Authors' Response to Reviews of

Factors limiting contrail detection in satellite imagery

Oliver G.A. Driver, Edward Gryspeerdt, Marc Stettler
Atmospheric Measurement Techniques, egusphere-2024-2198

RC: ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

We would like to thank the reviewers for their comments on our manuscript.

We would draw attention to several changes to the analysis, which are also referenced below in response to individual review comments that triggered them:

C.1 Changes to the detectable proportion of radiative forcing (Figs. 7, 12): The original manuscript weighted contrail segments on their estimated forcing per-unit-area to yield the "detectable fraction of forcing", which is not as clear an interpretation as the revised version. For linear contrails, it is clearer to weight on the forcing per length of contrail (which is also approximately equivalent to weighting on the total forcing of the segment, so estimates better reflect the detectable fraction of contrail climate impact). This change also makes these results consistent with other analyses in the paper (i.e. Figs. 9, 10, and 11).

C.2 The removal of the diurnal component of variability (Figs. 7, 12; Section 4.1.1). The original manuscript erroneously stated that this component of variability was found by subsetting on local time—the time thresholds were instead applied on UTC time. Thresholding on local time results in little added variability, except in the case of SW RF (which is entirely due to the lack of SW RF during the night). We believe the fact that seasonal variability in contrail properties alone has a significant impact on the RF is a more meaningful result, and that it is not useful to include the local solar time variability in the error bars (because this would weaken intercomparison between the differently-weighted results).

C.3 An additional cirrus background experiment where technically feasible (Fig. 8(f), Section 4.1.3). This acts as a point of comparison to the original test with background cirrus, which has properties not representative of naturally-forming midlatitude cirrus.

C.4 Additional 7 km resolution tests (Fig. 7), to act as a point of comparison for the impact of using significantly more coarsely-resolving imagers. This additional analysis requred testing of much longer synthetic contrail segments, because the detection algorithm requires a minimum number of pixels to be identified, (length has been changed from 50 to 150 km).

The manuscript has also been significantly restructured, and improvements made to the clarity of exposition.

Please find detailed responses to individual comments below, followed by the figures which have been significantly altered.

Line numbers refer to locations in the 'diff' file.

1. RC1: 'Comment on egusphere-2024-2198', Anonymous Referee #1, 19 Sep 2024

RC: ***Summary: The authors estimate the maximum observability of persistent linear contrails in current geostationary satellite imagery by applying a contrail detection algorithm similar to Mannstein et al. (1999) to synthetic thermal infrared images of contrails within an otherwise clear sky over a homogeneous background. By varying several parameters related to contrail observability singly, they determine how contrail optical properties and sensor resolution can affect the contrail observability. Although only 46 percent of the contrails produced in CoCiP model simulations are expected to be detectable in current 2-km imagery, the detectable contrails would represent nearly 80 percent of the net contrail radiative forcing. The authors also find that the lifetime-integrated LW radiative forcing increases nearly linearly with the observable lifetime of the simulated contrails, implying that to the first order, contrail longevity in satellite imagery is a function of the optical depth of the contrails.***

RC: ***General comments: The paper covers an interesting and important aspect of the satellite remote sensing of contrails. Accurate estimates of the detectability of contrails from geostationary satellites are needed, and this study provides a straightforward way to establish the upper bounds of contrail observability. Although the studys methodology and conclusions appear to be sound, the paper is weakened by its poor organization. The reader is presented with many details throughout the paper that are often presented in a confusing manner. The manuscript would be strengthened by rearranging the exposition in a more comprehensible manner. The authors present some excellent results so I recommend publication after a complete reorganization of the text.***

AR: We thank the reviewer for their generous reading of our manuscript and detailed comments. The reviewer's contribution to make this a stronger overall manuscript is appreciated. In response, we have made an effort to reorganise in a more coherent way, and believe that this improves the readability.

In particular, we have broken results sections into further subsections when appropriate, adjusted the structure of figures (including moving details for interpreting figures away from the main text, and into the captions), moved methodological details which are only relevant for certain analyses to the relevant results sections, and removed references to analyses later in the text. We have also improved the clarity of the text where possible. Changes **C.1** and **C.2** have also enabled simplifications.

RC: ***Specific comments:***
Line 106 synthetic is misspelled.

AR: Thank you for spotting, corrected in the revised manuscript.

RC: ***Line 305: I suggest for better writing style, do not begin a sentence with a number.***

AR: The text has been rephrased to avoid this.

RC: ***Line 316: diurnal is misspelled.***

AR: Corrected where still occurs the revised manuscript.

RC: ***Introduction: All of the discussion in the manuscript pertains to linear contrails only, but this is not acknowledged in the paper.***

AR: We have added the specifier that the contrail detection algorithm acts on and relies upon the linearity of contrails.

Lines 31–33
The observations use infrared images—particularly split-window brightness temperature images (Lee, 1989)~~which —~~which highlight optically thin ice clouds with small crystals against the surface and background liquid clouds. Both the strong signal and the linearity of the object are used to detect the presence of a line-shaped contrail.

We have also specified the same for neural net algorithms.

51–52
Each of these studies uses convolutional neural nets to detect the contrails. These rely on datasets of extended line-shaped contrails used as 'training data' to produce an algorithm that is able to detect linear contrails.

In addition, a discussion of treating segments of contrails has been added (enabling variability along the length of the contrail, and acknowledging that contrails are not straight lines and may advect and evolve differently along their length).

136–140
Alongside B and IWP_0, the effective radius (r_{eff}) is another characteristic contrail property (describing the size of constituent ice crystals). Each synthetic contrail is assumed to have constant r_{eff}. In reality, contrails are not straight lines, and their properties and evolution may vary along their length. The consequence of such variations is considered in this work by using the detectability of these test contrails to inform the observability of modelled contrail 'segments', the overall contrail's properties may vary from segment-to-segment.

The limitation to the linear phase of a contrail's evlolution is also now mentioned in the discussion of 'applicability of methodology'.

642–643
This tracking method also stands to enable the consideration of detected contrails beyond their initial linear phase.

Finally, we specify 'linear contrails' instead of just 'contrails' in several places.

RC: ***It is not clear why the first 10 - 30 min of the contrail lifetime when contrails are not observed would affect the overall contrail cirrus radiative forcing significantly. If a contrail is detected, the extra forcing from the pre-observable contrail could be approximated and added to the overall lifetime radiative forcing estimate. The contrails are thin and narrow at this time so the overall radiative impact is small, as suggested by the results in Fig 9(c) and 9(d).***

AR: A note has been added to mention the suggested technique, which could be applied to contrail forcing estimates. This process would require the approximation of the elapsed time to the first detection, and assumptions about its properties up to this point (which would be difficult to retrieve if the contrail is only producing a minimally-detectable signal). We agree that the pre-detectable period would have minimal radiative impact; the delayed onset of observation is mostly an obstacle when attributing contrails to flights—a discussion has been added here to this effect.
We'd also note the interval before observation has been amended in response to a comment from another review.

48–51
Initial detection in geostationary images has been found to occur ~~10–30~~ 10–45 minutes after formation ~~(Chevallier et al., 2023; Gryspeerdt et al., 2024)~~(Chevallier et al., 2023; Gryspeerdt et al., 2024; Geraedts et al., 2024) , indicating that contrails are unobservable for at least the earlier part of their evolution. Although the contrails are assumed to have small radiative impact during this time, which could be estimated if accurately matched to a generating flight, the delayed onset is an obstacle when attempting to attribute observed contrails to specific aircraft (as was an aim of each of these studies).

