https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

A flight-first approach for contrail detection using ground-based cameras

Dominik P. Nakoneczny[1], Jonathan P. Itcovitz[1], Denis Vida[2,3], Luc Busquin[4], Tristan H. Abbott[5], Marc E. J. Stettler[1]

[1]Department of Civil & Environmental Engineering, Imperial College London, London, SW7 2AZ, United Kingdom
[2]Department of Physics & Astronomy, University of Western Ontario, London, ON, N6A 3K7, Canada
[3]Institute for Earth and Space Exploration, University of Western Ontario, London, ON, N6A 5B8, Canada
[4]ContrailCast LLC, Phoenix, AZ, United States of America
[5]Contrails.org, 4110 Carillon Point, Kirkland, WA, United States of America

Correspondence to: Marc E. J. Stettler (m.stettler@imperial.ac.uk)

Abstract. Validating forecasts and models used in contrail mitigation efforts requires high fidelity observations. The ability to match individual contrails with their progenitor flights in such observations is vital. High-resolution ground cameras can offer such high-fidelity flight attributions through direct observations of contrail formation, something that low-resolution observation methods (e.g., geostationary satellites) cannot due to their optical limitations. Previous "detection-first" approaches using ground cameras have employed flight matching after detecting unattributed contrails and can thus still struggle with the attribution task due to having no inherent mechanism to distinguish between newer and older contrails, particularly suffering when in-frame contrail densities are large. Here, we present a "flight-first" approach that focally searches for contrails within dynamic bounding boxes around wind-advected flightpaths, enabling inherent flight matching for young contrails while reducing computational costs. We demonstrate our approach using ground cameras from the Global Meteor Network (GMN), the SAM2 asemantic segmentation model, and a simple random forest classifier model (RFM). In a case study of 3,506 flights across three weeks in the Southwestern USA, we detect 265 contrails with a precision of 0.99 and a recall of 0.74. The flight-first approach uniquely allows us to pinpoint where detections were missed by the SAM2-RFM pipeline, identifying cloud obscuration as a prominent driver of failure. We thus propose that our off-the-shelf approach can be deployed at-scale as a primary low-cost layer in a wider contrail detection strategy for model evaluation, serving to identify places and times for follow-up observations with more advanced observation techniques. The GMN is well suited to such a task given its global coverage and standardised components.

1 Introduction

Condensation trails (contrails), the linear ice clouds that form behind aircraft, are generated via water vapour condensing and freezing onto particulate matter stemming from the aircraft engine exhaust. Contrails can have both a warming and cooling effect through their interactions with shortwave and longwave radiation. However, their global net radiative forcing effect is positive, and hence they have an overall warming effect on the climate (Haywood et al., 2009) with the global annual mean contrail radiative forcing estimated between 34.8 and 74.8 mWm^{-2} in 2019 (Teoh et al., 2024a). Aviation as a whole contributes 3.5 (5-95% likelihood range of 4.0-3.4) % of total global anthropogenic radiative forcing (Lee et al, 2021), of which contrails make up approximately 57% based on central estimates reported by Lee et al., (2021).With the future growth of the aviation sector, this contrail climate impact will only grow, with the potential to triple by 2050 without mitigative action (Singh et al., 2024; Smith et al., 2026).

The lifetimes of contrails are relatively short (e.g., Gierens & Vazquez-Navarro (2017) found an average of 3.7±2.8 hours for persistent contrails) and their climate impacts are highly non-linear, with some airspace regions seeing just 2% of flights being

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

responsible for 80% of the contrail radiative forcing (Teoh et al., 2020). Together, these offer a promising opportunity for the mitigation of a large portion of aviation's annual climate impact in the near future: If a theoretical switch could be flipped turning off contrails tomorrow, it would provide the world with better living conditions during the time it takes to reduce long-term climate forcers, such as CO_2.

1.1 The need for observational data: Satellites & ground-based cameras

One of the barriers to contrail mitigation is accurately forecasting where contrails are going to form, persist, and have the greatest climate effect. To predict the formation step, we can utilise the Schmidt-Appleman criterion (SAC; Schmidt 1941, Appleman 1954, Schumann 1996). This describes the ambient atmospheric conditions with which the water-vapour-rich aircraft engine exhaust can mix and reach the conditions required for liquid particles to form on engine-derived particulate matter, which subsequently freeze to form ice particles. For these ice particles to persist, the ambient air around the initial contrail plume must be supersaturated with respect to ice (RH_{ice} > 100%). Atmospheric regions that fulfil such persistence criterion are commonly termed ice-supersaturated regions (ISSRs). ISSRs tend to be laterally wide but vertically thin, with lateral extents of hundreds of kms and vertical extents typically below 2000ft (Gierens et al., 2000; Spichtinger et al., 2003; Teoh et al., 2020), while exhibiting finer scale variability with a median path length of 3km (Reutter et al., 2020). Contrail ice particles formed in subsaturated conditions will quickly sublimate, leading to short-lived contrails.

Contrail formation and evolution can be predicted using models, such as the Contrail Cirrus Prediction model (CoCiP; Schumann, 2012). The ability to estimate contrail climate effects based on meteorological forecasts enables mitigation through flight deviations, mostly minor vertical adjustments on the order of up to ±2,000ft (e.g., Sausen et al. 2023, Martin-Frias et al. 2024, Sonabend-W et al., 2024, Sankar et al. 2026). However, to be used operationally and confidently, such models must be evaluated using observations, ideally throughout their modelling chain (i.e., formation, persistence, climate effect). A primary source of such evaluation data is flight-attributed contrail observations, which enable direct comparisons of forecasts with reality (meteorological predictions and contrail model outputs). Flight-attributed contrail observations can also enable the evaluation of contrail mitigation strategies. In the longer term, they can potentially support near real-time "nowcasting" of ISSRs and/or persistent contrail regions (PCRs).

Contrail detection efforts to date have largely been based on imagery from geostationary satellites, such as the GOES-R series. Such observations take repeated images of the same field of view at regular intervals, making them ideal for tracking persistent features like contrails. Current satellite-based detection techniques typically involve feeding multispectral composite imagery into deep learning algorithms that output masks of pixels predicted to contain contrails. These models are trained on human-labelled datasets (Meijer et al., 2022; Ng et al., 2023; Gyrspeerdt et al., 2024) and achieve reasonable accuracy when evaluated against said datasets (Hoffman et al., 2023; Nirob et al., 2023). However, there exists no observational ground-truth dataset, meaning that the accuracy of these models is limited to contrails that are detectable on satellite imagery by human-labellers. This could lead to an overestimate of real-world performance and underestimate of the true contrail population (e.g., Euchenhofer et al. 2025), particularly for faint, short-lived, or late-stage contrails with non-linear morphologies, or those that exist in complex backgrounds (Driver et al., 2024). Developing such a ground-truth observational dataset would, therefore, be a powerful next step in observational capabilities for contrails model evaluation.

One of the main attractions to using geostationary satellites is that they offer wide geospatial coverage over land and ocean. However, this comes at a cost of temporal and spatial resolution, resulting in potential losses of contrail detections and inter-frame motions (Driver et al., 2024). For example, the GOES-R series satellites' Full Disk images have a 10-minute frame separation and a 2 km pixel resolution at nadir (stretching to much coarser resolution away from the image centre), in comparison to initial contrail widths of ~100m (Schumann, 2012). Contrails can only be detected in GOES imagery after significant lateral spreading resulting in mean times to first detection in imagery of 17-41 minutes (Gereadts et al., 2024; Gryspeerdt et al., 2024). Resolution limitations are highlighted when attempting flight attribution, where it can severely

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

undermine confidence, especially in congested airspaces with overlapping contrails (Sarna et al., 2025). The build-up of windspeed errors with time (i.e., the position of advected flight tracks becomes more uncertain over time by ~12 km/h) further complicates flight attribution (Chevallier et al., 2023). While geostationary satellite methods remain essential for contrail monitoring on a global scale, especially over oceans, their limitations in spatial and temporal resolutions, and limited confidence in flight attribution motivate the exploration of alternative observational strategies.

Ground-based observational methods are a promising complement to satellite-based contrail detection. Their finer spatial and temporal resolution enables observations of faint and/or short-lived contrails that may be missed in satellite imagery, although this is a trade-off against a narrower field of view (FOV) and typically more limited multi-spectral capabilities. Schumann et al. (2013) first introduced contrail-flight matching in ground-based observations, utilising photogrammetric analysis on a multiple camera system and manual flight waypoint comparison to enable comparisons to CoCiP simulations. While the study was limited to a case study of four contrails, others have followed at larger scales, including a similar but larger study by Rosenow et al. (2023), as well as Low et al. (2024) with a methodology driven by CoCiP simulations and ADS-B flight tracks (see Table S1). Such studies demonstrated the feasibility for ground-based flight-attributed contrail detection but also highlighted the need for automated processes to handle large volumes of footage without extensive manual verification.

1.2 The "Flight-first" approach compared to "Detection-first" approaches

To date, methods for automated contrail detection in ground-based imagery have relied on computer vision techniques or machine-learning-based image segmentation models to identify contrail instances which are then matched to progenitor flights in a post-processing step. We term this the "detection-first" approach, in which the performance of the segmentation model is crucial to finding contrails. Several studies have implemented such methodologies (see Table S1 for details). Most recently, Jarry et al. (2025) constructed a large dataset (>4500 contrails) of such detections utilising an All-Sky camera and a detection algorithm trained on hand-labelled data. This was then leveraged by Dalmau et al. (2025) to develop a contrail-flight attribution framework, implementing likelihood matching scores (greedy assignment) based on dry advection. The resulting dataset represents the most advanced and extensive one to-date. However, one noted limitation was the presence of classification errors, where "new" contrails were misclassified as "old" and consequently wrongly omitted from certain flight-matches. Detection-first methodologies are naturally exposed to such classification errors, without an inherent mechanism to distinguish between contrails formed within the frame and those that have advected in after forming outside of the frame. This limitation to flight-matching is likely to be further pronounced in crowded airspaces with many persisting contrails.

This study introduces a "flight-first" approach that is built around determining whether an individual flight formed a contrail, rather than detecting instances of contrails within a frame. This framework naturally accounts for both contrails that form within the camera's field of view and those that form outside of it before subsequently advecting into view, avoiding the ambiguity between "new" and "old" contrails that 'detection-first" methods must resolve in post process. By beginning with known flight trajectories rather than detected image features, both positive and negative detections are inherently flight-matched, with negative detections arising from the explicit evaluation of the particular flight rather than from the assumed absence of contrails matched to that flight.

