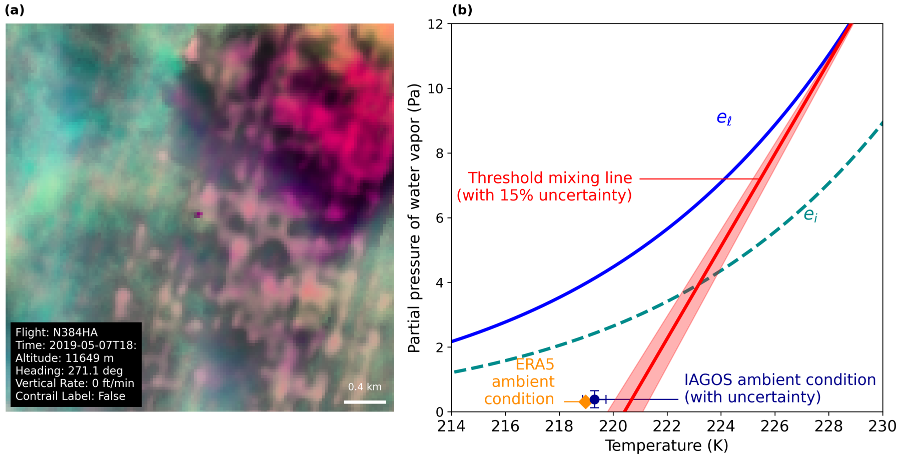

**Figure A3.** Same as Figure A1, but for a Landsat false colour composite.

These three parameters $s$, Re, $F$, calculated as 0.51, 0.97, and 0.07 respectively, define the parameters necessary for generating a distribution of precision and recall values under the null hypothesis that both datasets are random samples drawn from the same underlying process.

Using the baseline parameters $(s, H, F)$, we generate $m = 10000$ samples of size $n$, where $n$ corresponds to the sample sizes of the Landsat and Sentinel-2 datasets. From each synthetic contingency matrix, we compute precision and recall as

$$\widehat{\mathrm{Pr}}_i = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FP}_i}, \tag{A2}$$

$$\widehat{\mathrm{Re}}_i = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i}. \tag{A3}$$

This yields empirical distributions of $\widehat{\mathrm{Pr}}$ and $\widehat{\mathrm{Re}}$ for both the Landsat and Sentinel-2 sample sizes, representing the expected variability in precision and recall due to random sampling alone. For each realization, we compute the differences in performance metrics between the two datasets with

$$\Delta\widehat{\mathrm{Pr}}_i = \widehat{\mathrm{Pr}}^{(i)}_{\mathrm{sentinel}} - \widehat{\mathrm{Pr}}^{(i)}_{\mathrm{landsat}}, \tag{A4}$$

$$\Delta\widehat{\mathrm{Re}}_i = \widehat{\mathrm{Re}}^{(i)}_{\mathrm{sentinel}} - \widehat{\mathrm{Re}}^{(i)}_{\mathrm{landsat}}. \tag{A5}$$

These samples define empirical distributions of precision and recall differences under the null hypothesis of identical underlying performance. We denote the observed differences in precision and recall between the Sentinel-2 and Landsat as $\Delta\mathrm{Pr}_{\mathrm{obs}}$ and $\Delta\mathrm{Re}_{\mathrm{obs}}$, which we can read from Table 5 to be respectively: -0.084 and -0.034. The probability that the observed difference arises from sampling variability alone is estimated as

$$P_{\mathrm{Pr}} = \mathbb{P}\left(\Delta\widehat{\mathrm{Pr}} \leq \Delta\mathrm{Pr}_{\mathrm{obs}}\right), \tag{A6}$$

$$P_{\mathrm{Re}} = \mathbb{P}\left(\Delta\widehat{\mathrm{Re}} \leq \Delta\mathrm{Re}_{\mathrm{obs}}\right), \tag{A7}$$

where the probabilities are evaluated using the empirical cumulative distribution functions of the samples.

We find for the two datasets, evaluated with IAGOS, a probability $P_{\mathrm{Pr}} = 0.005$ or $P_{\mathrm{Re}} = 0.060$ that the observed differences are explainable by random sampling along. In other words, this means that the observed difference in precision and recall are unlikely to be explainable by random sampling effects alone.

### A3 Aircraft performance model comparison

We used the BADA 4.2 and Poll–Schumann implementation in *pycontrails*. Not all aircraft types present in the IAGOS dataset (see Table 4) are explicitly represented in the BADA 4.2 files. For aircraft types without a direct match, a closely related aircraft type available in BADA 4.2 was used. For the Poll–Schumann model, aircraft were identified using their ICAO codes. The mapping between IAGOS aircraft types, the corresponding BADA 4.2 types, and the Poll–Schumann ICAO codes is shown in Table A1.

### A3.1 Overall engine efficiency

In Figure A4 we compare the overall engine efficiency, for both BADA 4.2 and PS, for all 543 IAGOS collocations. We can split the overall engine efficiencies into the three flight phases where climb corresponds to vertical rates greater than 100 $\mathrm{ft/min}$, cruise to absolute vertical rates below 100 $\mathrm{ft/min}$, and descent to vertical rates less than -100 $\mathrm{ft/min}$. The average values of the computed overall engine efficiencies for each phase and for each aircraft are shown in Table A2.

We see from this comparison that generally, during the climb and cruise phases, the overall engine efficiencies correlate well, with a root-mean-squared error of 0.0322, 0.0326, 0.0176, and a correlation coefficient $r$ of 0.900, 0.609, and 0.515, for

**Table A1.** Mapping between IAGOS aircraft types and corresponding BADA 4.2 and Poll–Schumann aircraft representations.

| **IAGOS** | **BADA 4.2** | **PS** |
|---|---|---|
| A330-202 | A330-203 | A332 |
| A330-203 | A330-203 | A332 |
| A330-243 | A330-243 | A332 |
| A330-302 | A330-301 | A330 |
| A330-343 | A330-341 | A330 |
| A340-313 | A340-313 | A343 |

the A332, A333, and A343 respectively. For these two phases, the largest difference is found for the A333, where BADA 4.2 650 computes overall engine efficiencies that are on average approximately 0.03 lower than PS. However, during the descent phase, the difference between the two models is larger since the *pycontrails* implementation of BADA 4.2 computes an overall engine efficiency of 0 during descent. In contrast, PS computed efficiencies between 0.1 and 0.3 for these cases.

Figure A5 gives an alternative visualisation of this difference. Here, the overall engine efficiencies for BADA 4.2 and PS are shown agains the vertical rate. At negative vertical rates, BADA 4.2 does consistently compute the overall engine efficiency as 655 0. Overall, as seen from the marginal distribution on the right, PS computes higher engine efficiencies than BADA 4.2, with on PS computing engine efficiencies 0.01 to 0.02 higher than BADA 4.2.