RC: ***Section 2: This section contains many details about the contrails, but nothing about the surface conditions. Do the surface conditions change? What surface temperature is used in the calculations? Also, there is no information here about the viewing zenith angles and solar zenith angles used in the calculations.***

AR: A more detailed description of the treatement of the background has been added (including explicitly the assumption of a homogeneous, unit emissivity surface).
The US standard atmosphere is used in the calculations, including the surface temperature of $15\,^{\circ}\mathrm{C}$. An explicit statement of this temperature has been added to the text.
Variations in the background have not been considered here, with the choice of a unit emissivity, homogeneous surface serving the aim of a maximally-observable case (to give an upper-limit on detectable fraction).
Viewing zenith angle been approximated as zero—adding an assumption of observation of a contrail at nadir. This has also been added to the text. Radiative forcing calculations take into account the ambient meteorology and the solar zenith angle (estimates are from the CoCiP runs using pycontrails (Shapiro et al., 2023), based on Schumann et al. (2012)).

150–152
The viewing zenith angle has been taken to be zero, introducing an assumption that observation occurs near the satellite's nadir. The surface is assumed to have unit thermal emissivity (similar conditions to the ocean), and the surface temperature is treated as 15 °C (a property of the temperature profile). GOES-ABI imager bands are simulated using the REPTRAN representative wavelength parameterisation (Gasteiger et al., 2014).

RC: ***The discussion of the background cirrus layer is out of place and would be better located in section 4.1. Why is the effective radius of the background cirrus layer only 5 microns? Most cirrus layers have much larger mean particle sizes. For example, Wang et al (2019) and Yi et al. (2017) report a global mean of ice cloud effective radius around 30 microns. An effective radius of 5 microns is odd for a midlatitude cirrus layer and a more realistic value may affect the observability results in section 4.1.***

AR: We have moved this discussion to Section 4.1.3 as suggested.
The choice of particle size is a result of the technical limitation on the radiative transfer model used—only one ice cloud parameterisation is possible at a time, and 5 microns is the only effective radius shared between the two parameterisations used.
For a point of comparison, we have calculated the observability threshold against a layer of cirrus with 30 µm effective radius and similar optical thickness for $r_{\mathrm{eff}} > 5$ µm (change **C.3**). The below discussion has also been added.

489–495
The chosen effective radius (due to technical limitations) would be relatively unphysical for a midlatitude natural cirrus layer (although it may represent, for example, observation against a contrail cirrus

outbreak). To provide a comparison, a similar test against a 30 μm cirrus layer has also been performed for contrails whose r_{eff} is high enough that radiative transfer simulations can be made. The IWP has also been increased to approximately maintain an optical thickness of 3. The derived threshold is shown in Fig. 8(f). Against this background, contrails with a higher effective radius become observable, demonstrating that this cirrus layer would pose less of an obstacle to contrail detection. From the data available for contrails below 10 μm, the two background cirrus thresholds appear comparable at lower effective radii.

RC: ***Figure 5: This figure is difficult for the reader to follow. Many details are missing or unclear, both in the figure caption and in the accompanying text. The observability test results for 2 km wide contrails are presented as three separate subplots, but the results are condensed into one subplot for the other two contrail widths. The caption also describes parts (d) and (e) as histograms, rather than plots of contrail observability. Parts (b) and (c) are not clearly explained. What do the values in part (b) represent? The fraction of CoCIP contrails of a particular effective radius and IWP that would be detected? What do the values in part (c) represent? The fraction of contrail segments with a positive or negative contrail radiative forcing (which is positive (LW forcing?) and which is negative (SW forcing?)?***

AR: This figure has been restructured for clarity—the revised version is included at the end of this document. The histograms (now panels (d)–(i)) are intended to show the modelled distribution of contrail properties in the parameter space (width, r_{eff}, IWP), alongside the derived observability thresholds.
To be explicit, the values in panels (d)–(f) represent number of segments in the bin, relative to the maximum number of segments. The values in (g–h) are again the contrail population but this time weighted by the mean RF of contrails in that bin. The units are relative to the largest (absolute) value for any one bin at that width.

`Fig.5 caption`
Illustrative slices from a ~~3-D~~ contrail observability test ~~, including~~for three contrail-width bins (chosen from 30 total bins): ~~observability quantified using effective width~~ 0.2 km- (a, d, g), ~~and histograms of CoCiP populations weighted by occurrence~~ 2 km- (b~~)~~, e, h) and ~~net radiative forcing~~ 10 km-wide (c~~)~~, ~~for 2 km wide contrail segments~~f, i) contrails. ~~Similar~~ The observability threshold is plotted over the derived detection probability (a–c), and histograms ~~are shown for 0.2 km wide contrails~~ of CoCiP population 1 (~~d~~d–f) and ~~15 km wide~~ of the same population weighted by the mean net RF of the contrails in each bin (~~e~~g–i). The histogram values (d–i) are relative to the magnitude of the extreme value in each plot. The 'adjusted threshold' represents the observability threshold adapted to include all theoretically observable contrails, without the high-width algorithm deficiency. Contours of the contrail optical thickness (estimated based on IWP_0 and r_{eff}) are also shown.

RC: ***Why do 2-km and 10-km wide contrails produce only positive forcing, but the 0.2-km wide contrails produce both positive and negative forcing? The results appear to indicate that only contrails with net positive forcing would be detectable by GOES imagery, except a few (?) optically thick and very narrow contrails. This surprising result would be an important outcome of this research, and should be discussed more.***

AR: Because the values of panels (g)–(i) are weighted by the mean forcing for a bin, a positive (or negative) value does not indicate that all contrails with these properties are warming (or cooling)—this will depend on the conditions. We hope that our changes to Fig. 5 described previously help to make this clear.

RC: ***The discussion in lines 271 through 277 about the adjusted observability threshold and the cause of unobservability is confusing and doesnt add much information. Perhaps it can be removed?***

AR: The presentation of results here has been thoroughly restructured to avoid confusing the reader. The descriptions of the adjusted threshold and of the 'causes of unobservability' have been simplified significantly (with the specific wording removed), and integrated with the description of results. We hope this makes the exposition of these results more fluid, and these other discussions feel more natural.

The discussion of results in Section 3.2 has been heavily revised to improve the clarity of this section, including moving details needed for interpreting the figure to captions, and moving more relevant details to this section, such as how the p_{obs} threshold is constructed.
A full diff of Section 3.2 is not included here due to the widespread revisions, but we ask that you refer to the revised manuscript.

RC: ***Section 3.3***
Line 294: population 2 (Fig 8). The authors have several references like this one, in which figures and results not yet presented in the text are mentioned, confusing the reader and weakening the exposition of the research. Please remove these references or place them at more appropriate locations in the text.

AR: Thank you for the suggestion—we have worked to improve the flow of the revised manuscript, removing these references.

RC: ***Lines 309 through 310: What is the ratio between contrails with strong (what constitutes strong forcing?) LW forcing compared to those with strong SW forcing? This section once again hints that only unusually thick yet narrow contrails have an overall negative (SW greater than LW) forcing. Is this because only small solar zenith angles are used in the radiative transfer calculations? I would expect SW forcing to dominate when the sun is low.***

AR: We believe this wording was imprecise and would remove.
On reflection, making the comparison between contrails with strong forcing when using per-unit-area quantities made the interpretation of these results unclear, contributing to the decision to make change **C.1**. Following this change, this result is no longer seen—instead the observable fraction of forcing doesn't vary much with resolution and with component of forcing.