The aim of this work is not to propose a definitive implementation of the flight-first approach, but rather to introduce a modular flight-first framework that can be implemented using different feature extraction and classification models. The aim of the framework is to generate high-fidelity flight-attributed datasets aimed at formation detections and targeting observations from other sources. We demonstrate this framework here by (1) implementing a proof-of-concept model using a general asemantic segmentation model for feature extraction and a Random Forest classifier and (2) evaluating the performance of the model on a case study of data from the Global Meteor Network (GMN) (Sect. 2). We show that the model is capable classifying persistent contrails with very high precision (>0.99) when the evaluation dataset is curated to eliminate cases where the asemantic

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

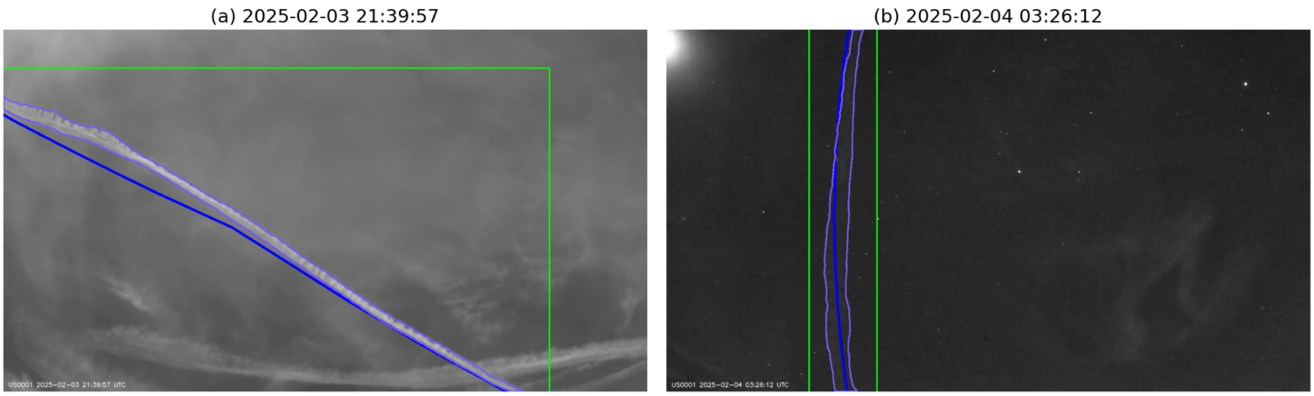

Figure 1: Two example frames from a GMN camera in the central-western United States, demonstrating the (a) daytime and (b) nighttime capture modes. The green bounding rectangle is the dynamically adjusted prompt box, the advected flightpath is in blue, and the outline of the detected and segmented contrail is the purple contour.

segmentation model extracts irrelevant features (Sect. 3). Although precision drops slightly when evaluated on an uncurated version of the evaluation dataset, we conclude by arguing that simple modifications to the framework (e.g., replacing the asemantic segmentation model with a segmentation model trained specifically on contrails, or adding an additional classifier to screen for cases where feature extraction fails) provide promising paths toward deploying flight-first contrail detection models that attain very high precision at scale (Sect. 4).

2 Materials and methods

This study presents a fully automated pipeline for the detection, segmentation, and tracking of contrails. It is the first to use cameras from the Global Meteor Network (GMN), detailed in Sect. 2.1, for fully automated contrail detection from a "flight-first" perspective, demonstrating its ability to generate a high-fidelity flight-attributed ground-based contrail detection dataset. Here, we present details of our analysis pipeline, including the data sources we leveraged (Sect. 2.2 & 2.3), the segmentation framework we built based on SAM2 (Sect. 2.4), and the simple random forest object classifier we designed to handle the segmentation output (Sect. 2.5 & 2.6). We also describe here the logic we implemented to define contrail termination within our analysis pipeline (Sect. 2.7) as well as the associated classification of contrail ages (Sect. 2.8).

2.1 Ground based footage

We use footage from the GMN as our source of ground-based contrail observations. GMN is a network of wide-angle ($88^{\circ} \times 48^{\circ}$) low-cost cameras in continuous monitoring operation that are designed primarily for monitoring meteor showers and meteorite-dropping fireballs (Vida et al. 2021). The GMN operates at a video resolution of 1280 x 720 px and most cameras use Raspberry Pi single-board computers as the processing unit. The setup is optionally capable of operating at 1080p, but a lower resolution was chosen to keep the data size and processing requirements manageable. Most cameras are deployed in single-camera configuration, but about half of the sites have two of more cameras, and some provide all-sky coverage with 6+ cameras. Several features make GMN's cameras attractive for contrail observations. Although limited to land, they have a wide global network, resulting in observations that span several airspaces across the globe and enable robust analysis in a

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

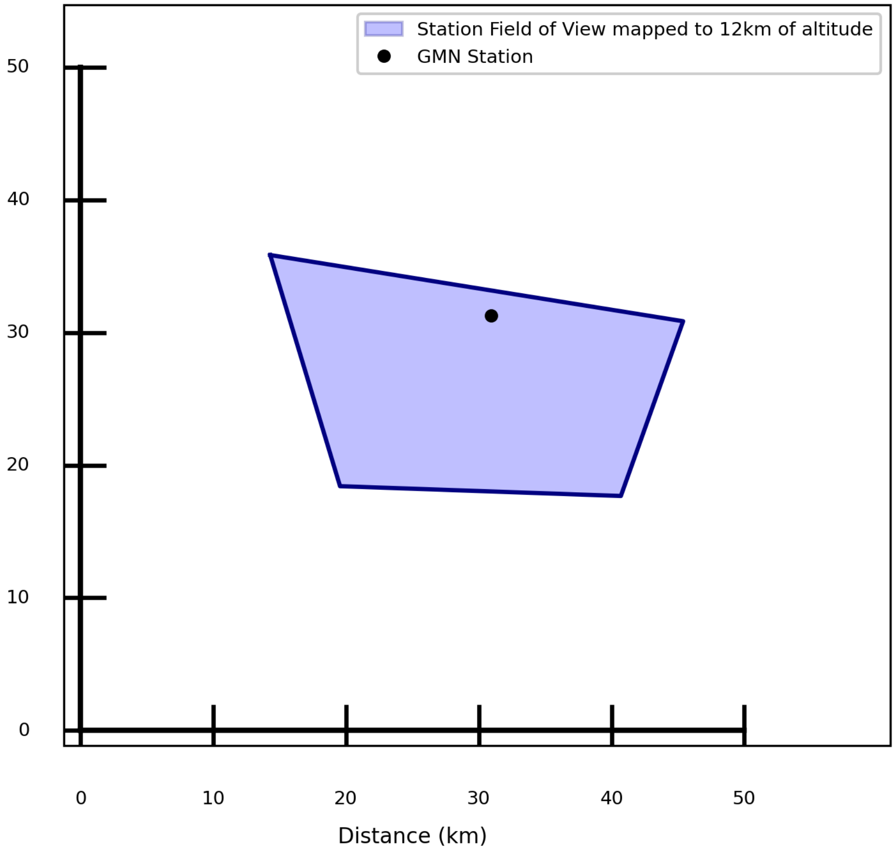

Figure 2: Projected field of view (FOV) for the GMN camera whose footage is shown. Note that distances are shown in km from the centre of the camera FOV to anonymise the location of the camera site.

variety of different atmospheric and air traffic conditions. Many GMN stations also overlap with geostationary satellite fields of view, enabling important co-location between ground- and satellite-based observations. Finally, all GMN cameras have timing accurate to 10s of milliseconds and have a mapped pixel calibration at an accuracy better than 1/3 px (~1 arc minute).

GMN cameras, when operating in continuous capture mode, capture a frame every five seconds, offering the temporal resolution needed to detect contrail formation and observe their early evolutionary phases. For ultra short-lived contrails, this can include even their dissipation (see Sect. 3.1). This temporal resolution, combined with a high spatial resolution (average pixel resolution of 17.27m at 12,000 m of altitude at zenith) supports highly confident flight attribution and quantitative descriptions of contrails in their early lifetimes (see Sect. 2.5). GMN cameras additionally have near-infrared capabilities that facilitate night observations, automatically adjusting the exposure time and gain (brightness) to switch between daytime and nighttime for ideal observing at all times (Fig. 1), This capability is key for observing contrails during nighttime hours, where their warming effects are known to be pronounced. Finally, GMN's standardised camera hardware and open-source software ensures consistent data capture and storage infrastructure, allowing for a generalisable data analysis pipeline that can be applied across stations and times, with station-specific differences being limited to camera tilt and orientation.

GMN data is provided as .mp4 files containing nighttime monochrome and daytime colour timelapses, a JSON dictionary with frame timestamps, and JSON dictionary containing the camera's calibration parameters for the associated observation period.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

2.2 Flight waypoint data

Flight data is an essential component of the flight-first approach that we present in this study. We use Automatic Dependent Surveillance-Broadcast (ADS-B) flight data sourced from Spire as the basis for our automated contrail detection (Shapiro et al., 2023). The dataset implemented is that of Teoh et al. (2024b), a cleaned and re-interpolated version of the Spire dataset that contains flight waypoints at an approximate one-minute resolution, indexed by unique flight ID's, and including aircraft metadata (e.g., aircraft type, tail number, origin-destination pair). We note that the coverage of the Spire-derived dataset is limited to terrestrial sources, although this causes no issues for ground-based observations.

We configure the ADS-B data for per-flight processing based on the ability of the target GMN camera to capture contrails made by each flight. Flights are classified as (i) directly intersecting the camera field of view (FOV), mapped at an altitude of 12km (Fig. 2), (ii) intersecting a padded FOV with a 350-pixel padding, whereby the flight track itself does not directly intersect the FOV but contrail candidates observed within the FOV may be attributed to those flights after advecting their flown trajectories, and (iii) non-intersecting. We refer to flights intersecting the padded FOV as "nearly intersecting" throughout this manuscript. Our analysis pipeline is thus divided into two branches: one branch for directly intersecting flights, whose contrails would form in-frame as the aircraft progresses, and the other branch for nearly intersecting flights, whose contrails would form outside the FOV and advect in. While detection and termination logic varies between the branches, the core skeleton of the analysis pipeline remains similar. The branched flight-by-flight treatment of directly/nearly intersecting flights ensures an exhaustive dataset with inherent negative detections. We define negative detections in an observational sense when conditions are favourable (see Sect. 3.2). Such an approach is vital in our flight-first approach and is a key difference between our dataset and those produced using detection-first approaches, which infer negative detections associated with flights that they are unable to match with detected contrails.

In our processing, GMN imaging frames are batch-loaded, and waypoints are linearly interpolated to align with the five-second resolution of the footage. Given the temporal resolution of the ADS-B data at source, we expect these interpolated aircraft positions to be sufficiently accurate for the needs of our analysis (i.e., within tolerance of other errors). For each frame, waypoints are advected using windspeed data (Sect. 2.3) and functionality within the pycontrails package (Shapiro et al., 2023). The geographic coordinates of the flight waypoints are first transformed into horizontal coordinates (azimuth and elevation) relative to the camera, then are projected onto the image plane to obtain image pixel coordinates (x and y). This mapping was implemented using GMN toolkits (Vida et al., 2021). Once the flightpath (or advected flightpath) enters the FOV, our contrail detection is performed around the path (Sect. 2.4). This process repeats until all flight waypoints are processed or an early termination criterion is flagged (Sect. 2.7).

An additional piece of logic for nearly intersecting flight processing is the calculation of the mean drift towards the frame. At each timestep, the Euclidean distance between every mapped pixel of the advected flightpath waypoint and the frame's centre is computed and the minimum is taken as a proxy for assessing relative motion of the flight track towards the FOV. We describe this as a proxy because the method is not suitable for precise spatial measurements due to projection distortions. The moving average of the 15 most recent minimum distances is calculated, with a positive trend indicating the flightpath is advecting away from the FOV. In such cases, advection is terminated early to avoid unnecessary computation.

2.3 Meteorological data

To account for the dry advection of contrails under the influence of ambient winds, meteorological data is sourced from the European Centre for Medium-Range Weather Forecasts (ECMWF) high resolution reanalysis product (ERA5 HRES; Hersbach et al. 2020). Specifically, the eastward and northward wind components are retrieved at 12 pressure levels between 750 hPa and 150 hPa, spanning typical cruising altitudes of commercial aircraft. Interpolation of gridded data to our advected flightpaths is carried out quadrilinearly in accordance with the methods outlined in pycontrails (Shapiro et al., 2023).

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

2.4 Image segmentation & classification

We utilise Meta's Segment Anything Model 2 (SAM2; Ravi et al. 2024) for automated flightpath-aware contrail segmentation via a dynamic flightpath-informed prompt box. SAM2 is an asemantic zero-shot segmentation model; it does not require task specific pretraining and segments out what it deems to be coherent objects (More information on SAM2 in Sect. S2). The resulting output is a binary mask of background and foreground pixels, without further semantic classification apart from instancing multiple objects.