**Table A2.** Comparison of mean climb, cruise, and descent efficiencies from PS and BADA 4.2 APMs for collocations in IAGOS dataset using IAGOS weather data. Climb corresponds to vertical rates greater than 100 $\mathrm{ft/min}$, cruise to vertical rates with absolute values lower than 100 $\mathrm{ft/min}$, and descent to vertical rates less than -100 $\mathrm{ft/min}$.

| | A332 | A333 | A343 |
|---|---|---|---|
| PS Climb | 0.291 | 0.315 | 0.307 |
| PS Cruise | 0.311 | 0.337 | 0.323 |
| PS Descent | 0.133 | 0.215 | 0.152 |
| BADA 4.2 Climb | 0.285 | 0.275 | 0.289 |
| BADA 4.2 Cruise | 0.312 | 0.314 | 0.313 |
| BADA 4.2 Descent | 0.000 | 0.062 | 0.000 |

### A3.2 Influence on predictions using the Schmidt-Appleman criterion

We present the Schmidt-Appleman criterion (SAC) evaluation using BADA 4.2 engine efficiencies in Table A3. Compared to the PS results shown in Table 5, precision either increases slightly (by 0.01 to 0.02) or remains constant, while recall decreases

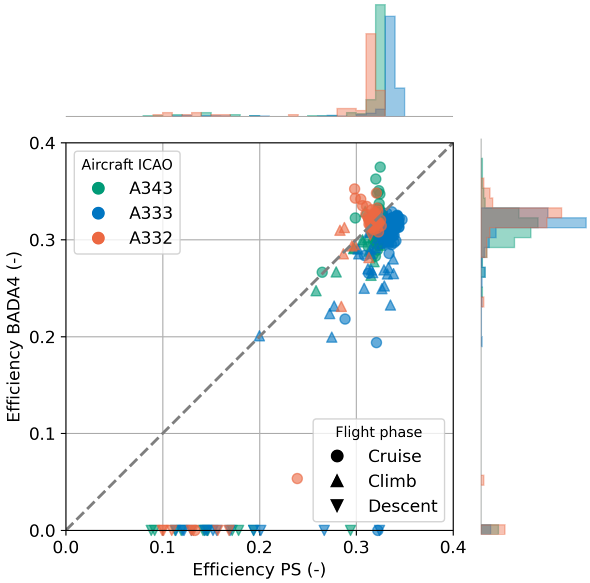

**Figure A4.** Scatter plot of overall engine efficiency in Poll-Schumann and BADA 4.2 at the scanning time of the satellite. Colour indicates IAGOS aircraft type, marker indicates flight phase. Histograms show the distribution of engine efficiencies shown with values scaled per aircraft type.

660 in all cases by approximately 0.01 to 0.02. Both differences are consistent with BADA 4.2 underestimating engine efficiency relative to PS, which results in fewer contrail predictions by the SAC. This reduces the number of false positive predictions.

**Table A3.** Summary of performance metrics for the SAC compared against the satellite observations for IAGOS and ERA5 weather data. SAC computed with BADA 4.2 overall engine efficiencies. Bold row indicates best performing dataset.

| Satellite | Weather Data | TP | FP | FN | TN | Precision | Recall |
|---|---|---|---|---|---|---|---|
| Landsat | ERA5 | 78 | 12 | 4 | 82 | 0.867 | 0.951 |
| Landsat | IAGOS | 76 | 9 | 6 | 85 | 0.894 | 0.927 |
| Landsat | IAGOS (within uncertainty) | 80 | 5 | 2 | 89 | 0.941 | 0.976 |
| Sentinel | ERA5 | 188 | 21 | 7 | 150 | 0.900 | 0.964 |
| Sentinel | IAGOS | 187 | 7 | 8 | 164 | 0.964 | 0.959 |
| **Sentinel** | **IAGOS (within uncertainty)** | 193 | 1 | 2 | 170 | **0.995** | **0.990** |

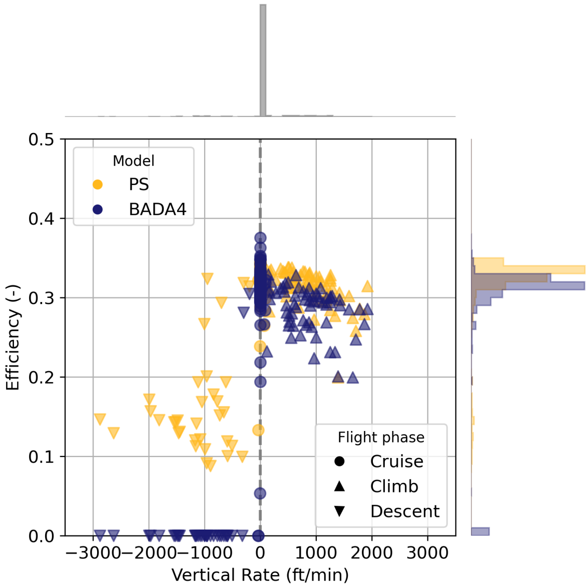

**Figure A5.** Scatter plot of overall engine efficiencies against the vertical rate at the scanning time of the satellite, coloured according to the model used. The histograms on the vertical axis show the distribution of efficiencies per model, and those on the horizontal axis show the distribution of vertical rates.

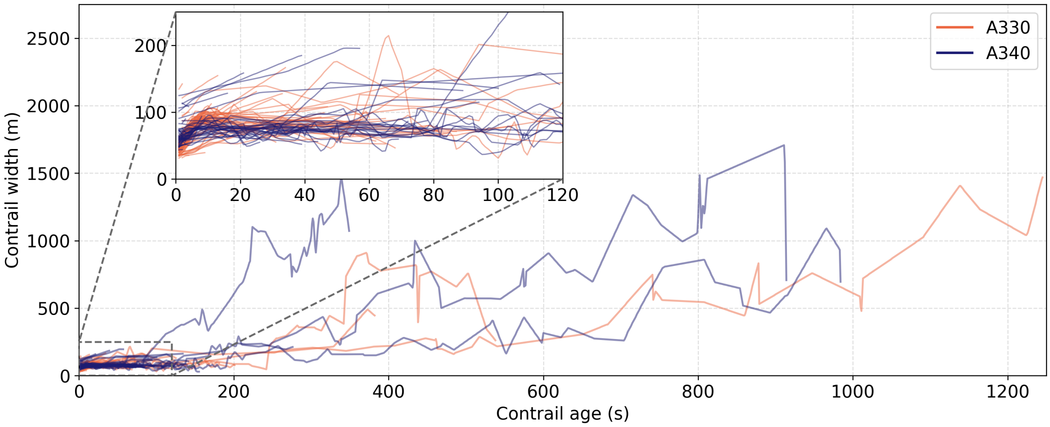

**Figure B1.** Evolution of contrail width as a function of age for all annotated contrails. The jet and vortex phase (first 120 s) are further visualised in the inset. Colours indicate aircraft family.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

## Appendix B: Additional results for contrail width evolution

## Appendix C: Additional details for the collocation process

### C1 Satellite tile search

We used *pycontrails* (Shapiro et al., 2026) to identify Sentinel-2 and Landsat tiles containing IAGOS aircraft. This implementation uses a Google BigQuery table of Sentinel-2 satellite tiles, which provides the bounding box and nominal scanning time of the tile. For Landsat scenes, we query the USGS Landsat Collection 2 Bulk Metadata Service. For each available IAGOS flight trajectory, we then use these bounding boxes and the nominal scanning time to ascertain whether any trajectory waypoints might be found within an image. In cases where tiles overlap near the edges, only a single tile is kept to avoid duplication.

### C2 Sentinel-2 Metadata and Processing Steps

For the parallax and scanning time corrections (described in subsection C4) we require viewing angle data at pixel-level ($10\ \mathrm{m}$) resolution as well as the satellite ephemeris and the detector time offsets.

We extract viewing angles (zenith and azimuth) from the tile metadata file (MTD_TL.xml, (Copernicus, 2024)) for every Sentinel-2 tile. These angles are provided at a coarse resolution of $25 \times 25\ \mathrm{km}$. The zenith angles are independent of the detector used to scan a pixel, meaning that we can interpolate them directly to the higher pixel-level resolution using bilinear interpolation. However, the azimuth angles exhibit staggered discontinuities at the detector borders (as shown in Figure C1a) which need to be taken into account when interpolating this data. We utilise the pixel-level detector footprint mask provided in the Quality Indicator folder (QI_DATA, (Copernicus, 2024)) to ascertain which detector captured each pixel (see Figure C1b). We then separate the coarse azimuth values by detector, bilinearly interpolate each detector's azimuth angles, and subsequently merge them using the detector mask. The resulting high-resolution azimuth viewing angles are shown in Figure C1c.