> **425–432**
> ~~Contrails that have a strong LW forcing are typically more observable than those with a strong SW forcing. This is because the cooling (SW-forcing) population tends to consist of narrower (younger) contrails. This is visible by comparing the SW-dominant and LW-dominant regions of the net forcing CoCiP population in Fig. 5. This may be caused in part by the combination of~~ This interpretation is also consistent with the findings that the observable fraction of contrail forcing does not vary significantly much with imager resolution or by component of forcing (because the few optically thick contrail segments causing much of the ~~daytime-only nature of SW forcing and the daytime bias of air traffic (Schumann and Heymsfield, 2017). As a result, the discrepancy between LW and SW components is less in the higher-resolution imager~~forcing are sufficiently wide that the resolution dependence is less significant).

As discussed in the response to a previous comment on zenith angle, the radiative forcing estimates on each contrail account for the solar zenith angle.

RC: ***Section 4.1***
Lines 314 through 315: The descriptions of Pobs-derived threshold uncertainty and seasonal and diurnal variability should be placed here, not in sections 2.2 and 2.3.

AR: As suggested, the discussions of uncertainty have been moved here.
We recognised several structural issues with the exposition in Section 4.1, informed by your review, so have restructured the results of this section, including separating into an initial presentation of the central results (Section 4.1.1), followed by the speculatively adjusted versions to illustrate limitations of algorithm (Section 4.1.2) and background conditions (Section (4.1.3)).

RC: ***The populations 1 and 2 are labeled inconsistently throughout the paper (population 2 versus CoCiP population, instantaneous population versus population 1). There is also a baseline population, which appears to be different from populations 1 and 2? Please use only one set of descriptors for the contrail populations, otherwise it is very confusing.***

AR: References to the populations have been reframed in terms of populations 1 and 2.

RC: ***Lines 323 through 329: The Mannstein et al (1999) algorithm was developed for one AVHRR satellite, which has different resolution, sensor sensitivity, and noise characteristics than the GOES imagers described here, and the threshold value identified by Mannstein et al (1999) cannot be applied to other AVHRR imagers, let alone the GOES imagers. The detection algorithm can be adjusted to detect more contrails, but given the differences between the AVHRR and GOES imagery performance, the numbers presented here are speculative, especially without any information about false positive detections.***

AR: The discussion of the different designs of the two sets of thresholds has been expanded to better motivate this as not just trying a different algorithm designed for the wrong instrument, but to make clear that the McCloskey et al. (2021) implementation may lock in the limitations of human observation due to the necessity of an assumed 'ground truth' when tuning, for which only human-labelled data is available.

447–454
The algorithm used contains thresholds ~~as distributed with McCloskey et al. (2021),~~ tuned on human-labelled datasets to ~~maximise precision and recall—making these results reflect the contrails which would be observable as distinct objects by~~ optimise precision and recall, as distributed with McCloskey et al. (2021). Real contrails which were not identified by the human labellers would be spuriously considered 'false positives' if picked up by the detection algorithm. This means that the algorithm reflects the contrails theoretically observable by the human labellers. ~~A~~ Mannstein et al. (1999) chose a brightness temperature difference threshold uniquely low compared to existing algorithms at the time (0.2 K), and much lower than the threshold in the tuned algorithm (1.33 K), aiming to use the contrail's geometry rather than thresholding to ensure reliability.

The adjusted algorithm has been tested on synthetic satellite observations that do not contain a contrail, in order to check that detection remains restricted only to actual contrail occurrences (and we acknowledge that without this the analysis would be meaningless—an algorithm that detected 'contrails everywhere' would have 100 % detection efficiency). There are no false positives in the idealised images caused by the adjusted threshold.

455–470
As a test, a less-conservative detection algorithm was ~~tested, based on reducing the threshold~~ constructed by reducing applied thresholds from the human-labelled tuned ~~value~~ values to the least-conservative between ~~this and the value identified~~ these and the values used by Mannstein et al. (1999). ~~The~~ It should be noted that such an algorithm poses an increased risk of false positive detection so may only be practically applied when this risk can be reduced (such as the targeted observation of a contrail known

to exist, or suspected to exist in a timestep following one where it is more confidently detectable). The algorithm of Mannstein et al. (1999) was designed for use with a different imager, so such algorithm adjustments need testing for their specific application. When applied to 10000 realisations of the clear sky instrument noise field (i.e. synthetic image with no contrail) no false positive detections occurred using this new algorithm, indicating that it is suitable to speculate on potential achievable detectabilities, albeit in this very controlled case. The threshold derived using this ~~adjustment~~ less-conservative algorithm is shown with with an evolving contrail in Fig. 8(~~b~~d). In this case—using the same high-width adjustment of Section 3.2—(89 ± 1) % of segments and more than ~~95~~99 % of all components of forcing is theoretically observable in fields simulated with 2 km resolution. ~~This~~

Observation-independent ground truth data does not exist, so these detectabilities are unachievable for general observation (because false positives cannot be controlled for). However, this initial analysis suggests that the observability limit derived here can be relaxed if the risk of false positives can be ~~avoided~~mitigated, for example in the case of targeted observation of specific contrails based on advected flight tracks, where a likely location of the contrail is known.

RC: ***Lines 339 through 340: Wider contrails with similar optical thickness are detected, because less-optically-thick parts of the contrail exist which remain detectable. I have no idea what the authors are trying to say here.***

AR: This description of the behaviour of the observability threshold discontinuity has been rephrased to clarify.

481–488

~~A high-IWP limit of detectability seen in this observability threshold (as in of Section 3.1, Fig. 3(b))~~The minimum width required for detection shows a discontinuity at ~~high-IWP.~~ high-IWP_0, similar to the limit in Fig. 3(b). In this case the combination of background cirrus and contrail causes does not cause a brightness temperature difference signal at the centre of the contrail as it is sufficiently optically thick to be opaque in both the channels that are differenced. ~~Wider contrails with similar optical thickness are detected, because less-optically-thick parts of the contrail exist which remain detectable.~~ The high-IWP_0 discontinuity in the threshold does not completely prevent detection, but significantly increases the minimum detectable width. This is because the IWP profile (eq. 1) varies gradually-enough in wide contrails that their edges remain detectable.

RC: ***Section 4.2***
Fig 9 is very complicated but the description in the text is not always helpful. As a result, it is not clear, for example, what a post-observable contrail segment is, what LW RF/length and what LW RF/length/segment mean. The comment on lines 372-373 that narrow and optically thin contrails must be avoided in contrail mitigation trails is so vague as to be meaningless. Should the production of narrow and optically thin contrails be avoided, or is it the counting of such contrails in an assessment of the mitigation trial that should be avoided? Revise the figure caption and accompanying text so that the results of the figure are more understandable.

AR: We have revised Section 4.2 to focus the text on the exposition of results, with details to enable the interpretation of figures moved to the caption.