A limitation of this approach is that SAM2 has no inherent understanding of contrails as a semantic class, and therefore does not specifically target contrails, instead it relies on general visual cues such as contrast and structural coherence. While SAM2 can, in principle, be fine-tuned for contrail-specific detection, this would require a curated hand-labelled dataset and introduce potential bias. This lack of pretraining reduces the risk of bias from a hand-labelled dataset and thus increases generalisability across varied atmospheric and imaging conditions.

We treat this "limitation" as a design choice. While training a computer vision model specifically for contrail detection has already been demonstrated in the literature (Pertino et al., 2024; Huffel et al., 2025; Jarry et al., 2025), we instead focus on integrating a general-purpose and off-the-shelf segmentation model in an automated and physically informed pipeline. We address the potential limitation through upstream processes, such as flightpath-informed prompting (Sect. 2.4.1), and downstream processes, such as mask cleaning (2.4.2) and a classification process for verifying if mask outputs correspond to a contrail or not (Sect. 2.5 & 2.6).

While SAM2 includes an inbuilt video tracker for propagating masks across frames, we do not adopt this feature. Instead, segmentation is carried out independently on each frame on a flight-by-flight basis. The only information carried forward is the advected flightpath, as well as dynamic estimates of the contrail's width and deviation from the flightpath to guide the input box's dynamic padding. This decision reflects the primary purpose of the pipeline being to detect whether a flight has formed a contrail or not. As most flights are expected to not leave a contrail, SAM2 will be asked to mostly segment contrail negative scenes, where the use of a video tracker could risk in propagating false positives, which SAM2 cannot distinguish from true detections due to its asemantic nature. Maintaining independence between frames thus allows for better control through enabling separate single-frame cleaning processes and temporally informed volumetric mask cleaning processes (Sect. 2.4.2).

2.4.1 Input box strategy

As a zero-shot segmentation model, SAM2 is able to accept prompts to focus its segmentation on a specific image region. The input box strategy that we implement uses the axially aligned bounding box, defined as a rectangle whose edges are parallel to the image coordinate axis. We define its corners by the minimum and maximum pixel coordinates of the advected flightpath with some small additional padding. This padding is dynamically adjusted based on the evolving maximum from a deque of the five most recent means of half of the mask width and the distance between the mask's centreline and the advected flightpath, with an additional 20 pixels laterally and vertically.

Several potential pre-processing approaches were considered and discarded, which we outline here both to inform and encourage future research. One approach involved rotating each frame such that the flightpath axially aligns with the bounding box, allowing for a tighter cropping and reduced empty space. However, rotated masks would still require substantial cleaning, particularly for intersecting or curved contrail objects, and could introduce frame-specific distortions. Another considered possibility was an extension to the rotation strategy by locally warping the mapped flightpath to remove curvature, yielding an even tighter fitting input box. However, this type of warping would require detected masks to be remapped to the original GMN camera's coordinate space, adding complexity and risking distorting metrics used for object classification (Sect. 2.5).

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

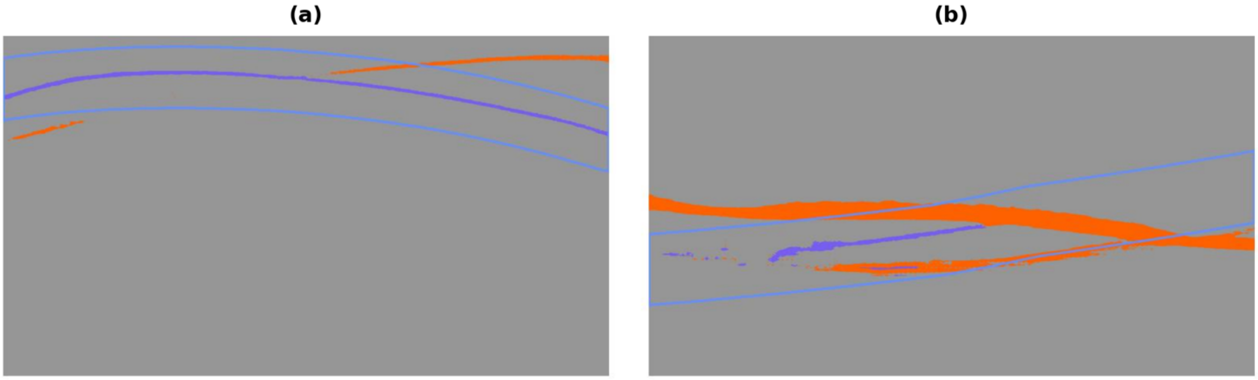

Figure 3: Two examples of the individual mask cleaning process, showing the stepwise decision logic implemented for selecting/rejecting components of the uncleaned mask based on the candidate flight data. Discarded mask components are shown in orange, the contours of the padded flightpath envelope are shown in blue, and the retained mask in purple.

Finally, a third considered option involved manipulating pixel values without any spatial transformations, suppressing all pixels outside of the padded flightpath envelope by setting them to a constant intensity (e.g., mean pixel intensity) to remove any background discontinuities like clouds. The rationale was that reducing distracting background features could help guide the segmentation process more effectively towards the potential contrail, given that SAM2 operates based on local contrasts. Despite the potential benefits in aiding faint contrail detection, especially for advected contrails, we found that such a method can excessively focus the model's input. Removing too much of the global context risks weak irrelevant features becoming dominant enough to be segmented out by SAM2. We thus suggest that a contrail should have sufficient visual strength to be distinguishable from the background without requiring heavy pre-processing, especially when guided by a well-constructed input box.
We find overall, therefore, that sufficient guidance is required to aid the segmentation of the relevant contrail, but not to the extent that guidance artificially enhances irrelevant features. Combining the flightpath-informed and dynamically adjusted bounding box with a more sophisticated post-processing mask cleaning was found to be favourable, rather than overly pre-processing raw footage. However, future work could yet explore more intelligent pre-processing or enhancing SAM2's segmentation process and reducing the weight on downstream post-processes.

2.4.2 Mask cleaning

Both satellite- and ground-based detection methods face limitations associated with fragmented, intersecting, or merged contrails (e.g., Ortiz et al. 2025). The mask cleaning strategy detailed here, enabled inherently by our flight-first based approach, offers a potential solution to these limitations, especially in crowded airspaces. As the pipeline prioritises refined post-processing over extensive pre-processing, initial masks may contain some irrelevant features, which the mask cleaning process addresses directly. If left unfiltered, these artefacts can corrupt downstream assessments of contrail detections, lead to the tracking of false positives, falsely increase a contrail's persistence, or merge distinct contrails into single objects.
To ensure that only the contrail of interest is being processed for each flight, and to guard against segmentation artefacts, a two-stage mask cleaning strategy is employed: the first stage operates on a frame-by-frame basis to enforce spatial coherence with the known flightpath; the second stage only activates for positive contrail detections processing the entire time series of computed masks in a 3-dimensional space of pixel coordinates and time, enforcing temporal coherency.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

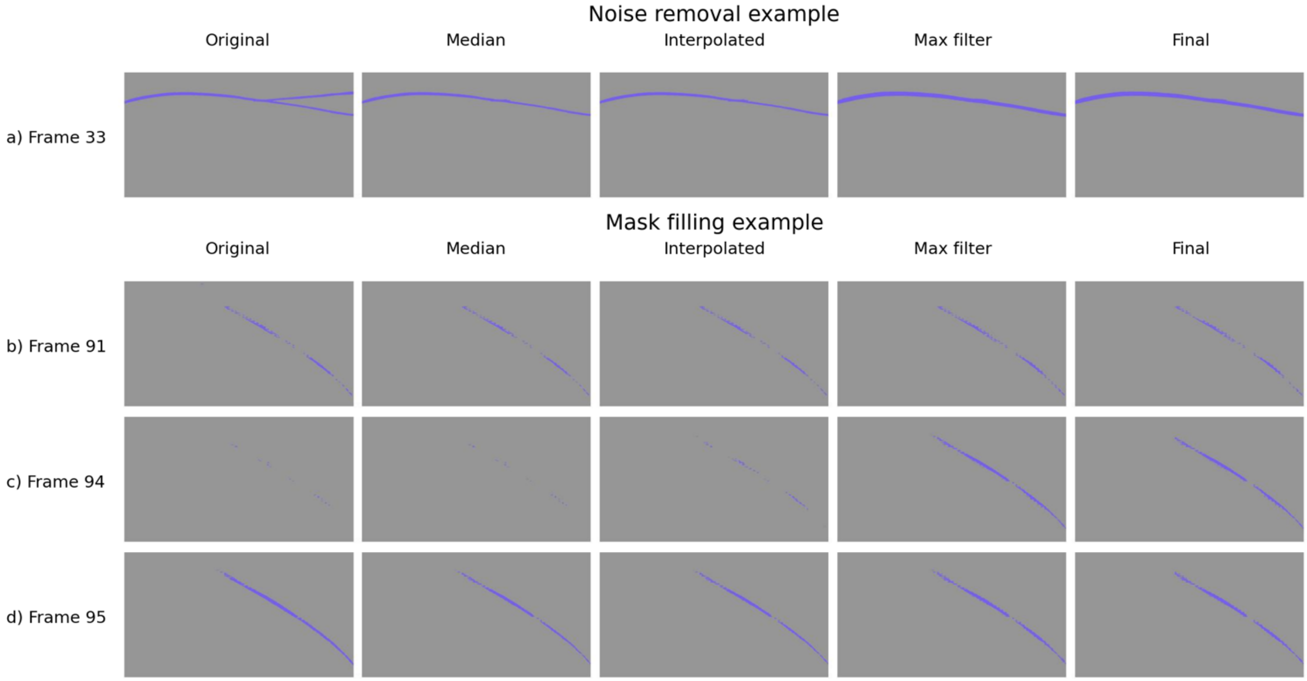

Figure 4: Examples of volumetric mask cleaning techniques, including (a) removal of noise from an intersecting contrail instance via use of a median filter applied temporally, and (b-d) mask gap filling via the maximum filter, with the initially empty frame 94 being flanked by valid frames 91 & 95.

2.4.3 Individual mask cleaning

In the first cleaning stage, each frame is processed independently (Fig. 3). The binary mask is first decomposed into connected regions using connected component labelling through SciPy's label() functionality (i.e., region proposal; Virtanen et al., 2020). Each region is then evaluated separately. Small regions, defined as fewer than 20 pixels during the formation stage and 30 for a formed contrail, are discarded as noise. In parallel, a padded flightpath envelope is constructed by interpolating the mapped flight path and drawing a polyline of a configurable thickness either side of it, initially set to 20 pixels during formation then is based on the evolving deque of the most recent five mask widths and mask-to-flightpath distances with an additional 50 pixels once the contrail has been formed (Fig. 3). This accounts for contrails that widen or drift away from their estimated locations. For advected traces, this padding is increased to account for greater mapping discrepancies stemming from build-ups of windspeed errors.

For each remaining region, the fraction of pixels within the padded flightpath envelope (region overlap ratio) is calculated. Regions with overlap above a set upper threshold (0.7) are kept, and those below a lower threshold (0.4) are discarded[1]. These thresholds were set empirically but can be adjusted based upon background observation conditions or the level of confidence required (e.g., relaxation for advected traces). Ambiguous regions in between these thresholds are further subdivided into subregions via OpenCV's watershedding algorithm (Bradski, 2000) as they display potential overlap between the contrail of

[1] *In one variant of the method, only pixels lying within the padded flight path envelope were kept, reliably excluding outside noise. However, it relied on there being accurate flightpath envelope mapping and padding to not exclude any of the actual contrail or include noise. Moreover, this variant was unreliable for intersecting contrails.*

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

interest and another contrail or cloud (e.g., Fig. 3). This process determines boundaries between subregions by simulating a "flood" from high-confidence deep- interior seed pixels (i.e., "peaks") that are identified by a normalised distance transform. The subregions are then kept or discarded against the upper overlap threshold. This watershedding algorithm is also invoked for large regions exceeding 1000 pixels, which did not get discarded based on the lower threshold but commonly have multiple branchpoints.