The ephemeris data is available in the Satellite Ancillary Data in the Datastrip metadata file (MTD_DS.xml, (Copernicus, 2024)). The ephemeris contains the dated positions ($x$, $y$, and $z$) in Earth-Centered, Earth-Fixed (ECEF) coordinates in $\mathrm{mm}$, along with the GPS time. The GPS time has been converted to UTC, using an 18 s offset (Seeber, 1993).

The detectors in the Sentinel-2 MSI alternate between pointing forward and backwards. For each satellite image, the Sensor Configuration metadata (Copernicus, 2024) provides the syncing times for each detector. The time delay per detector $\Delta t_{\text{detector}}$ is found by computing the difference between the individual detector's syncing time and the average syncing time of all detectors.

### C3 Landsat Metadata and Processing Steps

After determining relevant Landsat scenes intersecting IAGOS aircraft, we use the USGS/EROS Machine-to-Machine API (U.S. Geological Survey) to download scene imagery, per-pixel viewing angle data, and other necessary metadata to compute aircraft collocations. Table Table C1 shows the exact request.

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

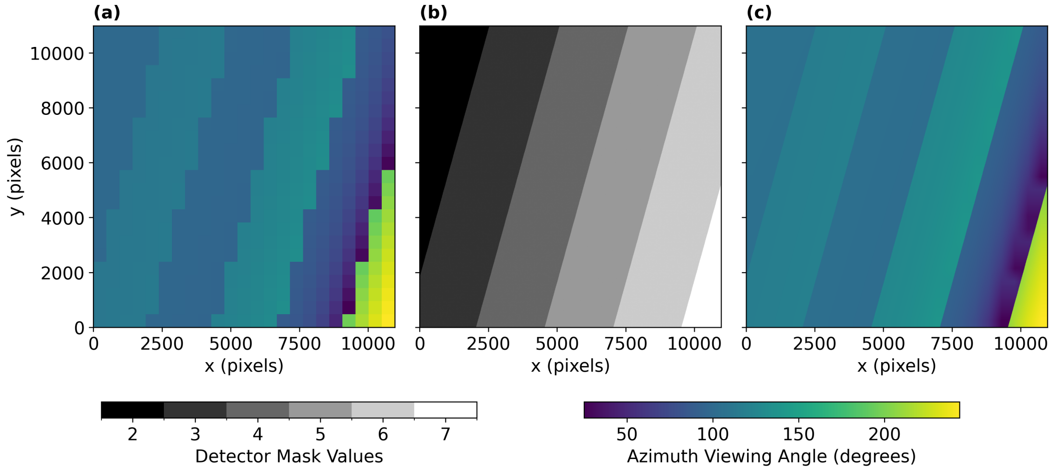

**Figure C1.** (a) Viewing azimuth angles for a single Sentinel-2 satellite image. (b) Detector pixel mask. (c) Viewing azimuth angles interpolated to pixel-level resolution.

**Table C1.** Files retrieved via the USGS/EROS M2M API for each candidate Landsat scene.

| File Name Suffix | Description | Purpose |
|---|---|---|
| MTL.TXT | Scene metadata | Radiometric rescaling factors |
| ANG.TXT | Angle metadata | Ephemeris waypoints |
| B2.TIF | Blue band | True colour RGB |
| B3.TIF | Green band | True colour RGB |
| B4.TIF | Red band | True colour RGB |
| B9.TIF | Cirrus band | False colour RGB |
| B10.TIF | Thermal infrared (11 $\mu$m) | False colour RGB |
| B11.TIF | Thermal infrared (12 $\mu$m) | False colour RGB |
| VAA.TIF | View Azimuth Angle | Per-pixel viewing angle |
| VZA.TIF | View Zenith Angle | Per-pixel viewing angle |

The per-pixel viewing zenith and azimuth angles (VZA and VAA) are provided at native spatial resolution and therefore do not require interpolation. The satellite ephemeris positions contained in the metadata are converted from ECEF coordinates to the scene's UTM coordinate reference system, and the satellite ground-track direction is estimated from consecutive waypoints. Unlike Sentinel-2, Landsat products do not explicitly provide the detector identifier for each pixel. To approximate detector membership, we leverage the view azimuth geometry: pixels are classified as forward- or aft-looking relative to the mean ground-track direction, thereby separating the focal plane into even and odd detector groups. A fixed temporal offset of $\pm 2$ s is then assigned according to this classification to approximate the scanning time offset $\Delta t_{\text{detector}}$ (see subsection C4).

https://doi.org/10.5194/egusphere-2026-1171
Preprint. Discussion started: 25 June 2026

## C4 Collocation Algorithm

We first transform the flight trajectory coordinates from IAGOS to the local UTM coordinate system used by the relevant Sentinel-2 or Landsat tile. Then, each waypoint ($x_{\text{UTM}}$, $y_{\text{UTM}}$, and $z$) at $t_{\text{flight}}$ is corrected for parallax displacement using an iterative approach starting with ($x_{\text{UTM}}$, $y_{\text{UTM}}$). The position is updated according to

$$x_{i+1} = x_i - z \cdot \tan\left(\text{VZA}(x_i, y_i)\right) \cdot \sin\left(\text{VAA}(x_i, y_i)\right), \tag{C1}$$

$$y_{i+1} = y_i - z \cdot \tan\left(\text{VZA}(x_i, y_i)\right) \cdot \cos\left(\text{VAA}(x_i, y_i)\right), \tag{C2}$$

where VZA denotes the viewing zenith angle, VAA the viewing azimuth angle and $i$ the iteration index. An iterative approach is necessary because as the position $(x_i, y_i)$ changes, the zenith and azimuth angles change as well. This has an especially large effect near the detector boundaries since the azimuth angle changes discontinously across the detector border (see Figure C1c). In this study, five iterations were used. The result of this iterative process at time $t_{\text{flight}}$ is denoted as ($x_{\text{proj}}(t_{\text{flight}})$, $y_{\text{proj}}(t_{\text{flight}})$) in what follows.

To estimate the scanning time, we first transform the satellite ephemeris into the UTM coordinate system of the satellite image. This produces a satellite ground-track ($x_{\text{sat}}(t_{\text{scan}})$, $y_{\text{sat}}(t_{\text{scan}})$). For each pixel with coordinates ($x_{\text{proj}}$, $y_{\text{proj}}$) a line perpendicular to the satellite ground track is drawn. The intersection point of this perpendicular line with the ground track is the location on the ground-track that is closest to this pixel. The corresponding scanning time $t_{\text{scan}}$ is then obtained as the time at which the ground-track position matches this point.

The detector-dependent time offset is obtained by identifying the detector corresponding to the pixel location ($x_{\text{proj}}$, $y_{\text{proj}}$). The scanning time offset $\Delta t_{\text{detector}}$ for Landsat (see subsection C3) or Sentinel-2 (see subsection C2) related to that detector is added to the previously estimated scanning times

$$t_{\text{scan}} \leftarrow t_{\text{scan}} + \Delta t_{\text{detector}}. \tag{C3}$$

We now have both the flight trajectory projected onto the ground surface ($x_{\text{proj}}(t_{\text{flight}})$, $y_{\text{proj}}(t_{\text{flight}})$), as well as the hypothetical times at which the projected trajectory would have been seen by the satellite ($x_{\text{proj}}(t_{\text{scan}})$, $y_{\text{proj}}(t_{\text{scan}})$). To determine the intersection point between the projected flight track and the satellite scanning time, the time difference

$$\Delta t = t_{\text{flight}} - t_{\text{scan}}, \tag{C4}$$

is computed. The exact intersection time, $t_{\text{col}}$, is then obtained by identifying the $t_{\text{flight}}$ where $\Delta t = 0$. The corresponding collocation location ($x_{\text{col}}$, $y_{\text{col}}$) is then computed by linearly interpolating ($x_{\text{proj}}(t_{\text{flight}})$, $y_{\text{proj}}(t_{\text{flight}})$) to $t_{\text{flight}} = t_{\text{col}}$.