Fig 9 caption

The time evolution of observability—the observability status of contrail segments (a, b) and of their proportions of forcing (c, d) formed globally within the time-evolving sample, using Jet-A1 fuel, for both 2 km (a, c) and 0.5 km (b, d) imagers. Using the adjusted observability threshold of Section 3.2 (Fig. 5; Fig. 8(a)), segments are categorised as 'observable', 'never observable', 'pre-observation' (have not yet been observable but will be observable later), or 'post-observation' (were observable earlier in their evolution, but are not at this age). Unobservable contrails have also been categorised as either too-narrow or too-optically-thin to be observed. Lines indicate: the number of persisting contrail segments (a,b), LW forcing (c,d), and the per-segment-mean LW forcing (c, d) due to contrails of a given age (relative to other CoCiP population 2 age-bins).

We have also amended the labelling 'RF/length' and 'RF/length/segment', which were confusing and inaccurate—they now read 'RF LW' and 'RF LW per segment'. The revised figure is shown at the end of this document.

These considerations also contibuted to the simplifications made in change **C.1**. A further description has been added to the text to clarify these 'overplot' quantities.

526–533

Each of the panels ~~are overplotted with relevant integrated quantities. The surviving proportion of segments is shown in Fig. 9(a,b) and the instantaneous LW~~ of Fig. 9 is overplotted with the relative contribution of contrails with a given age to the instantaneous population of contrails. Specifically, the proportions of contrail segments are shown alongside the fraction of persisting segments, equivalent to the fraction of segments that would be expected to be a given age for a given time. The proportions of forcing are shown with the contribution of contrails with a given age to the instantaneous global forcing due to contrails (as well as the per-segment mean forcing of contrails with a given age, approximately equivalent to the relative forcing per unit length ~~(both integrated over the surviving contrails and per contrail) is shown on Fig. 9.~~ of contrails with a given age).

These changes have enabled simplification of the text elsewhere.

The specific comment raised in original manuscript (that never-observable contrails should be "avoided" in trials) has been removed in favour of a note on the reason why this was stated—that they cannot be observationally verified. The point attempting to be made about considering observability for experimental trials is better left to the conclusions.

539–544

The population of never-observable contrails ~~(blue)~~ consists of both ~~narrow and optically thin~~ too-narrow and too-optically-thin contrails. This population is much reduced in the forcing-weighted analyses (Fig. 9(c,d)), suggesting that keeping track of these contrails is less important for monitoring the radiative forcing. ~~For trials attempting to mitigate the formation of contrails~~However, it is ~~important that these contrails are avoided as they would lead to false negatives—inference ofno contrail when one has actually formed~~apparent that the formation (or lack of) of a significant proportion of contrail segments would not be validatable using satellite-borne imagers.

RC: ***Line 375 - The pre-observable population in the forcing-weighted case (Fig 9(c) and 9(d)) appear to be only a few percent of the fraction of contrails for the 2-km resolution case, and much less for the 0.5-km resolution case, How can the authors claim they play an important radiative role (especially for the 0.5-km case)?***

AR: On reflecting, it is clear in Fig. 9(c,d) that this claim is incorrect. We have removed this comment.

> **545–547**
> Pre-observable contrails tend to be too narrow to be observed, becoming observable when they have broadened sufficiently. A pre-observable population persists into the forcing-weighted case, ~~so these contrails also play an important radiative role. The~~ but the high-resolution imager provides a significant improvement in accessing this population.

RC: ***Lines 382-383: As a check for consistency with the analysis of Fig. 7, the proportion of observable contrails has been integrated with the corresponding total. The corresponding total of what? Im assuming we are looking at all of the population 2 contrails here (once again called something different [time-evolving contrails] in the Fig 10 caption)?***

AR: This description has been rephrased to clarify the methodology—that we are just calculating detection efficiencies for population 2.

> **553–556**
> As a check for consistency with the analysis of Fig. 7, the ~~proportion of observable contrails has been integrated with the corresponding total~~ detection efficiencies of CoCiP population 2 contrails and of their RF have been estimated. This is performed by integrating the product of the observable proportion of contrails (or of LW RF) with the persisting fraction of segments (or relative contribution to the total LW RF) over the time since contrail formation.

The caption to Fig. 10 has been revised to keep the name of the population consistent.

RC: ***Line 397: To confirm, does the (21 +73 -11) min symbol mean the average first detection was at 21 min, while the minimum time for the first detection was 21 - 11 = 10 min, and the maximum time for the first detection was 21 + 73 = 94 min?***

AR: This is a correct interpretation (with 'minimum' and 'maximum' being first and third quartiles). We recognise these symbols were confusing, so they have been removed in favour of explicitly stating this for clarity.

> **571–579**
> Finally, the distribution of the times ~~at which a contrail~~between a contrail's formation and the time at which it first becomes observable are examined. This transition is highly skewed, so the median age of first observability is given, with the first and third quartiles ~~are given as an uncertainty range~~to give an impression of the variability. In 2 km resolution images, ~~detection first occurs at (21^{+73}_{-11}) min , whereas observation occurs at an age of (9^{+61}_{-4}) min~~ the median time for the onset of detectability occurs at 21 min with first and third quartiles at 10 min and 94 min. The onset occurs at a median 9 min after formation in hypothetical 0.5 km images (between 5 and 70 min quartiles). This is in good agreement with the delays to onset of ~~GOES~~ GOES-ABI (ca. 2 km resolution) observation reported by Chevallier et al. (2023) and Gryspeerdt et al. (2024)~~.~~, although these studies do not capture the higher limits estimated here. This is likely due to the increased difficulty in matching observed contrails to generating flights when the delay to observation increases.

RC: ***Fig 11: This plot seems to imply that the longest-lived contrails, that is the ones detectable in the GOES imagery the longest, are simply the most optically thick? This is another important result that needs to be highlighted in the conclusions.***

AR: The figure was labelled with incorrect units in the original manuscript and was not clearly described by the caption—the forcings are lifetime-integrated so should carry the units $\mathrm{J\,m^{-1}}$ (of an energy forcing multiplied by the width).The figure and caption have been revised.
As a result, this analysis doesn't imply that longest lived contrails are instantaneously the most strongly forcing (i.e. the most optically thick). The analysis is instead only intended to establish that the longest-observed contrails also have the most impact over the total period they persist for. (so they are the most radiatively important). While this follows naturally if 'observable lifetime' is assumed to be correlated to 'persistence lifetime' this requires co-variation between the lifetime and the contrail properties (therefore the forcing) to be neglected. Therefore, the result remains important as it enables 'observed lifetime' to be interpreted as a proxy for 'radiative importance' (i.e. lifetime energy forcing, relative to shorter-observed contrails).
We hope that our revisions to the caption (below) and figure labelling improve interpretability of this result.

> **Fig. 11 caption**
> The relationship between lifetime-integrated energy forcing per unit length, and the observable lifetime. The ~~bold~~ strongly-shaded region denotes uncertainty in the mean, and the lightly-shaded region is the variability (the standard deviation for contrails with a given observable lifetime).

RC: ***Section 5.2 The authors discussion here seems contradictory at times. The authors state that detection probability is not just a function of contrail optical depth, but earlier in the paper they characterized the cause of unobservability in terms of either too narrow or too optically thin. An argument can be made from the results in Fig 5(b) that the effective radii of most contrail segments fall between 2 and 10 microns, and that from Fig 5(a) between 2 and 10 microns the detection probability threshold follows the line of constant optical thickness (I estimate the thickness to be about 0.05-0.06). I dont disagree that particle size does matter, but to the first order, optical thickness appears to be the most important factor determining contrail observability for most of the CoCiP contrails used here.***

AR: This interpretation is not true in general. The panels of Fig. 5 highlighted (now Fig. 5(e) and (b)) only apply to the contrails approximately 2 km wide. The same is not true for the other two contrail widths illustrated in Fig. 5, nor for the other contrail widths for which illustrative slices of the observability test have not been included.
We agree that the optical thickness is important, but would also highlight contrail width as an important cause (hence some contrails are 'too narrow'). The exact optical thickness threshold does exhibit variability depending on r_{eff} and IWP, including at values predicted to occur by CoCiP (particularly as contrail width varies), so consideration of these factors remains important.
We have revised Section 3.2 to make this clearer.