2.4.4 Volumetric mask cleaning

Once a contrail stops being actively tracked (i.e., segmentation is terminated; see Sect. 2.7), corresponding individual masks for that flight are stacked into a three-dimensional mask volume, where time serves as the third dimension. A series of temporal mask cleaning operations is then carried out (Fig. 4).
The first operation applied to each mask volume is the removal of flickering noise via a one-dimensional median filter applied along the temporal axis (Fig. 4a). For each pixel, a sliding window of three consecutive frames is used to compute the median pixel value and set it across those frames. As this filter is also able to turn on pixels, which was found to introduce unwanted noise, Boolean conjunction between the original and median filter mask is performed to ensure pixels can only be turned off. As a result, pixels that are only turned on once out of three consecutive frames are suppressed. We note that this stage is only applied after a contrail has formed as to not corrupt formation masks.
The second operation involves interpolating over brief temporal dropouts to compensate for transient failures of SAM2 to accurately capture the contrail or any temporal obscuration. A time series is considered for each pixel and interpolated by activating "off" pixels if they are temporally flanked by "on" pixels, provided the temporal gap between them is no more than three frames (only counting in between frames), which was set empirically through case studies. This mask interpolation operation is supported via a one-dimensional maximum filter applied temporally at a window size of three frames. The maximum filter turns on all the pixels across the sliding window if there is at least one "on" pixel present (Fig. 4 b-d). Because it follows the median filter in operation order, which already suppresses transient flicker noise, the maximum filter does not introduce new noise. Instead, it extends valid voxels by propagating "on" pixels to adjacent frames. This further bridges gaps between frames and smooths out the capture of the contrail's evolution.
Despite these filtering operations being carried out, some spatiotemporally persistent noise may remain, such as background discontinuities, clouds, or other contrails. Our final operation thus involves removing persistent noise by labelling the mask volume into distinct clusters (i.e., voxels) that are discarded if they are below a minimum pixel count threshold, empirically set to 800 (Fig. 4 b-d). The resulting clean series of masks provides input for the computation of metrics for object classification and reduces the risk of corruption from artefacts or other contrails.

2.5 Computing metrics for contrail classification

Metrics describing the spatial and textural characteristics of each segmented mask are computed to distinguish genuine contrails from spurious detections and to support tracking their evolutions. The flight-first approach enables these metrics to be computed directly from the segmentation process rather than post hoc downstream of the segmentation. During the formation stage, when the flight is in frame, frames with absent masks, either directly from SAM2 or after mask cleaning, are accounted for by calculating an empty mask ratio. All metrics are stored in JSON dictionaries and recomputed again after the volumetric mask cleaning, with the updated metrics logged separately to be used for further analysis,
Both the mean and standard deviation, taken over the mask centreline, of our selected metrics are computed to capture the central tendency and variability of each feature. The contrail width, determined via skeletonization and a Euclidean distance transform, is particularly informative for formation detection as young contrails typically exhibit a narrow and uniform profile that widens over time. The distance between the mask centreline and the flightpath, also computed with a distance transform, assesses the alignment between the detected object and the mapped flightpath. Genuine contrails are expected to follow the

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

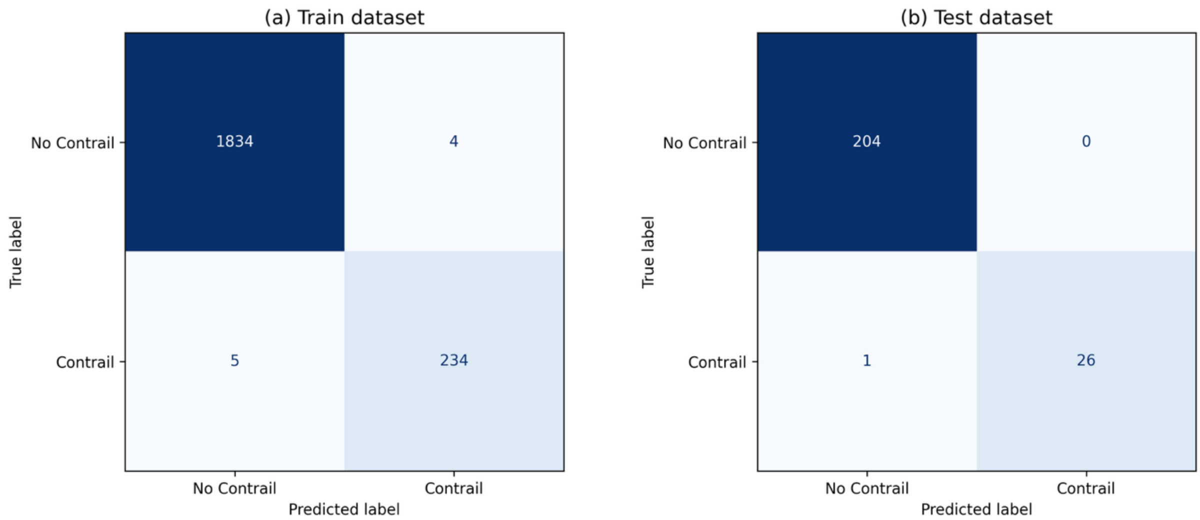

Figure 5: Confusion matrices for (left) the train dataset (Accuracy: 1.00, Precision: 0.98, Recall: 0.98) and (right) the test subset (Accuracy: 1.00, Precision: 1.00, Recall: 0.96) for the formation Random Forest model at the selected threshold of 0.35 (Sect. 2.6). The accuracy, precision, recall scores given as per standard definitions. Flights without valid masks are automatically labelled as negative and are excluded from training and testing, leading to better performance in practice.

flightpath closely, with only minor wind- and projection-related discrepancies, whereas a larger distance should flag irrelevant linear structures such as contrails from other flights or cirrus clouds. The pixel intensity, reflecting both the brightness and internal consistency of the mask, is also informative as genuine contrails will tend to have relatively bright and uniform pixel intensities. The ratio of the mean mask pixel intensity over the mean background pixel intensity serves as an additional check as contrails should appear brighter than their background, with a value below unity potentially indicating the presence of an irrelevant feature.

Finally, segmentation quality is evaluated through the number of valid regions retained after the mask cleaning, and SAM2's internal confidence score. A high number of valid regions indicates a fragmented and noisy detection, whereas a low number indicates more cohesion. Similarly, a high mask score indicates a strong and well-defined object, like a contrail, and a lower score indicates poor segmentation, which likely occurs when there are no well-defined objects present and SAM2 falls back on segmenting weak and/or diffuse background features. Finally, the number of pixels is recorded as, during the formation stage, genuine contrails should approximately have the same number of pixels as one another given that they have similar initial widths and are at similar distances away from the camera. To summarise, the computed metrics are the means and standard deviations of the mask width, max pixel intensity and the distance, the mask size, mask score, ratio of empty masks, number of valid regions, and ratio of the mean mask pixel intensity over background intensity.

2.6 Classification of contrails from metrics

For detection-first methods, establishing positive detections of contrails can be viewed as a semantic segmentation problem, with the objective of segmenting out contrail instances from the whole image frame. For flight-first methods, however, it becomes a classification problem, with the asemantically produced masks (from SAM2 in our case) only requiring classification of the found objects into positive and negative contrail detections. To automate this classification, we implement

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

a Random Forest classifier model (RFM), with one model each for flights directly and nearly intersecting the camera FOV. Both models were created using the scikit-learn library (Pedregosa et al., 2011). Initially, a Linear Regression model was also trained as a simpler alternative, but it achieved inferior performance and hence was not chosen. We suggest that simple RFMs provide a balance between complexity and capability.

An RFM is a machine learning model capable of capturing non-linear relationships by creating an ensemble of decision trees, where each tree is trained to create a split decision that maximises the purity of the outputted groups (e.g., contrails and non-contrails).

The RFM for directly intersecting flights uses the mean mask metrics (Sect. 2.5) derived from the formation stage, which are the frames encompassing the aircraft entering and exiting the frame. We label this RFM as the "local formation" classifier. The use of mean mask metrics accounts for empty masks by only utilising valid masks for the mean calculation and including the ratio of empty masks as an additional feature (see Sect. S4 for further information). This RFM was trained and tested on a curated dataset, representing 65% of the full dataset from the three-week case study (see Sect. 3.2),representing 2,308 flights out of a total of 3,506.The model was trained with 90% of this data subset and tested with the remaining 10% (Fig. 5). While the skewed train-test split may introduce a risk of overfitting, such a split was required to provide sufficient contrails for training, and comparison of accuracy, precision, and recall between training and test subsets suggests that overfitting is unlikely (Fig. 5). The positive cases for the RFM training included both short-lived and intermediate/persistent contrails. Cases where SAM2 failed to meaningfully segment out a human-labelled contrail (63 cases) were excluded from the training and testing of the RFM as the resulting mask metrics would not provide a reliable representation of the contrail's features and could introduce noise and degrade training.

The RFM's isolated evaluation thus reflects the performance achievable under the assumption that meaningful mask metrics are available, providing an estimate of the performance of the RFM independent of segmentation performance. However, the RFM is only one layer in the pipeline (SAM2 being another), and the purpose of the study is not to prescribe individual components, but to demonstrate the framework as a whole. This framework as whole is evaluated on the full case study, resulting in an overall 61-to-39% train-to-test split.

Flights that left no valid mask output, either directly from SAM2 or after mask cleaning, are automatically classified as negative detections by the wider pipeline and are hence also excluded from the training of the RFM. This confirmation of a negative detection is an inherent advantage of the flight-first approach in comparison to a detection-first approach, where negative detections may also arise from the possibility of an unsuccessful flight match. It should, however, be noted that the absence of a contrail from a ground observer's perspective does not strictly guarantee its true absence, for instance a contrail forming above a layer of lower-level cloud or against a layer of higher-level cloud (see Sect. 3.2). A pre-screening filter that characterises observation suitability would thus be a useful additional feature for flight-first approaches.

In a parallel processing stream, the RFM for nearly intersecting flights (labelled as the "advected" classifier) uses the mean metrics from the first 20 frames after five interpolated waypoints of the advected trace enter the field of view, accounting for likely mapped flightpath discrepancies due to wind errors. Only 33 such advected contrails were identified over the three-week

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

case study period (Sect. 3), despite the use of a 350-pixel padded region around the FOV. This is likely due to a low likelihood of a contrail forming outside the FOV, advecting into it, and persisting long enough to not dissipate yet remain visually distinguishable, i.e. not lost its linearity. Although the advected classifier showed promising performance on its full dataset, the low number of detected advected contrails limits further statistical analysis, and hence only the formation classifier is presented in subsequent analysis.

The RFM then outputs a probability score based on the means of the mask metrics collected over the aforementioned frame sequences, representing the probability of the segmented feature from the flight's masks corresponding to a contrail. The probability threshold of the RFM outputs, that being the score at which a detected object is deemed to be a genuine contrail, can be adjusted based on the operational priorities of the contrail detection task. For instance, a lower threshold would prioritise sensitivity to assure more contrails are detected at a cost of more increased positive detections (see Sect. 3.1.2). Correspondingly, a higher threshold would prioritise precision to yield a more robust and reliable set of detections at the expensive of potentially missing more ambiguous cases. This threshold may also be adjusted dynamically in response to observational conditions, such as setting a lower threshold for certain background scenes (e.g., cloudiness, airspace complexity). For the formation classifier RFM, the threshold was set at 0.35 based on Youden's J-statistic that seeks to balance sensitivity and specificity (Youden, 1950). Future work could be directed at dynamically determining suitable thresholds based on scenario and observational contexts.