## C5 Landsat collocation error

We present the collocation errors for Sentinel-2 in Figure 2, decomposed into components normal and tangential to the flight path. Similarly, collocation errors derived from Landsat-based metadata are shown in Figure C2. The analysis is based on approximately 1500 ADS-B flights for which the aircraft position was manually annotated and compared against the predicted locations. The corrected location predictions estimate the position of the aircraft within a distance of 621 $\mathrm{m}$ in 95% of the cases, compared to 3.523 $\mathrm{km}$ without corrections.

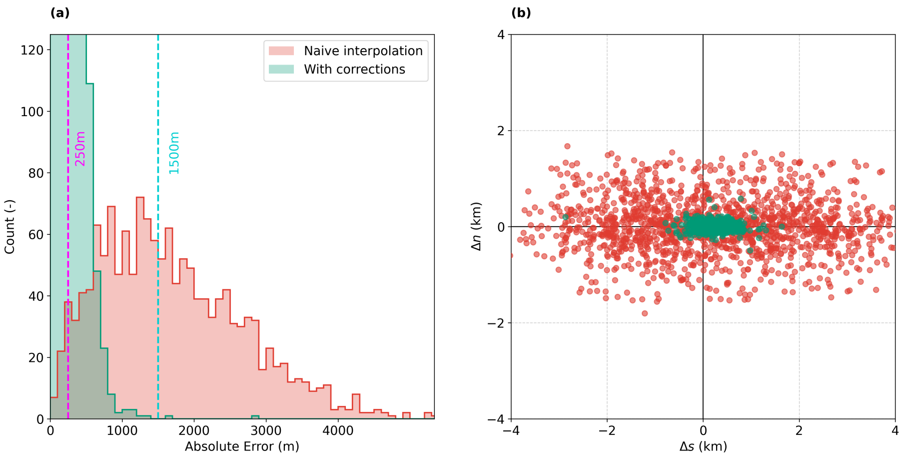

**Figure C2.** (a) Histogram of collocation errors for naive interpolation (red) and after scanning time and parallax corrections (green). The two vertical dotted lines indicate reference errors for 250 $\mathrm{m}$ (magenta) and 1500 $\mathrm{m}$ (cyan). (b) Scatterplot showing the errors normal, $\Delta n$, and tangential, $\Delta s$, to the flight path from the predicted positions to the manually annotated locations. Interpretation of direction of errors can be seen in Figure 2.

## Appendix D: Sample Sentinel-2 and Landsat Imagery

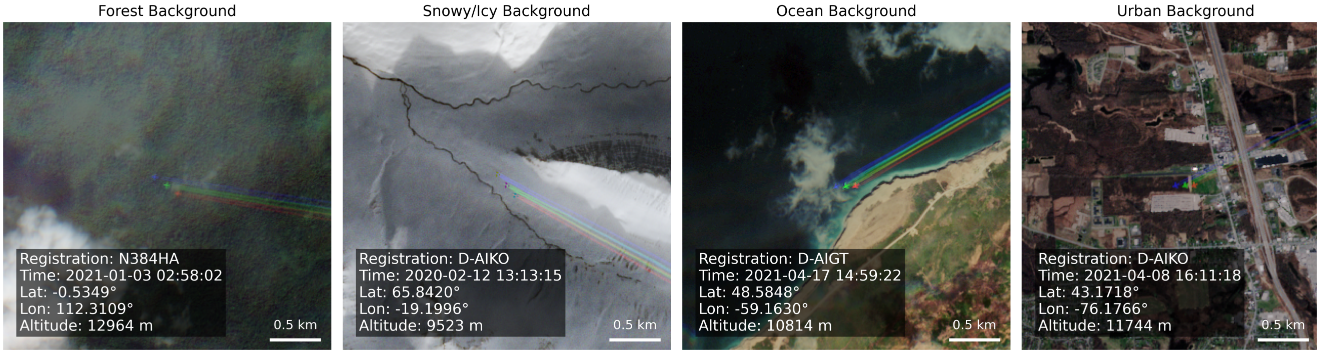

**Figure D1.** Examples of contrail-forming IAGOS aircraft in Sentinel-2 imagery with different background conditions. The aircraft, sensing time, location and altitude, as well as the scale, are provided for each example.

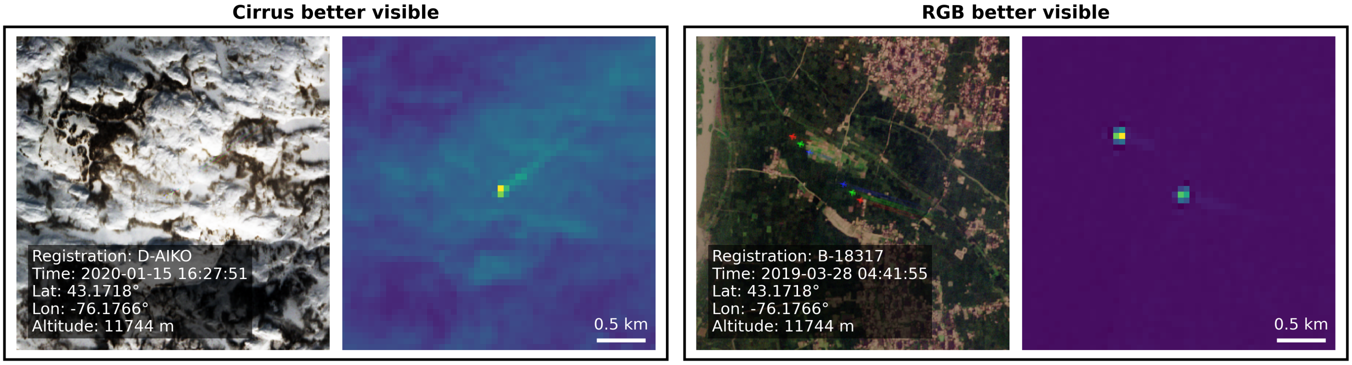

**Figure D2.** Sentinel-2 collocations comparing the visibility of contrails in true colour and cirrus band images. All images are displayed at the same scale. Note that the example on the right features the same aircraft twice, an artifact caused by the Sentinel-2 detector borders.

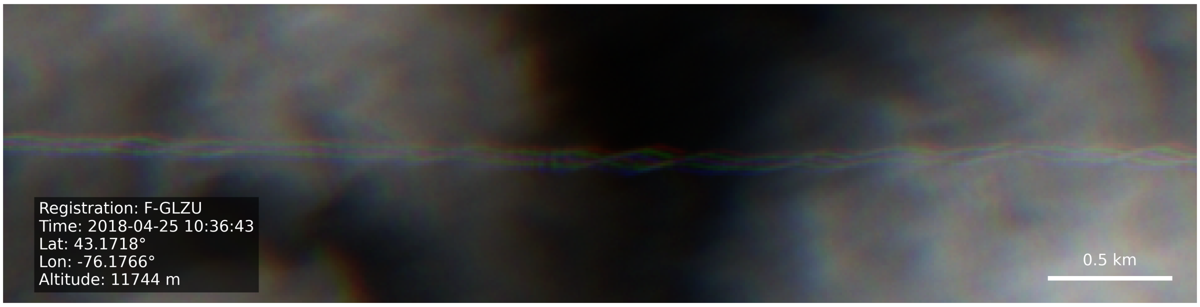

**Figure D3.** Contrail observed in Sentinel-2 imagery illustrating the occurrence of the Crow instability (Crow, 1970).

*Author contributions.* TW and VM conceptualised the study. TW, VM, and ZE developed the metholodogy and creation of the dataset. SR supported the analysis and interpretation of IAGOS data. TW and VM wrote the original manuscript. VM acquired funding. All authors have read, edited, and reviewed the manuscript and agreed upon the published version of the paper.