> **348–355**
> Contrails with a larger optical thickness tend to be more observable than those with a lower optical thickness~~. The observability threshold, $P_{\mathrm{obs}} > 0.5$, is shown as a black contour. Fig. 5(b) shows the observability threshold with the distribution of contrail segments with approximately this width, taken from the instantaneous contrail sample. Fig. 5(c) is as Fig. 5(b), with the population weighted by the net radiative forcing. For these 2 km contrails, the observability threshold splits the population (Fig. 5(b)), but the segments generating strong radiative forcings are generally observable~~ (, aligning well with expectations. Regardless of contrail width, contrails with optical thickness below approximately 0.05 were found to be undetectable, consistent with past work (e.g. Kärcher et al., 2009). Additionally, the

detectability of more optically thick contrails is seen to depend on the specific properties—particularly the contrail width, as well as r_{eff} and IWP_0.

2. RC2: 'Comment on egusphere-2024-2198', Anonymous Referee #2, 11 Oct 2024

RC: ***General comments***
In their manuscript Factors limiting contrail detection in satellite imagery, the authors investigate the principal limits of contrail observability using current generation geostationary satellite imagers under ideal conditions. A global dataset of CoCiP-modelled contrails was used to derive typical contrail properties at different stages of their lifecycle. Based on these properties, radiative transfer calculations were carried out to create synthetic satellite images corresponding to different contrails and different satellite imager resolutions. Finally, a simple contrail detection algorithm focusing mainly on line-detection was applied to those images to test the contrail observability. Using this setup, various sensitivity studies with respect to the contrail properties were conducted, indicating the main parameters, and the contrail observability at different stages of the contrail lifecycle was derived.

Both the study and the manuscript are well structured and written. Motivation, study approach results, conclusions and limitations were mostly described clearly. The study gives important insights in the limitations of contrail observability which are useful, e.g., for the satellite-based study of contrails, the validation of contrail models as well as the evaluation of live trials focusing on operational contrail avoidance.

Thus, the results are very important and the manuscript is recommended for publication after minor revisions.

AR: We thank you for reading our manuscript and contributing thoughtful comments. We have responded to each of your specific comments below, and revised the manuscript accordingly.
We believe the reviewer's contribution has made a better overall manuscript, particularly in the exposition of relevant background, and the inclusion of a more-coarsely resolved imager as a point of comparison.

RC: ***Specific comments***
L 1: Refine definition, e.g., Contrails (ice clouds, originally line-shaped, initiated by aircraft exhaust)

AR: We have adopted a slightly more specific definition based on the reviewer's suggestion.

> **Line 1–2**
> Contrails (~~clouds produced~~ ice clouds, originally line-shaped after initiation by aircraft exhaust) have a significant warming contribution to the overall climate impact of aviation.

RC: ***L 4: Stress the high temporal resolution of geostationary imagers, time-resolved can also mean once a day.***

AR: This has been adjusted to add emphasis.

> **4–6**
> Infrared imagers on geostationary satellites provide widespread ~~, time-resolved observations of~~ contrail observations, with sufficient time-resolution to observe the evolution of ~~contrail~~ their properties.

RC: ***L 10: Maybe replace instruments by satellite imagers***

AR: Revised to "satellite-borne imagers".

RC: ***L 12-14: Which resolution are you considering? I think if you continue to increase the resolution, you should eventually be able to observe all contrails. Could you specify your assumptions here?***

AR: Revised the text to emphasise that "the same imager" includes the same 2 km resolution as the previous result. High spatial resolution is indeed necessary, but is an insufficient condition to observe all contrails—sensitivity (i.e. ability to observe an optically thin contrail) is a second limiting factor. This distinction between too-narrow and too-optically-thin contrails (judged too much detail for the abstract) is made in the manuscript, showing that resolution improvements alone do not enable detection of the unobserved contrails.

RC: ***L 17: Maybe replace observation requirements by observation conditions***

AR: Agree that this is unclear. Revised to highlight that different applications require observation of different contrails.

> **19–20**
> However, there is a highlighted need to assess the observability of ~~specific contrails depending on the observation requirements of a given~~ contrails where the observation conditions may vary by application.

RC: ***L 39: Explain why human labellers are considered here (i.e., mention human-labeled training data)***

AR: A description of the need for training data has been included (changes included with the below comment).

RC: ***L 41: Compare also Geraedts et al. (10.1088/2515-7620/ad11ab, 2024), stating Most flight segments start matching contrails about half an hour after formation, with the mean time until first observation being 41 minutes***

AR: Revised to include the suggested citation. We have expanded time window for initial detection found previously in line with this additional result. Note we also comment on estimating the impact before initial detection, in response to another review (part of these changes also quoted above).

> **46–52**
> Initial detection in geostationary images has been found to occur ~~10–30~~ 10–45 minutes after formation ~~(Chevallier et al., 2023; Gryspeerdt et al., 2024)~~(Chevallier et al., 2023; Gryspeerdt et al., 2024; Geraedts et al., 2024) , indicating that contrails are unobservable for at least the earlier part of their evolution. Although assumed to have small radiative impact during this time, which could be estimated if accurately matched to a generating flight, the delayed onset is an an obstacle when making these attributions (as was an aim of each of these studies). Each of these studies uses convolutional neural nets to detect the contrails. These rely on datasets of extended line-shaped contrails used as 'training data' to produce an algorithm that is able to detect linear contrails.

RC: ***L 43-48: Maybe restructure section instead of jumping from Kärcher et al. (2009) to other papers and then back to Kärcher et al. (2009).***
L 43: Which model/simulation has been considered? Maybe note that there might be significant uncertainties in the contrail modelling as well.
L 48-49: Consider reformulating the statement depends not only on the optical thickness of a contrail, but also the microphysical properties, as the optical thickness depends also on the microphysical properties, so these are not distinct properties.

AR: We have restructured this paragraph and collected references to Kärcher et al. (2009) to improve readability. Included reference to the modelling technique and made this more balanced—that they were able to reconcile rather than implying any one source is 'more accurate'.
Also reformulated the reference to microphysics to clarify that microphysics contributes to optical thickness.