Once the RFM outputs a probability exceeding the selected threshold, the corresponding contrail begins to be tracked (Sect 2.7), without in a tracking and terminating process independent of the RFM.

2.7 Tracking and termination

Under the flight-first approach, each contrail is tracked simply by updating the input box with the coordinates of the mapped flightpath with dynamically adjusted padding. The tracking is then terminated based on three criteria (i) if the flightpath advects out of the camera FOV, (ii) after an absolute maximum tracking time, currently set to 30 minutes after the aircraft exits the frame, to safeguard against excess computation and prolonged tracking of spurious noise, or (iii) if a specified number of masks within a time window are classified as empty.

For (iii), the content and quality of the most recent masks are recorded using a fixed length deque. A mask is flagged as "empty" under three conditions: If SAM2 produces no raw mask output, if there is no mask output left after the cleaning stage, or if its confidence score is below an empirically set threshold of 0.05. The fixed length deque is currently set to store the last six frames for contrails detected by the local classifier and the last 10 frames for contrails detected by the advected classifier. These values were determined through examining case studies of detected contrails. A lower value may result in the premature termination of tracking and over-punishment of genuine contrails with mask gaps that would ultimately be filled in after volumetric mask cleaning. On the other hand, too high of a threshold unnecessarily prolongs tracking and artificially extends the estimated contrail lifetime.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

2.8 Categorisation of contrail using tracked lifetimes

We categorise detected and tracked contrails into three classes, in keeping with Low et al. (2024). We define short-lived contrails as those that dissipate within 2 minutes of forming. We define persistent contrails as those that are detected for longer than 10 minutes. Correspondingly, contrails that last between 2 and 10 minutes are categorised as intermediate.

We perform this categorisation based on the contrail "mean age", which is calculated by tracking the time since formation for each interpolated waypoint, and at each frame taking the mean of the subset of in-frame points. This provides a measure of how long the in-frame portion of the contrail has existed for. This method also generalises well to advected contrails and ones that move across the frame, as it tracks the ages of waypoints from their time of formation before that segment of the contrail enters the field of view, providing an objective system for contrail classification based on age. For contrails that persist and advect out of the field of view, the maximum observed "mean age" will be shorter than the total contrail lifetime. However, this does not significantly affect the classification framework, as contrails typically form across a substantial portion of the field of view and advect gradually, resulting in observable durations that exceed the 2-minute threshold for a short-lived contrail. The classification is therefore robust for distinguishing short-lived contrails from ones with sustained persistence (see Sect. S3 for further explanation).

Objective lifetime-based contrail categorisation is important for providing validation for meteorology as short-lived contrails are indicators of contrail-forming conditions being met without persistence conditions, while longer-lived contrails indicate the additional presence of these persistence conditions. It is also important for evaluating a contrail's potential for further analysis (e.g., observation from geostationary satellite, assessing climate impact).

3 Results and discussions

Using our pipeline of input box construction, SAM2 segmentation, mask cleaning, metric calculation, RFM classification, and tracking, we collected footage from GMN's US0001 station located in Albuquerque, New Mexico over a three-week period in February 2025. This camera and time period have been chosen for the case study as it was qualitatively noticed that they contain many visible contrails. Of the 21 days analysed, seven were deemed to have poor contrail detection conditions due to limited formation or heavy cloud coverage. The pipeline was still run on two of these poor detection days to test the robustness of the pipeline under less favourable conditions. There are thus 16 total observation days, encompassing 3,506 flights that directly intersected the camera's FOV, and over 10 thousand advection candidate flights. These 16 days form the case study that we present here.

A manual review of all the collected footage was used to establish a ground truth of negative detections and flights that produce intermediate or persistent contrails. The review process consisted of identifying contrails in the footage and attributing them to flights through a temporal comparison and geometrical comparison with ADS-B data overlaid onto the frame. The pipeline was evaluated on both the full dataset of 3,506 flights and further on a subset representing flights that were not included in the training of the RFM("unseen" data). 1,378 flights were included in this unseen subset, making up 39% of all data, including

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

the days of the 31/01/2025 and 20/02/2025, the test dataset for the RFM, cases without meaningful segmentation output, and other cases.

Visual examples of segmentation are provided in Sect. 3.1, the overall pipeline is evaluated in Sect. 3.2, and potential uses of the pipeline are discussed in Sect. 3.3, including a comparison to meteorology and modelling.

3.1 Case studies of segmentation and detection

To illustrate the visual performance of the analysis pipeline under a range of conditions, we present here a selection of examples from our case study period. In each example, the advected flight path is plotted in blue, the input box in green, and the cleaned segmentation output is overlaid as a white contour. These examples not only demonstrate successful contrail detection but also the pipeline's ability to accurately segment contrails as they evolve temporally and spatially, highlighting the robustness of the two-stage mask cleaning process.

3.1.1 Locally formed contrails

We find that under clear sky conditions, which are the most common for the location of our chosen station, the pipeline performs visually and statistically well both for persistent (Fig. 6) and short-lived contrails (Fig. S1 and S2). We note that performance is retained during the night (Fig. S3), as well as during the day (Fig. 6). This speaks to the potentially scalability of the flight-first approach, in particular through leveraging the standardised nature of the GMN but also hints at a limitation of ground-based camera networks in that their use depends strongly on observation conditions.

We observe that in addition to handling constant-width contrails with relative ease (Fig. 6), the flight-first approach can handle more complicated cases of width-evolving contrails or ones that intersect other contrails (a use-case in which other methods may struggle; Fig. 7). The pipeline maintains consistent focus on the contrail of interest, allowing continuous tracking despite the presence of overlapping contrails and in spite of the target contrail becoming nonlinear and dispersed, highlighting the effectiveness of the flight-first approach's dynamic bounding box and mask cleaning process. This is a particularly important feature for establishing ground truth observations in regions with dense air traffic, where other observation sources struggle with flight-matching fidelity. Conversely, the pipeline correctly and inherently identifies flights that do not leave a contrail as true negatives (Fig. S3).

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

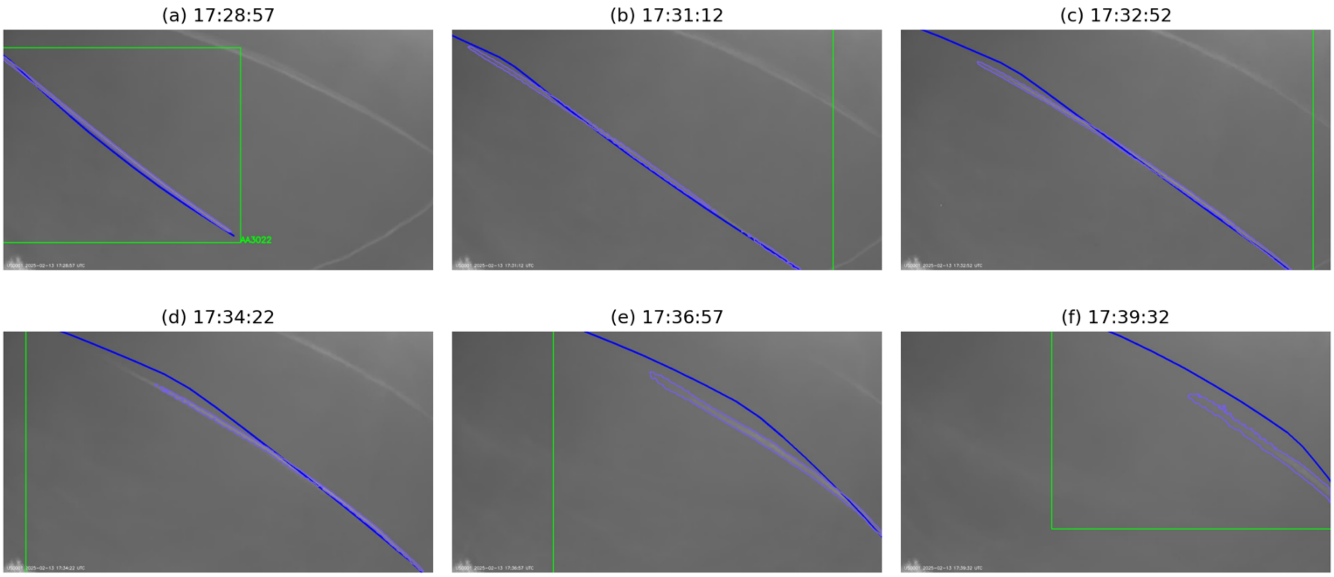

Figure 6: The overlaid masks from the common case of a persistent and consistent width (non-growing) contrail segmented successfully against a homogenous sky background. The advected flightpath is drawn in blue, the input box in green, and the mask is contoured in purple.

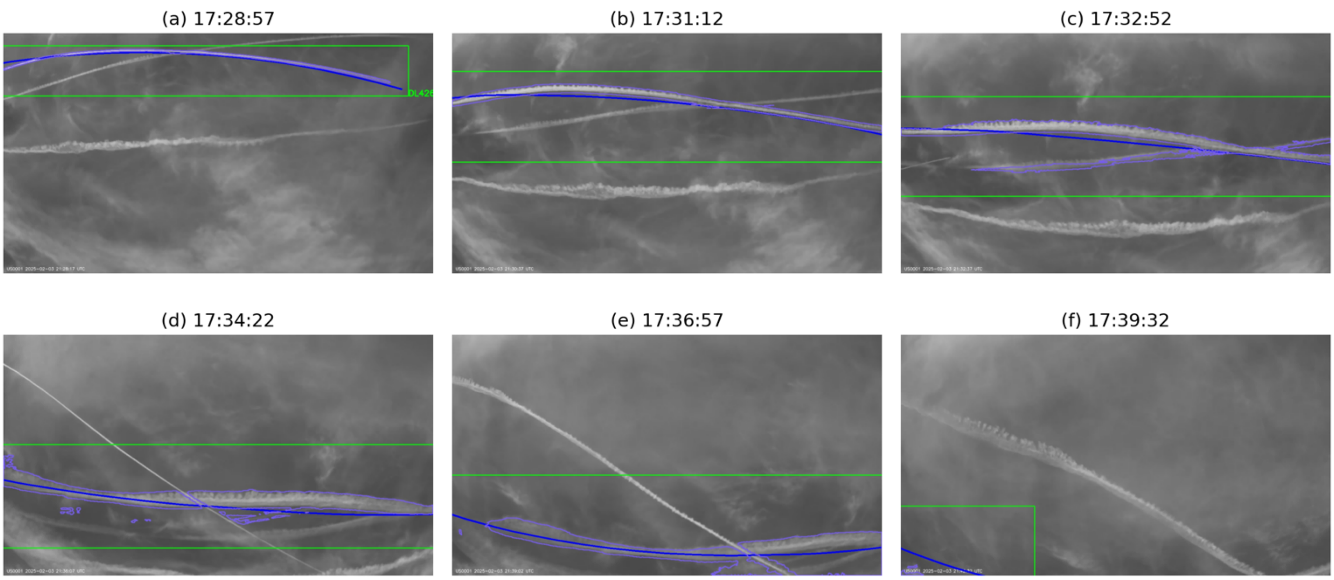

Figure 7: The overlaid masks from a width-evolving persistent contrail that is intersected by other contrails and is against a complicated background. The contrail is correctly tracked and segmented, highlighting the effectiveness of the mask cleaning process. The advected flightpath is drawn in blue, the input box is in green, and the mask is contoured in purple.