*Competing interests.* The contact author has declared that none of the authors has any competing interests.

*Acknowledgements.* MOZAIC/CARIBIC/IAGOS data were created with support from the European Commission, national agencies in Germany (BMBF), France (MESR), and the UK (NERC), and the IAGOS member institutions (http://www.iagos.org/partners). The participating airlines (Lufthansa, Air France, Austrian, China Airlines, Hawaiian Airlines, Air Canada, Iberia, Discover Airlines, Cathay Pacific, Air Namibia, Sabena) supported IAGOS by carrying the measurement equipment free of charge since 1994. The data are available at http://www.iagos.fr thanks to additional support from AERIS. We would also like to thank Andreas Petzold (FZ Jülich) for his helpful feedback and suggestions.

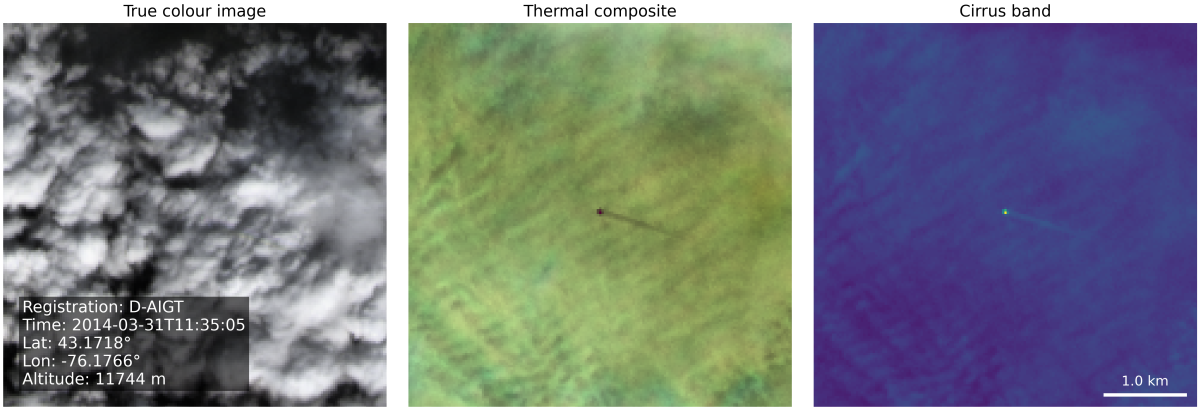

**Figure D4.** Landsat collocation comparing the visibility of a contrail in the true colour image, false colour composite based on infrared bands, and monochromatic cirrus band image. All images are displayed at the same scale.

## References

Alduchov, O. A. and Eskridge, R. E.: Improved Magnus Form Approximation of Saturation Vapor Pressure, Journal of Applied Meteorology, 35, 601–609, https://doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2, 1996.

Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., and Hanssen, R. F.: ESA's sentinel missions in support of Earth system science, Remote Sensing of Environment, 120, 84–90, https://doi.org/10.1016/j.rse.2011.07.023, 2012.

Berkes, F., Neis, P., Schultz, M. G., Bundke, U., Rohs, S., Smit, H. G. J., Wahner, A., Konopka, P., Boulanger, D., Nédélec, P., Thouret, V., and Petzold, A.: In situ temperature measurements in the upper troposphere and lowermost stratosphere from 2 decades of IAGOS long-term routine observation, Atmospheric Chemistry and Physics, 17, 12 495–12 508, https://doi.org/10.5194/acp-17-12495-2017, 2017.

Binet, R., Bergsma, E., and Poulain, V.: Accurate Sentinel-2 Inter-Band Time Delays, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-1-2022, 57–66, https://doi.org/10.5194/isprs-annals-V-1-2022-57-2022, 2022.

Borella, A., Vignon, , Boucher, O., and Rohs, S.: An Empirical Parameterization of the Subgrid-Scale Distribution of Water Vapor in the UTLS for Atmospheric General Circulation Models, Journal of Geophysical Research: Atmospheres, 129, e2024JD040 981, https://doi.org/10.1029/2024JD040981, 2024.

Boulanger, D., Thouret, V., and Petzold, A.: IAGOS Data Portal, https://doi.org/10.25326/06, 2020.

Buchholz, B., Kühnreich, B., Smit, H. G. J., and Ebert, V.: Validation of an extractive, airborne, compact TDL spectrometer for atmospheric humidity sensing by blind intercomparison, Applied Physics B, 110, 249–262, https://doi.org/10.1007/s00340-012-5143-1, 2013.

Buck, A. L.: New Equations for Computing Vapor Pressure and Enhancement Factor, Journal of Applied Meteorology and Climatology, 20, 1527–1532, https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2, 1981.

Burkhardt, U., Bock, L., and Bier, A.: Mitigating the contrail cirrus climate impact by reducing aircraft soot number emissions, npj Climate and Atmospheric Science, 1, 37, https://doi.org/10.1038/s41612-018-0046-4, 2018.

Carver, R. W. and Merose, A.: ARCO-ERA5: An Analysis-Ready Cloud-Optimized Reanalysis Dataset, American Meteorological Society, Denver, CO, USA, https://ams.confex.com/ams/103ANNUAL/meetingapp.cgi/Paper/415842, 2023.

Chevallier, R., Shapiro, M., Engberg, Z., Soler, M., Delahaye, D., Chevallier, R., Shapiro, M., Engberg, Z., Soler, M., and Delahaye, D.: Linear Contrails Detection, Tracking and Matching with Aircraft Using Geostationary Satellite and Air Traffic Data, Aerospace, 10, https://doi.org/10.3390/aerospace10070578, 2023.

Copernicus: Copernicus browser, https://browser.dataspace.copernicus.eu/, last access: 23 February 2026.

Copernicus: Sentinel-2 products specification document, Technical Report S2-PDGS-CS-DI-PSD, Issue 15.0, Copernicus Programme, European Space Agency, https://sentiwiki.copernicus.eu/__attachments/1692737/S2-PDGS-CS-DI-PSD%20-%20S2%20Product%20Specification%20Document%202024%20-%2015.0.pdf?inst-v=fd76cb18-465c-4195-8a03-c55c5e0070db, 2024.

Crow, S. C.: Stability theory for a pair of trailing vortices, AIAA Journal, 8, 2172–2179, https://doi.org/10.2514/3.6083, 1970.

Driver, O. G. A., Stettler, M. E. J., and Gryspeerdt, E.: Factors limiting contrail detection in satellite imagery, Atmospheric Measurement Techniques, 18, 1115–1134, https://doi.org/10.5194/amt-18-1115-2025, 2025.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sensing of Environment, 120, 25–36, https://doi.org/10.1016/j.rse.2011.11.026, 2012.

Engel, A., Bönisch, H., Brunner, D., Fischer, H., Franke, H., Günther, G., Gurk, C., Hegglin, M., Hoor, P., Königstedt, R., Krebsbach, M., Maser, R., Parchatka, U., Peter, T., Schell, D., Schiller, C., Schmidt, U., Spelten, N., Szabo, T., Weers, U., Wernli, H., Wetter, T., and Wirth, V.: Highly resolved observations of trace gases in the lowermost stratosphere and upper troposphere from the Spurt project: an overview, Atmospheric Chemistry and Physics, 6, 283–301, https://doi.org/10.5194/acp-6-283-2006, 2006.

Euchenhofer, M. V., Prashanth, P., Parke, S. A., Eastham, S. D., and Waitz, I. A.: Contrail Observation Limitations Using Geostationary Satellites, Geophysical Research Letters, 52, e2025GL118 386, https://doi.org/10.1029/2025GL118386, 2025.

EUROCONTROL: User manual for the base of aircraft data (BADA) family 4, EEC Technical/Scientific Report 12/11/22-58, EUROCONTROL Experimental Centre (EEC), https://www.eurocontrol.int/model/bada, 2016.