53–69

~~Kärcher et al. (2009) established that simulated and observed distributions of contrail optical thicknesses differ, and can be reconciled using optical-thickness-dependent detection efficiencies. Previously, studies had focused on relative detection efficiency due to surface inhomigeneities and overdetections due to false positives (Mannstein et al., 1999; Meyer et al., 2002; Minnis et al., 2005; Palikonda et al., 2005; Meyer et al., 2007) . The Kärcher et al. (2009) approach goes some way to enabling estimates of contrail coverage which align more closely with model output, albeit with detection efficiency inferred empirically based on the population simulated by the model~~Early studies of linear contrail detection in satellite imagery centre discussions of detection efficiency around the background conditions: surface inhomogeneities driving detection efficiency losses or false-positive overdetections (Mannstein et al., 1999; Meyer et al., 2002; Minnis et al., 2005; Palikond

. ~~It is also clear that contrail detectability depends not only on the optical thickness of a contrail, but also the microphysical properties (Yang et al., 2010). Increasingly prevelant~~ Later, Kärcher et al. (2009) established that the properties of the contrail also affect its observability. It was shown that the observed distribution of contrail optical thickness differs from the distribution produced by a model (CCSIM, an analytical model shown to be consistent with large eddy simulations). The observations underestimated the occurrence of optically thin contrails (with optical thickness < 0.2), relative to the model—an empirically-inferred optical-thickness-dependent detection efficiency was able to reconcile model with observation. Although optical thickness ought to be a good predictor of contrail observability, underlying microphysical properties (such as particle size and concentration) will have individual effects on the observability of contrails, depending on the techniques used for their detection (Yang et al., 2010).

RC: ***L 43-54: Check whether you can restructure this section to make it easier to follow. E.g., you mention first Kärcher et al. (2009) established that simulated and observed distributions of contrail optical thicknesses differ, and can be reconciled using optical-thickness-dependent detection efficiencies. Then CNNs are mentioned. And then some sentences later you introduce the concept again, writing If these satellite-observed contrails are to be used to evaluate model simulations on contrails, it is essential that the properties of the observing system are taken into account***

AR: We hope that by completing the account of the development of techniques including neural nets, then discussing limitations following a similar development process (aligning with the above revisions of the discussion of Kärcher et al. (2009)) make this easier to follow. We think it is important to give an initial introduction to CNNs before discussing optical thickness detection efficiency to make it clear that this applies to CNNs as well as older algorithms.
We have added to the point on satellite simulation to draw contrast between the value of satellite simulation in contrast to human-labelled data as a benchmark.

70–81

The adoption of convolutional neural net algorithms ~~are typically benchmarked against human-identifiable contrails used as training data (Meijer et al., 2022; Ng et al., 2023; Gryspeerdt et al., 2024) . If these satellite-observed contrails are to be used to evaluate model simulations on contrails, it is essential that the properties of the observing system are taken into account—a~~ brings with it a new set of limitations to discuss regarding their detection efficiency. These algorithms are typically trained and benchmarked against datasets of contrails identified by humans in satellite imagery, such datasets are also used as 'training data' (Meijer et al., 2022; Ng et al., 2023; Gryspeerdt et al., 2024). Using human-labelled datasets as a benchmark neglects the cases which are unobservable by human labellers (which are likely also unobservable by algorithmic methods), potentially leading to over-estimated detection efficiency.

A concept known as "satellite simulation" (e.g. Bodas-Salcedo et al., 2011) ~~. With~~ can aid in the essential analysis of contrail observability while not depending on the observational limits of a human labeller. These analyses should seek to account for varying instrument properties~~and a~~ , the wide variety of contrail ~~micro~~ micro- and macrophysical properties, and the background conditions; it is not clear that a simple optical depth threshold is suitable for this purpose.

RC: ***L 56: Maybe better: CoCiP produces predictions which generally align with observations regarding the order of magnitude and principle age-dependencies for micro- and macrophysical properties***

AR: Agree that the proposed replacement is more precise while remaining succinct! Adopted in the text (full revisions below).

RC: ***L 56-60: Four sentences starting with CoCiP. Try to reformulate due to style reasons.***

AR: We have also reformulated as suggested for style, and with the aim of making this discussion of applications less confusing (full revisions below).

RC: ***L 59-60: Note that CoCiP does not check for persistence, see Schumann (2012): An explicit criterion for persistency as a function of supersaturation is not necessary in CoCiP. In case of very low temperatures, short-lived contrails may form from the emitted water vapor even in totally dry air. I think the pycontrails implementation works similarly.***

AR: Agree that this was inaccurate in the text—thank you for picking up! Revised accordingly.
Full revisions for this discussion of model validation are shown below.

82–97
~~Validation of models,~~ Well-understood observations are required for a large number of reasons, for example, to validate models like CoCiP (the 'Contrail Cirrus Prediction model'; Schumann, 2012)~~requires well-understood observations~~. CoCiP produces predictions which generally align with observations ~~, insofar as their properties are resolved (Schumann et al., 2017). CoCiP predictions have been used for predicting contrail formation in order to guide~~ regarding the order of magnitude and principle age-dependencies for micro- and macrophysical properties (Schumann et al., 2017). Planning for in-situ observations ~~(Voigt et al., 2017). CoCiP~~ has been informed by CoCiP applied to forecast data (Voigt et al., 2017). The model has also been used to consider potential consequences of lower non-volatile particulate emissions following climate action, such as due to the adoption of Sustainable Aviation Fuel (SAF) (Teoh et al., 2022b). CoCiP is based on simple, well-understood criteria for formation and persistence, described in Schumann (1996). Most fundamentally, this is driven by Schmidt–Appleman temperature threshold for mixing cloud formation~~and ice-supersaturation conditions for persistence (with appropriate adaptations, for example, for propulsion efficiency)~~. The properties of the formed contrail are then allowed to evolve, subject to ice water content changes in response to ambient supersaturation, diffusion, and particle number loss processes (Schumann, 2012). Meteorological input limits the predictability of persistent contrail formation, particularly uncertainty in relative humidity values at flight altitudes (Gierens et al., 2020; Agarwal et al., 2022). Model adaptations, such as the correction of ~~relative humidity measurements~~ input relative humidity from meteorology (Teoh et al., 2022a) or alterations to contrail processes like ice crystal formation and loss mechanisms (Schumann et al., 2017) ~~,~~require well-understood contrail observations for validation.

RC: ***L 66: Specify tactical avoidance application***

AR: Added definition and clarified the discussion.

98–107

Beyond model validation, other applications have a varied range of specific observational needs. ~~Climate monitoring and tactical avoidance applications (Geraedts et al., 2023; Chevallier et al., 2023; Sausen et al., 2023)~~ For example, applications include climate monitoring and operational tactical avoidance (action taken in response to current condi. These applications require observation of as high a proportion of strongly forcing contrails as possible (~~and~~ at least at some point in their evolution)~~. The same applications~~ , and often need to match contrails to flights ~~and dynamically avoid contrails~~or take quick action, so observation as quickly after formation as possible is required. ~~Avoidance trials~~ The planning of experimental trials seeking to analyse changes to contrail formation or persistence will benefit from an understanding of the dependence of observability on the properties of contrails and the surrounding conditions, so that they can be confident that unobserved contrails are indeed unformed contrails (Molloy et al., 2022).

RC: ***L 73-74: Suggest to reformulate: Contrail detectability is tested in otherwise-clear-sky synthetic satellite images by applying a contrail detection algorithm.***

AR: Adopted as suggested, thank you! (Revisions included below)

RC: ***L 74: Where does this baseline contrail population come from? Is it modelled?***

AR: Clarified at this point of introduction that these are CoCiP-derived and from Teoh et al. (2024). (Revisions included with below)

RC: ***L 83: What means radiative importance conclusions?***

AR: Clarified these as "comparisons of the relative radiative impact of contrails".
Changes to this outline are shown below.