To illustrate where the pipeline faces challenges and limitations, a set of false negative examples have been analysed reflecting common examples of unsuccessful segmentation (Fig. 8 and Fig. S5-S7). A key limitation stems from the use of SAM2, whose asemantic and general-purpose nature means that it relies mostly on local contrasts and is not trained to detect faint contrails, especially in complex scenes. Since the RFM was only trained on flights with viable SAM2 segmentations, these weak

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

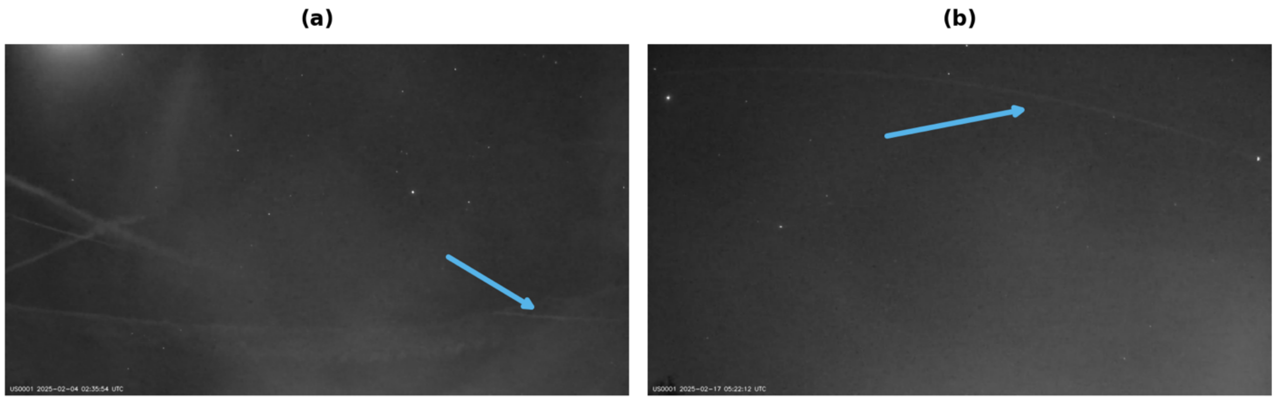

Figure 8: : (left) An example of a missed contrail due to being partially obscured by cloud cover, with arrows indicating its location. (right) An example of a missed contrail caused by SAM2 failing to segment out the contrail due to low local contrast and the contrail thus appearing faint.

detections were not used for training, leading to this blind spot in the pipeline. This reflects the trade-off with using a generalisable and off-the-shelf model. Despite these limitations, the pipeline still achieves a precision of 0.99 and recall of 0.74 (Table 1), indicating that these failure modes are relatively uncommon within the case study. The advantage of the flight-first approach is that we are able to pinpoint the underlying causes for contrails being missed in the segmentation bounding boxes around the advected flightpaths, helping us to understand the conditions under which our approach can be best utilised. Such investigation is more challenging to perform precisely using detection-first methods as they rely on less interpretable semantic segmentation of the global scene within each frame. Our false negatives generally fall into six non-mutually exhaustive categories[2] (i) peripheral contrails (Fig. S5), (ii) partially obscured contrails (Fig. 8 (left) and S6), (iii) visually faint or low-contrast contrails (Fig. S7), (iv) severely complex cloud backgrounds, (v) flight path discrepancies (Fig. 8 (right) and S8), and (vii) anomalous sources of bright pixel patches, such as solar/lunar presence or lens flares.

3.1.2 Advected contrails

Similarly to the locally formed contrails, the performance of the advected traces pipeline is demonstrated with examples of successful segmentations (Fig. 9) that is retained at night as well (Fig. S8). However, we do not find sufficient examples of advected contrails coming into the camera FOV as to conduct a thorough analysis of our pipeline. One of the challenges with accounting for advected contrails is the accumulated windspeed error, which leads to discrepancies in the mapped advected flight path (Fig. 10). This can be partially combatted is with higher padding for the input box and mask cleaning flight envelope. However, excessive padding may lead to unrelated features being segmented before the contrail advects into the field of view or other features in the larger input box and cleaning envelope.

---

[2] *We note that a small number of flights lacked ADS-B data entirely, mostly military flights, or showed discrepancies pre-advection.*

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

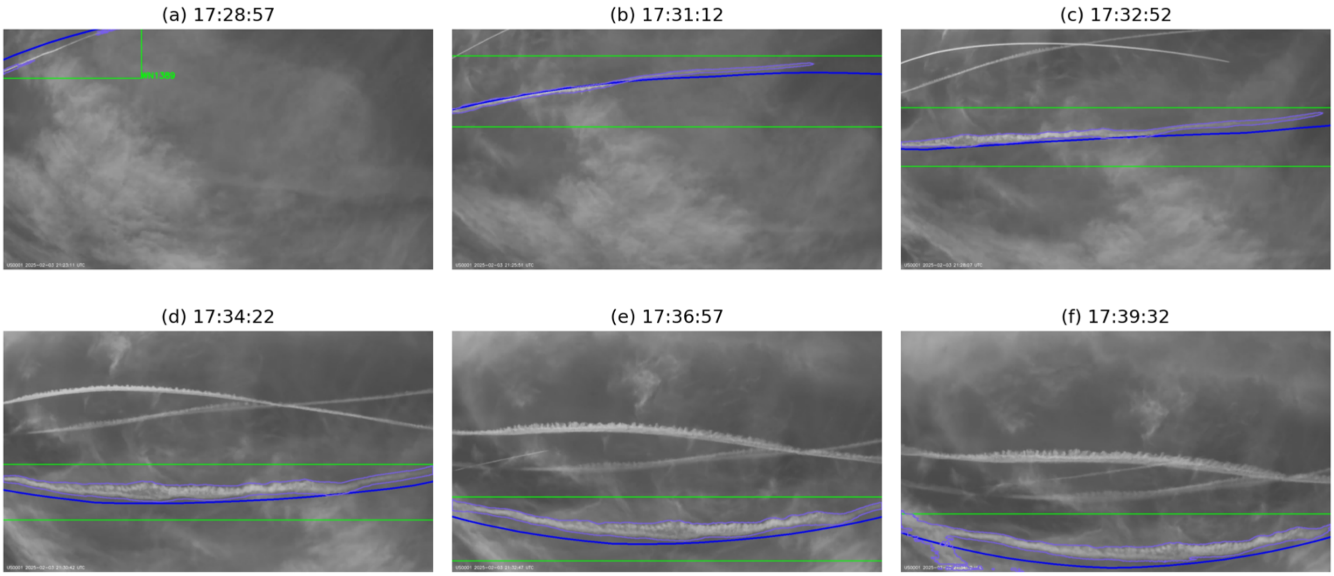

Figure 9: The overlaid masks from a width-evolving persistent contrail that formed outside of the field of view, advected in after 110 seconds, and was present in the frame till the mean age of its visible trace was 890 seconds. The advected flightpath is drawn in blue, the input box is in green, and the mask is contoured in purple.

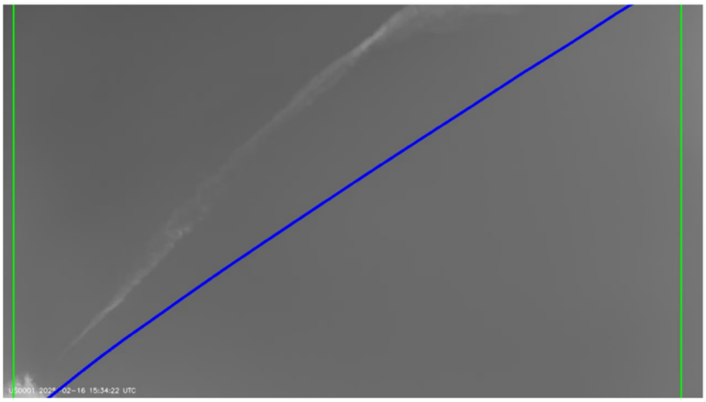

Figure 10: Overlaid masks from a missed advected contrail mask, caused by a large discrepancy with the mapped advected flightpath. The advected flightpath is drawn in blue, the input box is in green, and the mask is contoured in purple.

3.2 Overall pipeline performance

To evaluate the performance of the pipeline, we focus on and compute accuracy metrics for detecting intermediate and persistent contrails only as these represent the largest contrail-derived contributions to aviation's climate impact and are the most likely to be detected by other observation sources. While short-lived contrails are relevant as indicators of the Schmidt-Appleman

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

Table 1 : Summary of detections from the three-week case study period, including detections of short-lived, intermediate/persistent, and advected contrails. Values are divided into the complete labelled dataset, those used in RFM training, and those used in testing. See footnote for definition of MCC.

Statistic	Values for full dataset	Unseen Subset	RFM Training Subset
Short-lived locally formed contrails	75	-	-
Intermediate/ Persistent locally formed contrails	190	64	-
Advected Contrails	33	-	-
Total Contrails	298	64	193
True Positives (intermediate/persistent)	190	32	158
False Positives (intermediate/persistent)	2	2	0
True Negatives (intermediate/persistent)	3,247	1,312	1,935
False Negatives (intermediate/persistent)	67	32	35
Recall for intermediate/persistent	0.7393	0.5000	0.8187
Precision for intermediate/persistent	0.9896	0.9412	1.000
MCC for intermediate/persistent	0.8462	0.6760	0.8967

criterion (SAC; Schumann 1996) being fulfilled, we propose that tailoring the analysis and the model development towards detecting intermediate and persistent contrails better utilises the opportunities that ground-based camera observations represent , which are fit for the purposes of this study's objectives (Sect. 1.2). Modifying the pipeline to better detect short-lived contrails can be explored as future work.

Accuracy metrics were also only calculated for intermediate/persistent contrails over the entire dataset that included no contrail, short-lived, and intermediate/persistent contrails. We define true negatives in an observational sense as flights for which both a human observer and the detection pipeline did not detect an intermediate or persistent contrail. We acknowledge that this definition does not guarantee the absence of a contrail as contrails may be obscured by cloud cover. This observational limitation is common to ground-based observations methods, where detectability is constrained to visual observability. However, the impact of this limitation may be managed by restricting observations to periods with favourable observing conditions with an appropriate pre-processing step. We include Short-lived contrails in the negative category for the calculation of accuracy metrics. Successfully detected short-lived contrails (75 flights) were subsequently classified as short-lived by the lifetime classification stage and therefore we classed them as negative intermediate/persistent detections.

We evaluated the pipeline's performance over both the full three-week case study period and on the subset of data that was unseen by the RFM, representing 39% of the full dataset (Table 1). The training and testing data for the RFM was curated (Section 2.6) and hence the unseen subset does not represent a random sample of the data. Instead, it includes flights that were purposely left out of the RFM data subset, such as segmentation failures, which explains its lowered recall of 0.500 that was influenced mostly by the performance of the segmentation model.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

Out of the two poorer contrail-observability days, one day yielded only true negative detections, while three short-lived contrails were detected on the other day (rest true negatives). We suggest that this demonstrates the model's ability to maintain performance in unfavourable conditions without yielding a high number of false positives, although application of our methods to a larger observation period would be required to confirm this.