Franks, S., Storey, J., and Rengarajan, R.: The New Landsat Collection-2 Digital Elevation Model, Remote Sensing, 12, https://doi.org/10.3390/rs12233909, 2020.

Freudenthaler, V., Homburg, F., and Jäger, H.: Contrail observations by ground-based scanning lidar: Cross-sectional growth, Geophysical Research Letters, 22, 3501–3504, https://doi.org/10.1029/95GL03549, 1995.

Geraedts, S., Brand, E., Dean, T. R., Eastham, S., Elkin, C., Engberg, Z., Hager, U., Langmore, I., McCloskey, K., Yue-Hei Ng, J., Platt, J. C., Sankar, T., Sarna, A., Shapiro, M., and Goyal, N.: A scalable system to measure contrail formation on a per-flight basis, Environmental Research Communications, 6, 015 008, https://doi.org/10.1088/2515-7620/ad11ab, 2024.

Gierens, K., Kohlhepp, R., Dotzek, N., and Smit, H. G.: Instantaneous fluctuations of temperature and moisture in the upper troposphere and tropopause region. Part 1: Probability densities and their variability, Meteorologische Zeitschrift, 16, 221–231, https://doi.org/10.1127/0941-2948/2007/0197, 2007.

Gierens, K., Matthes, S., and Rohs, S.: How Well Can Persistent Contrails Be Predicted?, Aerospace, 7, 169, https://doi.org/10.3390/aerospace7120169, 2020.

Google: Sentinel-2 - Marketplace -Google Cloud console, https://console.cloud.google.com/marketplace/product/esa-public-data/sentinel2, last access: 23 February 2026, 2026.

Gryspeerdt, E., Stettler, M. E. J., Teoh, R., Burkhardt, U., Delovski, T., Driver, O. G. A., and Painemal, D.: Operational differences lead to longer lifetimes of satellite detectable contrails from more fuel efficient aircraft, Environmental Research Letters, 19, 084 059, https://doi.org/10.1088/1748-9326/ad5b78, 2024.

Heiselberg, H.: Aircraft and Ship Velocity Determination in Sentinel-2 Multispectral Images, Sensors, 19, 2873, https://doi.org/10.3390/s19132873, 2019.

Helten, M., Smit, H. G. J., Sträter, W., Kley, D., Nedelec, P., Zöger, M., and Busen, R.: Calibration and performance of automatic compact instrumentation for the measurement of relative humidity from passenger aircraft, Journal of Geophysical Research: Atmospheres, 103, 25 643–25 652, https://doi.org/10.1029/98JD00536, 1998.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoshizaki, H., Conti, R. J., Anderson, L. B., Redler, K. O., and Meyer, J. W.: Study of High-Altitude Aircraft Wake Dynamics. Task 1. Problem Definition, https://apps.dtic.mil/sti/html/tr/AD0754918/index.html, number: DOTTST903, 1972.

IAGOS: Data quality – IAGOS, https://iagos.aeris-data.fr/data-quality/, last access: 23 February 2026.

Immler, F., Treffeisen, R., Engelbart, D., Krüger, K., and Schrems, O.: Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes, Atmospheric Chemistry and Physics, 8, 1689–1699, https://doi.org/10.5194/acp-8-1689-2008, 2008.

Irons, J. R., Dwyer, J. L., and Barsi, J. A.: The next Landsat satellite: The Landsat Data Continuity Mission, Remote Sensing of Environment, 122, 11–21, https://doi.org/10.1016/j.rse.2011.08.026, 2012.

Iwabuchi, H., Yang, P., Liou, K. N., and Minnis, P.: Physical and optical properties of persistent contrails: Climatology and interpretation, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD017020, 2012.

Jarry, G., Dalmau, R., Very, P., Ballerini, F., and Bocu, S.-D.: GVCCS: a dataset for contrail identification and tracking on visible whole sky camera sequences, Earth System Science Data, 18, 1037–1059, https://doi.org/10.5194/essd-18-1037-2026, 2026.

Jeßberger, P., Voigt, C., Schumann, U., Sölch, I., Schlager, H., Kaufmann, S., Petzold, A., Schäuble, D., and Gayet, J.-F.: Aircraft type influence on contrail properties, Atmospheric Chemistry and Physics, 13, 11 965–11 984, https://doi.org/10.5194/acp-13-11965-2013, 2013.

Knight, E. J., Kvaran, G., Knight, E. J., and Kvaran, G.: Landsat-8 Operational Land Imager Design, Characterization and Performance, Remote Sensing, 6, 10 286–10 305, https://doi.org/10.3390/rs61110286, 2014.

Kunz, A., Schiller, C., Rohrer, F., Smit, H. G. J., Nedelec, P., and Spelten, N.: Statistical analysis of water vapour and ozone in the UT/LS observed during SPURT and MOZAIC, Atmospheric Chemistry and Physics, 8, 6603–6615, https://doi.org/10.5194/acp-8-6603-2008, 2008.

Kärcher, B.: Formation and radiative forcing of contrail cirrus, Nature Communications, 9, 1824, https://doi.org/10.1038/s41467-018-04068-0, 2018.

Lee, D., Fahey, D., Skowron, A., Allen, M., Burkhardt, U., Chen, Q., Doherty, S., Freeman, S., Forster, P., Fuglestvedt, J., Gettelman, A., De León, R., Lim, L., Lund, M., Millar, R., Owen, B., Penner, J., Pitari, G., Prather, M., Sausen, R., and Wilcox, L.: The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018, Atmospheric Environment, 244, 117 834, https://doi.org/10.1016/j.atmosenv.2020.117834, 2021.

Lewellen, D. C. and Lewellen, W. S.: The Effects of Aircraft Wake Dynamics on Contrail Development, Journal of the Atmospheric Sciences, 58, 390–406, https://doi.org/10.1175/1520-0469(2001)058<0390:TEOAWD>2.0.CO;2, 2001.

Li, Y., Mahnke, C., Rohs, S., Bundke, U., Spelten, N., Dekoutsidis, G., Groß, S., Voigt, C., Schumann, U., Petzold, A., and Krämer, M.: Upper-tropospheric slightly ice-subsaturated regions: frequency of occurrence and statistical evidence for the appearance of contrail cirrus, Atmospheric Chemistry and Physics, 23, 2251–2271, https://doi.org/10.5194/acp-23-2251-2023, 2023.

Liu, Y., Xu, B., Zhi, W., Hu, C., Dong, Y., Jin, S., Lu, Y., Chen, T., Xu, W., Liu, Y., Zhao, B., and Lu, W.: Space eye on flying aircraft: From Sentinel-2 MSI parallax to hybrid computing, Remote Sensing of Environment, 246, 111 867, https://doi.org/10.1016/j.rse.2020.111867, 2020.

Low, J., Teoh, R., Ponsonby, J., Gryspeerdt, E., Shapiro, M., and Stettler, M. E. J.: Ground-based contrail observations: comparisons with reanalysis weather data and contrail model simulations, Atmospheric Measurement Techniques, 18, 37–56, https://doi.org/10.5194/amt-18-37-2025, 2025.

Mannstein, H., Meyer, R., and Wendling, P.: Operational detection of contrails from NOAA-AVHRR-data, International Journal of Remote Sensing, 20, 1641–1660, https://doi.org/10.1080/014311699212650, 1999.

Mannstein, H., Brömser, A., and Bugliaro, L.: Ground-based observations for the validation of contrails and cirrus detection in satellite imagery, Atmospheric Measurement Techniques, 3, 655–669, https://doi.org/10.5194/amt-3-655-2010, 2010.

Martin Frias, A., Shapiro, M. L., Engberg, Z., Zopp, R., Soler, M., and Stettler, M. E. J.: Feasibility of contrail avoidance in a commercial flight planning system: an operational analysis, Environmental Research: Infrastructure and Sustainability, 4, 015 013, https://doi.org/10.1088/2634-4505/ad310c, 2024.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., and Jenstrom, D. T.: Landsat 9: Empowering open science and applications through continuity, Remote Sensing of Environment, 248, 111 968, https://doi.org/10.1016/j.rse.2020.111968, 2020.