108–125

This study establishes limits of ~~contrail observability~~ observability for the automated detection of line-shaped contrails as a function of the contrail properties, independently from models. ~~Contrail detection is simulated~~ The detectability of linear contrails is tested in otherwise-clear-sky synthetic ~~radiance fields,~~ satellite images by applying a contrail detection algorithm~~that reflects the contrails visible in-context~~. The derived observability threshold will then be compared with ~~a baseline contrail population and its~~ CoCiP-modelled populations of global contrails and their estimated radiative forcing (from Teoh et al., 2024), and the consequences for a range of applications will be considered. In Section ~~2.1, the simulated contrail images~~ 2, key components of the observability analysis are described, including ~~radiative transfer simulations used to model them . These images are tested using~~ the simulated contrail images and the models used to create them (Section 2.1), the line-filtering ~~detection algorithm described in Section 2.2. A~~ contrail detection algorithm (Section 2.2), and the modelled population of ~~contrails, detailed in Section 2.3, is used to consider the consequences of derived observability thresholds~~global contrails (Section 2.3). Observability assessments are made in Section 3, including varying single parameters (Section 3.1) and ~~derivation of~~ deriving an observability threshold against the key observability-driving properties (Section 3.2)—properties which form a parameter space in which the population of contrails is shown to be well-resolved (Section 3.3). The ~~resulting proportion of contrails and contrail forcing that can be observed are found~~ derived observability threshold is finally applied to the properties of contrails modelled to form in Section 4, resulting in estimates of the observable fraction (Section 4.1), ~~as well as~~ the evolution of this observability with contrail aging (Section 4.2)~~. The validity of radiative importance conclusions made from the observed lifetime is~~

~~scrutinised in Section 4.3, followed by a consideration of~~ , the lifetime radiative impact of contrails based on their observed lifetime (Section 4.3), and the changing observability as climate action is taken (Section 4.4).

RC: ***Fig. 1: Maybe reformulate: Schematic of the process for deriving the contrail detection efficiency by application of a contrail detection algorithm to synthetic satellite observations of a single contrail, using a specific imager and contrail detection algorithm, and a pre-calculated radiative transfer lookup table.***

AR: Adopted revision as suggested—thank you!

RC: ***Fig. 2: What denotes the title for the three panels (0.5 km, 1 km, 2km)? Are the simulated calibration error and the NEdT given somewhere in the plot? Which brightness temperature channel is shown?***

AR: Clarified caption and removed the titles (in favour of describing the resolutions, imager band, noise and calibration in the caption).

RC: ***L 86: You maybe want to check with Schumann et al. (10.1175/JAMC-D-11-0242.1, 2012) and Wolf et al. (10.5194/acp-23-14003-2023, 2023), both describing comprehensive radiative transfer calculations for contrails. You might want to mention here or elsewhere in the manuscript how your assumptions agree or disagree with those studies.***

AR: A comparison with these two pieces of work has been included in the revised manuscript.

171–181

Radiative transfer modelling of contrails has been previously performed by Schumann et al. (2012) and Wolf et al. (2023) . Both these previous works similarly use libRadtran (Emde et al., 2016) to perform calculations. Each use similar ice cloud parameterisation settings to this work, with Schumann et al. (2012) using an earlier set of scattering properties. The previous works similarly suffer from a lack of scattering properties for small crystals, and the Schumann et al. (2012) work turns to Mie calculations for this purpose, as is done in this work. For larger crystals, each of the two previous works choose to use a range of ~~resolutions.~~ different habits which may be found in contrails—this was omitted here with the aim of choosing a habit that was consistent with the Mie calculations. Similarly, the other works use a range of atmospheric profiles. The approach taken here is aligned with the aim of considering an idealised case to provide an upper limit of the detectability; variation in habit mixtures and atmospheric profiles will add further variability to the detection efficiency achievable in practice. The other works have additional considerations for solar radiation, including solar zenith angle variations—neglected in this work due to the focus on thermal radiation for the detection algorithm used.

RC: ***L 93-101: What is the setting for the ice water path IWP_0? What parameter space is covered for r_eff? Are slanted observations considered, i.e., non-zero viewing zenith angles?***

AR: The radiative transfer calculations are performed for the range given by the lookup table (ranges described). The ranges of contrail properties (i.e. IWP_0 and r_{eff} and width) varies based on the analysis, but the values have been added in Section 3.2.

320–331

~~IWP, r_{eff}, and contrail width~~ The analysis of Section 3.1 leads to the identification of a parameter space consisting of those properties which have a strong control on contrail observability~~(Fig. 3). This aligns with expectations: contrail observability ought to depend on the contrail optical thickness (to~~

~~first order, a combination of r_{eff} and IWP), and the interaction of the contrail's geometry with the imager's pixel resolution. A sensitivity to altitude was also observed for narrow contrails~~ : IWP_0, r_{eff}, and contrail width. The observability of a contrail is tested in this parameter space by covarying these three parameters. The population has been split into logarithmically-spaced bins in each dimension—including contrail width bins between 0.025 and 25 km wide, r_{eff} bins between 0.1 and 50 µm, and IWP_0 between 10^{-3} and $10^{3/2}$ $\text{g}\,\text{m}^{-2}$. We neglect altitude variations, because it only had a weak observability effect (Fig. 4), ~~although dependence on altitude is weak at the higher altitudes where most contrails exist. An altitude~~ so an altitude of 11 km ~~has been~~ is assumed for all ~~contrails—this~~ contrails. This altitude is approximately consistent with the modal altitude (Fig. 3(d)), and aligns with the aim of a maximally-observable case (Fig. 4).

The approximation of zero viewing zenith angle has been added to the description of the radiative transfer simulations, alongside details requested by another review (the changes are as quoted above).

150–152
The viewing zenith angle has been taken to be zero, introducing an assumption that observation occurs near the satellite's nadir. The surface is assumed to have unit thermal emissivity (similar conditions to the ocean), and the surface temperature is treated as 15 °C (a property of the temperature profile).

Viewing zenith angle effects are also briefly mentioned in the conclusions, where the effects of enhanced optical path and reduced imager resolution are discussed.

RC: ***L 145: Maybe only a language issue, but what is the Mannstein et al. (1999) style detector?***

AR: This was a slightly confusing phrasing, it has been revised for clarity.

205–207
This is a result of the contrail profiles terminating suddenly, producing a regional gradient in brightness temperature fields, ~~highlighted by the Mannstein et al. (1999) style detector~~which is picked out by the detection algorithms used in this work.

RC: ***L 175: I suggest to say that it is assumed that the CoCiP-based statistics aligns with reality. More validation is necessary at this point.***

AR: We agree, and have changed 'expected' to 'assumed'.

RC: ***L 181: Suggest to speak of contrail segments instead of waypoints.***

AR: Changed 'waypoints' to 'segments'.

RC: ***L 196: Please clarify which globally-consistent thresholds are meant.***

AR: The original submission misstated the diurnal binning as local solar time, and should have read UTC ('globally-consistent thresholds' was intended to describe the use of a simulataneous global daytime, and global nighttime bins—highlighting global observability variations rather than local).
Reflecting on this comment led us to find that variability due to time-of day did not contribute a significant source of variability, other than for SW RF, because no SW forcing occurrs during the night. As a result, the treatment of diurnal variability has been removed (change **C.2**).
Please note that the discussion of variability has been moved to section 4.1.1, in respose to feedback from another review.