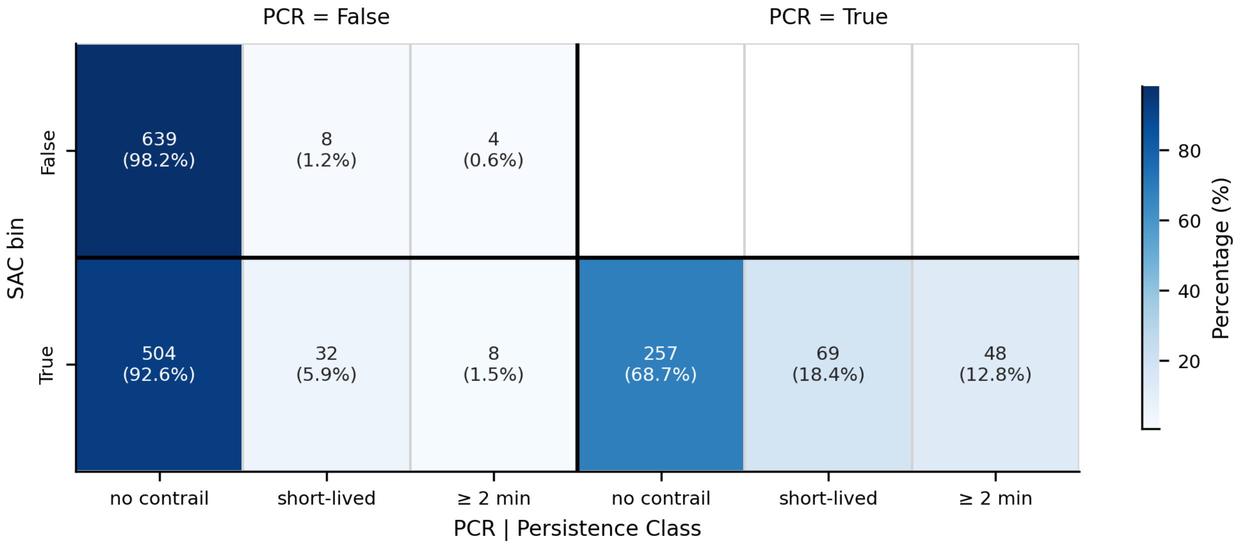

Figure 11 : Table of flight counts and percentages as a function of SAC/PCR criteria (meteorological data) and Persistence Class (ground-based camera observations), coloured by percentage. Percentages are computed within each SAC and PCR combination. For example, 69 (18.4%) flights within regions fulfilling both SAC and PCR were observed making short-lived contrails. Note the bin for SAC False and PCR True flights is empty as PCR conditions require SAC to be satisfied first.

Despite a skewed dataset (i.e., only 7.4% of flights produced an intermediate/persistent contrails) the pipeline is able to characterise both true positives and true negatives, indicated by a strong MCC score of +0.85[3]. Although not enough for further statistical analysis, the 33 advected contrails detected demonstrate that our "flight-first" pipeline can inherently distinguish between contrails formed within the frame and ones that advect in, which is a post hoc step for "detection-first" methods, which can attribute advected contrails to flights that directly intersect the FOV. Our true negative detections demonstrate the inherent negative detections of our pipeline, that contribute to our high-fidelity dataset. Labelling flights that do not produce contrails is also a post hoc step for "detection-first" methods.

3.3 Uses for a flight-first contrail detection dataset

The flight-first approach to contrail detection inherently includes both flight matching and negative detections. This compares to detection-first methods, where these are either post-processing steps or assumptions. We demonstrate here a key use case

[3] The MCC score is used to evaluate the performance of predictive models on skewed datasets. A score of +1 represents perfect prediction, 0 random guessing, and -1 indicates inferior performance to random guessing.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

enabled by the inherent properties of the flight-first approach using the dataset generated in this study, highlighting the potential for this approach to be scaled to generate wider dataset.

3.3.1 Comparison to meteorology & modelling

Detections of contrails within ground-based cameras can be used to evaluate predictions of contrail formation and the underlying meteorology on which they are based. All categorisations of contrails that we defined in Sect. 2.8 (short-lived, intermediate, persistent) undergo a formation stage, which can be captured under the high spatial and temporal resolution of the GMN cameras. The inherent flight matching of the flight-first approach then allows for interpolation of meteorological data to the accurate and precise location at which the formation occurred (given accurate ADS-B data). Using this process, we evaluate the predictions of contrail formation within our dataset using ECMWF ERA5 HRES meteorological data. For each flight, we simulate whether the portion observed in the camera FOV is expected to form a contrail (i.e., fulfilment of SAC) and whether the resulting contrail is expected to persist (i.e. SAC fulfilment within an ice supersaturated region (ISSR), flagged as a persistent contrail region (PCR)). Figure 11 summarises the distribution of flights across SAC and PCR classification, further separated into observed classes of no contrail, short-lived, and intermediate/persistent contrail grouped into one class (as per definition in Sect. 2.8).

We acknowledge that a meteorological comparison that is solely limited to a ground-based source is bound to have bias as a ground-based source cannot capture all contrails, e.g., some may be obscured by cloud cover. Similarly, other sources such as satellite-based observations cannot capture all contrails (Driver et al., 2025). Hence, ground-based observations should therefore be complemented with other sources in order to get a holistic picture for a meteorological and model comparison. Moreover, assessing the underlying causes for errors in formation prediction, such as errors in specific meteorological quantities in the forecast product (e.g., temperature, pressure, humidity) or errors in aircraft properties (e.g., engine efficiency, emissions indices), requires a more extensive dataset than presented and statistical methods.

We observe asymmetric agreement between observed contrail formation and flight segments located in SAC satisfying regions. Formation recall is strong, with 92.8% of contrails forming in SAC-positive regions. However, formation precision is weak as only 17.1% of SAC-positive flight segments resulted in an observed contrail (Fig. 11). This indicates an overidentification of contrail formation by the meteorological product. It has to be noted that these results only represent observable contrails (see Section 3.2).

Persistence is defined here in accordance with Sect. 2.8, based on the observable tracked lifetime of contrails within the camera field of view. As aforementioned, this does not significantly affect the classification framework. However, this study is limited by the use of a single ground-based camera, as contrails may advect out of view before their full physical lifetime is observed. Thus, while we can offer an analysis of contrail persistence within the classification framework, enabling the assessment of meteorological forecast predictions of ISSRs, observations extending well beyond the initial persistent phase are constrained by the single camera setup. There are currently no datasets based solely on ground-camera observations that assess whether the observed contrails will persist sufficiently long to become visible in other observations or have a substantial climate effect.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

Such observations may be possible with the GMN network in areas of high camera density where a contrail may be tracked across multiple FOVs . Alternatively, all-sky cameras with much wider FOVs could be utilised to extend the observable lifetimes of contrails while still using a single camera.

We observe asymmetric agreement between observed contrail persistence and meteorological conditions indicative of the presence of a PCR. We find recall to be relatively high, with 76.7% of persistent contrails occurring within PCRs. However, precision is low, with only 12.8% of PCR flight segments resulting in a persistent contrail. Among contrails formed within PCRs, 41.0% were observed to be persistent. This indicates an overidentification of persistence by the meteorological product. This bias may be compounded by the disparity of the ERA5 forecast resolution of 31km with the finer scale variability of atmospheric conditions. For instance, ISSRs were found to have a median pathlength of 3 km (Reutter et al., 2020), which is substantially below ERA5s spatial resolution.

Contrail models such as CoCiP are able to output estimates of the contrail age and time-resolved radiative forcing. Future work could explore using metrics that we use during classification (captured at formation) as proxies for radiative impact and estimated lifetime. If a contrail's persistence can be reliably predicted from its formation visual characteristics, paired with meteorological data, the functional scope of ground-based camera networks would be extended. This could also enable more accurate retrospective analysis, estimating how long ago an advected contrail was formed. We do not pursue this here, however, due to the limited scale of our dataset.

Finally, in our analysis, we readily observe accumulated windspeed errors (Fig. 10), a phenomenon that is well-known within the contrail observation community. However, future work could invert the problem, leveraging the fine resolution of flight-matched contrail observations across multiple timestamps to evaluate forecast products' predictions of windspeeds at cruise altitudes and even potentially provide data for assimilation into NWPs or similar tools.

4 Conclusions

We present here an automated flight-attributed pipeline for the detection, tracking, characterisation, and classification of both newly formed and advected contrails using footage from the Global Meteor Network (GMN). A key achievement is the adaptation of the asemantic and general-purpose SAM2 segmentation model into a contrail-aware system through the implementation of a dynamic input box strategy, individual frame and volumetric flightpath-aware mask cleaning frameworks, and a Random Forest classifier model built on top of metrics characterising contrail and image properties.

A novel feature of the pipeline is its flight-first nature, which leads to contrails detections being flight-attributed by default and further leads to inherent negative detections, both of which are post-processing or assumptive steps in detection-first approaches to ground-camera observations of contrails.

The formation contrail Random Forest classifier model (RFM) achieved a precision of 1.00 and a recall of 0.96 on the subset of data for its testing (90-to-10 train-to-test split from a curated set of 2,308 flights). The training and testing subset for the RFM was curated by excluding hand-labelled contrails for which SAM2 failed to produce meaningful segmentations, as these

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

would not yield representative mask metrics for the RFM. This evaluation thus reflects the classification performance of the RFM independent of the segmentation stage.

The most influential classification features in determining whether segmented features are genuine contrails being (i) mask quality metrics, such as object size, segmentation model score, or ratio of empty masks, and (ii) spatial and textural coherence metrics, such as mean and standard deviation of the width or standard deviation of the mask pixel intensity. A second model was also developed for advected traces, however, relatively few contrails were able to be observed in the examined duration of camera footage, preventing detailed statistical analysis as there were only 33 detected advected contrails out of 10,000 candidate flights for such contrails.

We deployed the analysis pipeline on footage from a period of three weeks in February 2025 for one GMN camera located in Albuquerque, New Mexico, involving 3,506 candidate flights for observing direct contrail formation and over 10,000 candidate flights for observing contrails advected into the camera FOV. Our analysis pipeline successfully identified 75 short-lived and 190 intermediate/persistent locally formed contrails, for which a full-dataset recall of 0.74 and a precision of 0.99 was achieved against a hand-labelled dataset (61-to-39 train-to-test split). A precision of 0.94 and a recall of 0.50 was achieved for the subset of data that was not included in the training nor testing of the RFM, representing 39% of the full dataset. This subset included segmentation failures (63 cases), that were excluded from the RFM subset, contributing to its lowered recall. Additionally, 33 advected contrails were detected, totalling 298 detected contrails. The low number of advected contrails detected prevented further statistical analysis on them.

The pipeline demonstrated robustness across varying conditions, such as diurnal variation and against moderately complex background cloud coverage, involving intersecting contrails and ones that break apart irregularly. A dedicated analysis of false negatives was conducted to identify common failure reasons, including peripheral or partially obscured contrails, visually faint or low- contrast contrails, highly complex cloud backgrounds, and flight path discrepancies. We acknowledge these failure modes but note that our pipeline aims to create a high-fidelity dataset rather than high quantity, so most of these edge cases are not important, especially since we are focussed on high-climate impact contrails, that tend to not be faint and optically thin, and ones that can be later observed on satellite imagery.

The high spatial and temporal resolution of ground camera observations combined with detections being inherently flight-matched allow for several use cases. One of these cases is a comparison with meteorological outputs, SAC and ISSR flags, that provided meaningful insights showing strong recall, 92.8% of all contrails were formed in regions predicted to fulfil SAC in ECMWF ERA5 HRES data and 76.7% of persistent contrails were formed in predicted ISSRs, but poor precision as only 17.1% of flights within SAC-positive regions formed contrails and 12.8% of flights within ISSRs formed a persistent contrail (41.0% of all contrails in ISSRs were observed to persist). This result is in line with the known behaviour of ERA5, whose coarse resolution can smooth out finer variations in humidity fields (Reutter et al., 2020).

We find that the main limitation of the pipeline to be its reliance on an asemantic segmentation model. While SAM2 is unbiased and not tailored to any specific use case, it is not optimised for faint or complex contrail scenarios. This reflects the trade-off between generalisability and task-specific pretraining potentially introducing bias. In addition, wind speed errors and advection

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

inaccuracies further limited the reliability of advected contrail detection. These identified limitations are not unique to the presented pipeline but reflect broader challenges inherent to contrail detection. Ground-based cameras should not be considered in isolation but as a part of a wider system of multiple cameras and different observational sources.