McCloskey, K., Geraedts, S., Jackman, B., Meijer, V. R., Brand, E., Fork, D., Platt, J. C., Elkin, C., and Van Arsdale, C.: A human-labeled Landsat-8 contrails dataset, in: Climate Change AI, Climate Change AI, https://www.climatechange.ai/papers/icml2021/2, 2021.

Meijer, V. R.: Satellite-based Analysis and Forecast Evaluation of Aviation Contrails, PhD thesis, Massachusetts Institute of Technology, https://hdl.handle.net/1721.1/155350, 2024.

Meijer, V. R., Kulik, L., Eastham, S. D., Allroggen, F., Speth, R. L., Karaman, S., and Barrett, S. R. H.: Contrail coverage over the United States before and during the COVID-19 pandemic, Environmental Research Letters, 17, 034 039, https://doi.org/10.1088/1748-9326/ac26f0, 2022.

Meijer, V. R., Eastham, S. D., Waitz, I. A., and Barrett, S. R. H.: Contrail altitude estimation using GOES-16 ABI data and deep learning, Atmospheric Measurement Techniques, 17, 6145–6162, https://doi.org/10.5194/amt-17-6145-2024, 2024.

Meyer, J., Rolf, C., Schiller, C., Rohs, S., Spelten, N., Afchine, A., Zöger, M., Sitnikov, N., Thornberry, T. D., Rollins, A. W., Bozóki, Z., Tátrai, D., Ebert, V., Kühnreich, B., Mackrodt, P., Möhler, O., Saathoff, H., Rosenlof, K. H., and Krämer, M.: Two decades of water vapor measurements with the FISH fluorescence hygrometer: a review, Atmospheric Chemistry and Physics, 15, 8521–8538, https://doi.org/10.5194/acp-15-8521-2015, 2015.

Märkl, R. S., Voigt, C., Sauer, D., Dischl, R. K., Kaufmann, S., Harlaß, T., Hahn, V., Roiger, A., Weiß-Rehm, C., Burkhardt, U., Schumann, U., Marsing, A., Scheibe, M., Dörnbrack, A., Renard, C., Gauthier, M., Swann, P., Madden, P., Luff, D., Sallinen, R., Schripp, T., and Le Clercq, P.: Powering aircraft with 100 % sustainable aviation fuel reduces ice crystals in contrails, Atmospheric Chemistry and Physics, 24, 3813–3837, https://doi.org/10.5194/acp-24-3813-2024, 2024.

Neis, P., , Herman G. J., S., , Susanne, R., , Ulrich, B., , Martina, K., , Nicole, S., , Volker, E., , Bernhard, B., , Karin, T., , and Petzold, A.: Quality assessment of MOZAIC and IAGOS capacitive hygrometers: insights from airborne field studies, Tellus B: Chemical and Physical Meteorology, 67, 28 320, https://doi.org/10.3402/tellusb.v67.28320, 2015a.

Neis, P., Smit, H. G. J., Krämer, M., Spelten, N., and Petzold, A.: Evaluation of the MOZAIC Capacitive Hygrometer during the airborne field study CIRRUS-III, Atmospheric Measurement Techniques, 8, 1233–1243, https://doi.org/10.5194/amt-8-1233-2015, 2015b.

Ng, J. Y.-H., McCloskey, K., Cui, J., Meijer, V. R., Brand, E., Sarna, A., Goyal, N., Arsdale, C. V., and Geraedts, S.: OpenContrails: Benchmarking Contrail Detection on GOES-16 ABI, arXiv [preprint], https://doi.org/10.48550/arXiv.2304.02122, 2023.

Ortiz, I., García-Heras, J., Jafarimoghaddam, A., and Soler, M.: Robust Evaluation of Neural Networks Trained on the OpenContrails Dataset, IEEE Transactions on Geoscience and Remote Sensing, 63, 1–17, https://doi.org/10.1109/TGRS.2025.3629628, 2025.

Paoli, R. and Shariff, K.: Contrail Modeling and Simulation, Annual Review of Fluid Mechanics, 48, 393–427, https://doi.org/10.1146/annurev-fluid-010814-013619, 2016.

Paoli, R., Nybelen, L., Picot, J., and Cariolle, D.: Effects of jet/vortex interaction on contrail formation in supersaturated conditions, Physics of Fluids, 25, 053 305, https://doi.org/10.1063/1.4807063, 2013.

Paugam, R., Paoli, R., and Cariolle, D.: Influence of vortex dynamics and atmospheric turbulence on the early evolution of a contrail, Atmospheric Chemistry and Physics, 10, 3933–3952, https://doi.org/10.5194/acp-10-3933-2010, 2010.

Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A. M., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., Boulanger, D., Marshall, J., Nédélec, P., Smit, H. G. J., Friess, U., Flaud, J.-M., Wahner, A., Cammas, J.-P., Volz-Thomas, A., and Team, I.: Global-scale atmosphere monitoring by in-service aircraft – current achievements and future prospects of the European Research Infrastructure IAGOS, Tellus B: Chemical and Physical Meteorology, 67, https://doi.org/10.3402/tellusb.v67.28452, 2015.

Petzold, A., Khan, N. F., Li, Y., Spichtinger, P., Rohs, S., Crewell, S., Wahner, A., and Krämer, M.: Most long-lived contrails form within cirrus clouds with uncertain climate impact, Nature Communications, 16, 9695, https://doi.org/10.1038/s41467-025-65532-2, 2025.

Poll, D. and Schumann, U.: An estimation method for the fuel burn and other performance characteristics of civil transport aircraft in the cruise. Part 1 fundamental quantities and governing relations for a general atmosphere, The Aeronautical Journal, 125, 257–295, https://doi.org/10.1017/aer.2020.62, 2021a.

Poll, D. and Schumann, U.: An estimation method for the fuel burn and other performance characteristics of civil transport aircraft during cruise: part 2, determining the aircraft's characteristic parameters, The Aeronautical Journal, 125, 296–340, https://doi.org/10.1017/aer.2020.124, 2021b.

Poll, D. I. A. and Schumann, U.: An estimation method for the fuel burn and other performance characteristics of civil transport aircraft; part 3 full flight profile when the trajectory is specified, The Aeronautical Journal, 129, 825–861, https://doi.org/10.1017/aer.2024.141, 2025.

Rolf, C., Rohs, S., Smit, H. G., Krämer, M., Bozóki, Z., Hofmann, S., Franke, H., Maser, R., Hoor, P., and Petzold, A.: Evaluation of compact hygrometers for continuous airborne measurements, Meteorologische Zeitschrift, 33, 15–34, https://doi.org/10.1127/metz/2023/1187, 2024.

Sarna, A., Meijer, V., Chevallier, R., Duncan, A., McConnaughay, K., Geraedts, S., and McCloskey, K.: Benchmarking and improving algorithms for attributing satellite-observed contrails to flights, Atmospheric Measurement Techniques, 18, 3495–3532, https://doi.org/10.5194/amt-18-3495-2025, 2025.

Sassen, K.: Contrail-Cirrus and Their Potential for Regional Climate Change, Bulletin of the American Meteorological Society, 78, 1885–1903, https://doi.org/10.1175/1520-0477(1997)078<1885:CCATPF>2.0.CO;2, 1997.

Sausen, R., Hofer, S., Gierens, K., Bugliaro, L., Ehrmanntraut, R., Sitova, I., Walczak, K., Burridge-Diesing, A., Bowman, M., and Miller, N.: Can we successfully avoid persistent contrails by small altitude adjustments of flights in the real world?, Meteorologische Zeitschrift, pp. 83–98, https://doi.org/10.1127/metz/2023/1157, 2024.