This study's scope is limited by the use of a single camera; The natural next step is to scale the automated pipeline across the GMN to facilitate the widespread collection of contrail formation and early persistence data. Such a network would support a more robust evaluation of numerical weather prediction and contrail models over a wider and more global range. The use of multiple cameras, with overlapping field of views, would further enable contrail triangulation for the estimation of contrail altitude evolution inherently from observations, while also extending the observable contrail lifetime as they advect between fields of view. This would help bridge the gap between observing contrail formation and later-stage satellite observations.

Scaling the framework across the wider GMN may introduce a greater diversity of observing conditions, including camera viewing angle and orientation, foreground objects, local climate, and illumination, despite standardised hardware and calibration. Such conditions may require station-specific adaptations to the individual feature segmentation, extraction, and classification components. However, the modular nature of the proposed flight-first framework accommodates such adaptations without altering the underlying framework.

Within such a network, the flight-first approach is particularly fit for identifying regions for targeted follow-up observations from complementary sensing sources, such as satellites. This is due to its high fidelity, from its inherent inclusion of negative detections as well as positive, and its lower computational costs compared to a universally running detection-first method; Frames are only processed in association with flights passing through the (padded) FOV, and only relevant portions of each frame are processed, making the image classification problem a lighter task than the corresponding image segmentation task that is required in a full frame detection-first approach. To aid in targeted follow-up observations, the development of methodologies for translating flight-first contrail detections into spatiotemporal domains would have to be carried out. This would allow the identification of biases and a full contrail lifetime to be studied.

Future work can also focus on the enhancement of the pipeline's current implementation, including pre-processing improvements (e.g. adaptive contrast adjustment, kernel smoothing, and flightpath distortion to allow for a tighter input box) or custom video tracking that only propagates confident contrail-positive pixels. Further investigation into the relationship between visual metrics against physical meaning, such as persistence or climate impact, could extend the functional scope of the ground-based cameras, potentially enabling more retrospective analysis, estimating how long ago an advected contrail was formed. Together, these developments highlight the potential of a flight-first approach to serve as a foundational component of multi-source contrail observation.

Code and data availability

SAM2 is open source software, accessible at https://github.com/facebookresearch/sam2. Pycontrails, which was used for advecting flights and running CoCiP, can be accessed at [https://py.contrails.org/.](https://py.contrails.org/) The camera

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

footage used in analysis can be requested via the GMN. Code for mapping flightpaths onto image frames is from the GMN and is available at https://github.com/CroatianMeteorNetwork/RMS. Code related to processing of the flight-first analysis pipeline is available upon request, as is the hand-labelled dataset used for validation.

Author contributions

DPN carried out formal analysis and investigation, designed and developed the methodology, and prepared the original manuscript draft and visualisations. JPI assisted with methodological development, and draft preparation. JPI and MEJS conceptualised the study and provided supervision. DV and LB were responsible for resources, software, data curation and funding acquisition for the Global Meteor Network. All authors contributed to the interpretation of results and reviewed the final manuscript.

Competing interests

We declare no competing interests.

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher

Acknowledgements

We thank Peter Eschman for the ground-camera data.

Financial support

DV and LB were partially funded by grants and donations from Breakthrough Energy, Google Research and the Western Space Catalyst Grant.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

Review statement

The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

References

Alejandra Martin Frias, Shapiro, M., Engberg, Z., Zopp, R., Soler, M., and Emil, M.: Feasibility of contrail avoidance in a commercial flight planning system: an operational analysis, Environmental research: infrastructure and sustainability, https://doi.org/10.1088/2634-4505/ad310c, 2024.

Appleman, H.: Design of a Cloud-Phase Chart *, Bulletin of the American Meteorological Society, 35, 223–225, https://doi.org/10.1175/1520-0477-35.5.223, 1954.

Ashraful Alam Nirob, Ahmed, S., Tahmidul Karim Takee, Adnan Rahman Eshan, Haque, S. U., Ehsanur Rahman Rhythm, and Rasel, A. A.: Contrail Analysis through Advanced Neural Network Architectures: Image Segmentation and Classification, 1–6, https://doi.org/10.1109/iccit60459.2023.10441226, 2023.

Bradski, G. (2000). The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, *25*(11), 120-123.

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, 578, https://doi.org/10.3390/aerospace10070578, 2023.

Dalmau, R., Jarry, G., & Very, P. (2025). Contrail-to-Flight Attribution Using Ground Visible Cameras and Flight Surveillance Data. *arXiv preprint arXiv:2510.16891*.

Dharmendra Kumar Singh, Sanyal, S., and Wuebbles, D. J.: Understanding the role of contrails and contrail cirrus in climate change: a global perspective, Atmospheric chemistry and physics, 24, 9219–9262, https://doi.org/10.5194/acp-24-9219-2024, 2024.

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.

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, https://doi.org/10.1029/2025gl118386, 2025.

Geraedts, S., Brand, E., Dean, T. R., Eastham, S., Elkin, C., Engberg, Z., Hager, U., Langmore, I., McCloskey, K., Ng, J. Y.-H., 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, 015008–015008, https://doi.org/10.1088/2515-7620/ad11ab, 2023.

Gierens, K. and Vázquez-Navarro, M.: Statistical analysis of contrail lifetimes from a satellite perspective, Meteorologische Zeitschrift, https://doi.org/10.1127/metz/2018/0888, 2017.

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 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, 084059–084059, https://doi.org/10.1088/1748-9326/ad5b78, 2024.

Han, H., Guo, X., and Yu, H.: Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest, 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), https://doi.org/10.1109/icsess.2016.7883053, 2016.

Haywood, J. M., Allan, R. P., Bornemann, J., Forster, P. M., Francis, P. N., Milton, S., Rädel, G., Rap, A., Shine, K. P., and Thorpe, R.: A case study of the radiative forcing of persistent contrails evolving into contrail-induced cirrus, Journal of Geophysical Research, 114, https://doi.org/10.1029/2009jd012650, 2009.

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., and Chiara, G.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, https://doi.org/10.1002/qj.3803, 2020.

Hoffman, J. P., Rahmes, T. F., Wimmers, A. J., and Feltz, W. F.: The Application of a Convolutional Neural Network for the Detection of Contrails in Satellite Imagery, Remote Sensing, 15, 2854–2854, https://doi.org/10.3390/rs15112854, 2023.

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.

K. Gierens and P. Spichtinger: On the size distribution of ice-supersaturated regions in the upper troposphere and lowermost stratosphere, Annales Geophysicae, 18, 499–504, https://doi.org/10.1007/s005850050907, 2000.

Lee, D. S.: The Contribution of Global Aviation to Anthropogenic Climate Forcing for 2000 to 2018, Atmospheric Environment, 244, 117834, https://doi.org/10.1016/j.atmosenv.2020.117834, 2020.

Low, J., Teoh, R., Ponsonby, J., Gryspeerdt, E., Shapiro, M., and Marc: 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.

Marc Shapiro, Zeb Engberg, Roger Teoh, Marc Stettler, & Tom Dean. (2023). pycontrails: Python library for modeling aviation climate impacts (v0.47.2). Zenodo. https://doi.org/10.5281/zenodo.8387328

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, 034039, https://doi.org/10.1088/1748-9326/ac26f0, 2022.

OpenContrails: Benchmarking Contrail Detection on GOES-16 ABI:

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.

Paolo Pertino, Pavarino, L., Lomolino, S., Miotto, E., Cambrin, D. R., Garza, P., and Emanuele Ogliari: Ground-Based Contrail Detection by Means of Computer Vision Models: A Comparison Between Visible and Infrared Images, 254–259, https://doi.org/10.1109/rtsi61910.2024.10761667, 2024.

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, Kalyan Vasudev, Carion, N., Wu, C.-Y., Girshick, R., Dollár, P., and Feichtenhofer, C.: SAM 2: Segment Anything in Images and Videos, arXiv (Cornell University), https://doi.org/10.48550/arxiv.2408.00714, 2024.

Reutter, P., Neis, P., Rohs, S., and Sauvage, B.: Ice supersaturated regions: properties and validation of ERA-Interim reanalysis with IAGOS in situ water vapour measurements, Atmospheric Chemistry and Physics, 20, 787–804, https://doi.org/10.5194/acp-20-787-2020, 2020.

Rosenow, J., Jakub Hospodka, Sébastian Lán, and Fricke, H.: Validation of a Contrail Life-Cycle Model in Central Europe, Sustainability, 15, 8669–8669, https://doi.org/10.3390/su15118669, 2023.

Sankar, T., Dean, T., Abbott, T., Blickstein, J., Martín Frías, A., Galyen, M., Grenham, R., Hodgson, P., McCloskey, K., Pechman, A. and Robarge, T., 2026. Efficacy of Scalable Airline-led Contrail Avoidance. *Journal of Environmentally Compatible Air Transport System Discussions*, *2026*, pp.1-25.

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.

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, https://doi.org/10.1127/metz/2023/1157, 2023.

Schmidt, E.: Die Entstehung von Eisnebel aus den Auspuffgasen von Flugmotoren, Schriften der Deutschen Akademie der Luftfahrtforschung, Verlag R. Oldenbourg, München, Heft 44, 5, 1–15, 1941.

Schumann, U.: Über Bedingungen zur Bildung von Kondensstreifen aus Flugzeugabgasen, 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.

Smith, J. R., Grobler, C., Hodgson, P. J., Mukhopadhaya, J., Shapiro, M. L., Mirolo, M., Stettler, M. E. J., Eastham, S. D., and Barrett, S. R. H.: The climate opportunities and risks of contrail avoidance, Nature Communications, https://doi.org/10.1038/s41467-026-68784-8, 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., and Norris, B.: Feasibility

https://doi.org/10.5194/egusphere-2026-5251
Preprint. Discussion started: 21 September 2026

test of per-flight contrail avoidance in commercial aviation, Communications Engineering, 3, https://doi.org/10.1038/s44172-024-00329-7, 2024.

Spichtinger, P., Gierens, K., Leiterer, U., and Dier, H.: Ice supersaturation in the tropopause region over Lindenberg, Germany, Meteorologische Zeitschrift, 12, 143–156, https://doi.org/10.1127/0941-2948/2003/0012-0143, 2003.

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., 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, 2024a.

Teoh, R., Engberg, Z., Shapiro, M., Dray, L., and Stettler, M. E. J.: The high-resolution Global Aviation emissions Inventory based on ADS-B (GAIA) for 2019–2021, Atmospheric Chemistry and Physics, 24, 725–744, https://doi.org/10.5194/acp-24-725-2024, 2024b.

Van Huffel, J., Ehrmanntraut, R., and Croes, D.: Contrail Detection and Classification using Computer Vision with Ground-Based Cameras, 2025 Integrated Communications, Navigation and Surveillance Conference (ICNS), 1–6, https://doi.org/10.1109/icns65417.2025.10976944, 2025.

Vida, D., Šegon, D., Gural, P. S., Brown, P. G., McIntyre, M. J. M., Dijkema, T. J., Pavletić, L., Kukić, P., Mazur, M. J., Eschman, P., Roggemans, P., Merlak, A., and Zubović, D.: The Global Meteor Network – Methodology and first results, Monthly Notices of the Royal Astronomical Society, 506, 5046–5074, https://doi.org/10.1093/mnras/stab2008, 2021.

Virtanen, P., Gommers, R., Burovski, E., Oliphant, T. E., Weckesser, W., Cournapeau, D., ... & Feng, Y. (2021). scipy/scipy: SciPy 1.6. 0. *Zenodo*.

Youden, W. J.: Index for rating diagnostic tests, Cancer, 3, 32–35, https://doi.org/10.1002/1097-0142(1950)3:1%3C32::aid-cncr2820030106%3E3.0.co;2-3, 1950.