Schumann, U.: On conditions for contrail formation from aircraft exhausts, Meteorologische Zeitschrift, 5, 4–23, https://doi.org/10.1127/metz/5/1996/4, 1996.

Schumann, U.: A contrail cirrus prediction model, Geoscientific Model Development, 5, 543–580, https://doi.org/10.5194/gmd-5-543-2012, 2012.

Schumann, U., Hempel, R., Flentje, H., Garhammer, M., Graf, K., Kox, S., Lösslein, H., and Mayer, B.: Contrail study with ground-based cameras, Atmospheric Measurement Techniques, 6, 3597–3612, https://doi.org/10.5194/amt-6-3597-2013, 2013.

Schumann, U., Baumann, R., Baumgardner, D., Bedka, S. T., Duda, D. P., Freudenthaler, V., Gayet, J.-F., Heymsfield, A. J., Minnis, P., Quante, M., Raschke, E., Schlager, H., Vázquez-Navarro, M., Voigt, C., and Wang, Z.: Properties of individual contrails: a compilation of observations and some comparisons, Atmospheric Chemistry and Physics, 17, 403–438, https://doi.org/10.5194/acp-17-403-2017, 2017.

Schäfer, M. and Jonáš, P.: ADS-B Positional Accuracy and Anomalies: A Comprehensive Analysis Using High-Resolution MLAT Data, in: 2025 Integrated Communications, Navigation and Surveillance Conference (ICNS), pp. 1–14, ISSN 2155-4951, https://doi.org/10.1109/ICNS65417.2025.10976935, 2025.

Seeber, G.: Satellite geodesy: Foundations, methods and applications, Walter de Gruyter, Berlin, ISBN 3-11-012753-9, 1993.

Shapiro, M., Engberg, Z., Teoh, R., Stettler, M., Dean, T., and Abbott, T.: pycontrails: Python library for modeling aviation climate impacts, https://doi.org/10.5281/zenodo.18260617, 2026.

Sonabend-W, A., Elkin, C., Dean, T., Dudley, J., Ali, N., Blickstein, J., Brand, E., Broshears, B., Chen, S., Engberg, Z., Galyen, M., Geraedts, S., Goyal, N., Grenham, R., Hager, U., Hecker, D., Jany, M., McCloskey, K., Ng, J., Norris, B., Opel, F., Rothenberg, J., Sankar, T., Sanekommu, D., Sarna, A., Schütt, O., Shapiro, M., Soh, R., Van Arsdale, C., and Platt, J. C.: Feasibility test of per-flight contrail avoidance in commercial aviation, Communications Engineering, 3, 184, https://doi.org/10.1038/s44172-024-00329-7, 2024.

Sonabend-W, A., Geraedts, S., Goyal, N., Ng, J. Y.-H., Van Arsdale, C., and McCloskey, K.: Observing long-lived longwave contrail forcing, https://doi.org/10.5194/egusphere-2025-3739, 2025.

Teoh, R., Schumann, U., Majumdar, A., and Stettler, M. E. J.: Mitigating the Climate Forcing of Aircraft Contrails by Small-Scale Diversions and Technology Adoption, Environmental Science & Technology, 54, 2941–2950, https://doi.org/10.1021/acs.est.9b05608, 2020.

Teoh, R., Schumann, U., Gryspeerdt, E., Shapiro, M., Molloy, J., Koudis, G., Voigt, C., and Stettler, M. E. J.: Aviation contrail climate effects in the North Atlantic from 2016 to 2021, Atmospheric Chemistry and Physics, 22, 10 919–10 935, https://doi.org/10.5194/acp-22-10919-2022, 2022.

Teoh, R., Engberg, Z., Schumann, U., Voigt, C., Shapiro, M., Rohs, S., and Stettler, M. E. J.: Global aviation contrail climate effects from 2019 to 2021, Atmospheric Chemistry and Physics, 24, 6071–6093, https://doi.org/10.5194/acp-24-6071-2024, 2024.

Tompkins, Gierens, and Rädel: Ice supersaturation in the ECMWF integrated forecast system, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.14, 2007.

U.S. Geological Survey: USGS machine-to-machine (M2M) API, https://m2m.cr.usgs.gov/, last access: 23 February 2026.

Voigt, C., Schumann, U., Jurkat, T., Schäuble, D., Schlager, H., Petzold, A., Gayet, J.-F., Krämer, M., Schneider, J., Borrmann, S., Schmale, J., Jessberger, P., Hamburger, T., Lichtenstern, M., Scheibe, M., Gourbeyre, C., Meyer, J., Kübbeler, M., Frey, W., Kalesse, H., Butler, T., Lawrence, M. G., Holzäpfel, F., Arnold, F., Wendisch, M., Döpelheuer, A., Gottschaldt, K., Baumann, R., Zöger, M., Sölch, I., Rautenhaus, M., and Dörnbrack, A.: In-situ observations of young contrails – overview and selected results from the CONCERT campaign, Atmospheric Chemistry and Physics, 10, 9039–9056, https://doi.org/10.5194/acp-10-9039-2010, 2010.

Voigt, C., Schumann, U., Jessberger, P., Jurkat, T., Petzold, A., Gayet, J.-F., Krämer, M., Thornberry, T., and Fahey, D. W.: Extinction and optical depth of contrails, Geophysical Research Letters, 38, n/a–n/a, https://doi.org/10.1029/2011GL047189, 2011.

Voigt, C., Kleine, J., Sauer, D., Moore, R. H., Bräuer, T., Le Clercq, P., Kaufmann, S., Scheibe, M., Jurkat-Witschas, T., Aigner, M., Bauder, U., Boose, Y., Borrmann, S., Crosbie, E., Diskin, G. S., DiGangi, J., Hahn, V., Heckl, C., Huber, F., Nowak, J. B., Rapp, M., Rauch, B., Robinson, C., Schripp, T., Shook, M., Winstead, E., Ziemba, L., Schlager, H., and Anderson, B. E.: Cleaner burning aviation fuels can reduce contrail cloudiness, Communications Earth & Environment, 2, 114, https://doi.org/10.1038/s43247-021-00174-y, 2021.

Wang, Z., Bugliaro, L., Gierens, K., Hegglin, M. I., Rohs, S., Petzold, A., Kaufmann, S., and Voigt, C.: Machine learning for improvement of upper-tropospheric relative humidity in ERA5 weather model data, Atmospheric Chemistry and Physics, 25, 2845–2861, https://doi.org/10.5194/acp-25-2845-2025, 2025.

Woldhuis, T., Meijer, V., Engberg, Z., and Rohs, S.: Dataset accompanying "Observing formation and early evolution of contrails formed by IAGOS aircraft using high-resolution LEO satellite imagery", https://doi.org/10.4121/2d66d65e-8041-4435-ab3c-0af3fdfc5d23, 2026a.

Woldhuis, T., Meijer, V., Engberg, Z., and Rohs, S.: Code accompanying "Observing formation and early evolution of contrails formed by IAGOS aircraft using high-resolution LEO satellite imagery", https://doi.org/10.4121/a1e4d5b4-5af2-4a13-8152-a1e8d49c297f, 2026b.

Wolf, K., Bellouin, N., Boucher, O., Rohs, S., and Li, Y.: Correction of ERA5 temperature and relative humidity biases by bivariate quantile mapping for contrail formation analysis, Atmospheric Chemistry and Physics, 25, 157–181, https://doi.org/10.5194/acp-25-157-2025, 2025.

Zhao, F., Xia, L., Kylling, A., Li, R., Shang, H., and Xu, M.: Detection flying aircraft from Landsat 8 OLI data, ISPRS Journal of Photogrammetry and Remote Sensing, 141, 176–184, https://doi.org/10.1016/j.isprsjprs.2018.05.001, 2018.