437–442

The ~~component due to seasonal and diurnal variability is also shown independently on Fig. 7. The combined uncertainty is dominated by the seasonal and dirunal component, indicating that the observability threshold is well-defined with respect to different realisations of the noise field~~impact of seasonal variability is assessed by taking the standard deviation of the observable fraction obtained when subsetting CoCiP population 1 based on the month of the year. Both the combination of the two uncertainties, and the uncertainty due to variability are shown in Fig. 7, and the variability in contrail properties clearly dominates the component due to instrument noise in this idealised case.

We have also noted in the conclusions that observing systems with a biased time of observation could bias the population properties.

RC: ***L 199: Maybe replace clear by apparent?***

AR: Changed this in the text.

RC: ***L 211: I assume some of these properties (e.g., width and depth) might be related? So a limitation of this approach is that some of the configurations you consider are not realistic?***

AR: Added a brief discussion of this limitiation in the text.

277–280

This analysis is intended to abstract the observability consequence of the each individual contrail property from the consequence of the others, so no attempt is made to ensure that contrails simulated in the course of these observability tests are realistic, including no accounting for properties which are likely to covary (such as width and depth).

RC: ***L 217: For IWP, no upper detection limit is shown for the 2km resolution in Fig. 3b. Do you expect this limit to be at even higher IWPs? Why arent the limits for both imager resolutions at the same value?***

AR: A slightly higher upper-limit on IWP_0 in the coarser-resolution imager is consistent with expectation but is worth expanding on. The difference is that a coarser pixel contains both the very optically thick centre and part of the more optically thin edges of the contrail further from the central peak IWP. The effect of the coarse pixel is then to 'average' this and reduce the effect of the high IWP. In this case, the averaging with lower-IWP parts of the contrail causes a stronger signal (because the brightness temperature difference is lost if the contrail is too optically thick). Using the higher resolution imager, for this contrail width, when the high IWP_0 limit is reached, the signal is not strong enough anywhere to produce a linear signal detected as a contrail using this algorithm.

286–290

The analogous limit for the coarser-resolution imager is not seen, despite the expectation that the opacity as a function of IWP is resolution-independent. The limit occurs at higher IWP_0 for the coarser-resolution imager because the effect of the peak IWP at the centre of the profile is 'averaged' over the pixel, which would include some lower-IWP of the profile (eq. 1) further from the centre. For this contrail, with this particular width, the signals in the synthetic image are simply not strong enough to cause a detection.

RC: ***L 223: It seems like the observable range in Fig. 3c can maybe related to the imager resolution? Maybe from 1/5 to 5 times the imager resolution? A consequence would be that for the newest generation of imagers some kind of downsampling is necessary to detect broader contrails, right?***

AR: This downscaling would work, but the high-width limit is specific to this detection algorithm (and this baseline contrail's set of microphysical properties). The upper limit on width would not be such a 'hard limit' for a CNN (which wouldn't have a pre-specified line filter of finite width). The discussion of Fig. 3(c) has been expanded to clarify this.

> **298–307**
> Wide contrails are also not detected. This is a limitation of the detection algorithm used—which uses line kernels that highlight linearly-extended regional gradients 1–4 pixels wide. The CoCiP population has a significant spread in width, including contrails with widths of up to several hundred kilometres, so this algorithm limitation would have a significant effect if untreated. ~~These wide unobservable contrails are~~ Detection occurs within different upper- and lower-limits of width for each of the imagers tested. The high-width limit is likely to be ~~detectable~~ overcome if a different approach is used, given that narrower contrails with similar microphysics ~~have been identified~~ are detectable using this algorithm. ~~To assess the impact of a less conservative contrail detection algorithm, detectability thresholds are adjusted to only include low-width unobservability behaviour, detailed in Section 3.2 (the adjusted threshold in Fig. 5).~~ For example, the observation could be downsampled onto a coarser grid before applying the detection algorithm. As a result, it is reasonable to consider all contrails wider than the narrowest-detectable contrail with a given set of microphysical properties to be detectable.

The width limits vary based on the microphysics (Fig. 8 shows the detection limits as the 'minimum width' for which a contrail can be detected). The detectability response of Fig. 3c doesn't therefore can't be straightforwardly applied as a rule-of-thumb.

RC: ***Fig. 7: The seasonal and diurnal variability complete coincides with the combined uncertainty, right? However, I think the seasonal and diurnal variability is hardly visible right now. I suggest to plot this differently.***

AR: We have altered the way we plot the errorbars to make the seasonal variability clearer. The revised figure and previous version is included at the end of this document.

RC: ***L 354: Why was the proportion of LW forcing and not the net forcing considered? The latter is the crucial quantity.***

AR: Using LW forcing enables a simpler discussion when considering a 'proportion', because LW forcing is a simple scalar. Considering a proportion of total net forcing doesn't make sense because individual strongly positive or negatively forcing contrails could have undue impact when LW and SW components largely cancel on average, and the proportion would no longer be constrained between 0 and 1. The alternatives of 'total absolute net forcing' or 'net forcing of net warming contrails' were considered but ultimately deemed less-meaningful than a discussion of LW forcing. A comment describing this has been included.

> **520–522**
> LW forcing has been used, because its positive definite nature simplifies calculations and reduces variability in forcing estimates due to solar zenith angle (meaning the results are instead focussed on the link between contrail properties and radiative importance).

A similar comment has been added to Section 4.3.

593–594
LW energy forcing is used (as in Section 4.2) rather than net forcing, so that net forcing variability due to the variations in SW forcing with time-of-day do not overwhelm the variability estimates.

Consideration of these comments contributed to changing to more-coherent definitions for the earlier proportions (change **C.1**).

RC: ***L 445: The Mannstein algorithm checks only for linearity and is applicable to detect contrails in the early phase of the lifecycle, where this linearity is present. In principle, detectability might be increased by combining different detection methods or apply tracking procedures to observe contrails also beyond their linear stage, e.g., Vazquez-Navarro et al. (2010, 10.5194/amt-3-1089-2010), Vazquez-Navarro et al. (2015, 10.5194/acp-15-8739-2015).***

AR: While true that the algorithm checks only for linearity, all of the synthetic contrails are inherently linear, and the CoCiP populations are treated as individual linear segments. Combined with our minimally-restrictive approach of imposing only a minimum-width limit to detectability, the results presented should still stand of those assessed by application of a detection algorithm. However, we agree that targeted observation in successive timesteps after initial detection is likely to increase detectability (consistent with the results from the less-conservative detection algorithm tested in Section 4.1.2). Reference to this, and its implication for contrail tracking, has been added to Section 5.1.

638–643
Conversely, observability may be increased when targeted observation of a contrail known to exist is possible (where a less-conservative detection algorithm could be applied, as discussed in Section 4.1.2). This is well-illustrated by the work of Vazquez-Navarro et al. (2010) and Vázquez-Navarro et al. (2015) , demonstrating that contrails detected using a higher-resolution non-geostationary imager enabled targeted observation in coarser-resolution geostationary observations for contrail evolution to be tracked. This tracking method also stands to enable the consideration of detected contrails beyond their initial linear phase.

RC: ***In the paper you considered imager resolutions of 0.5 and 2 km. MSG/SEVIRI used in previous studies had resolution of 3 km at nadir. Can you make any comments on its performance?***

AR: We recognise that it could be useful to expand this analysis to a range of coarser instruments. We chose to include a 7 km resolution test case, which is representative of MSG-SEVIRI at mid-latitudes (change **C.4**). We think providing a significantly coarser case (rather than an incremental one) can provide more insight. The derived results are shown in the revised Fig. 7 and Fig. 8 (included at the end of this document).

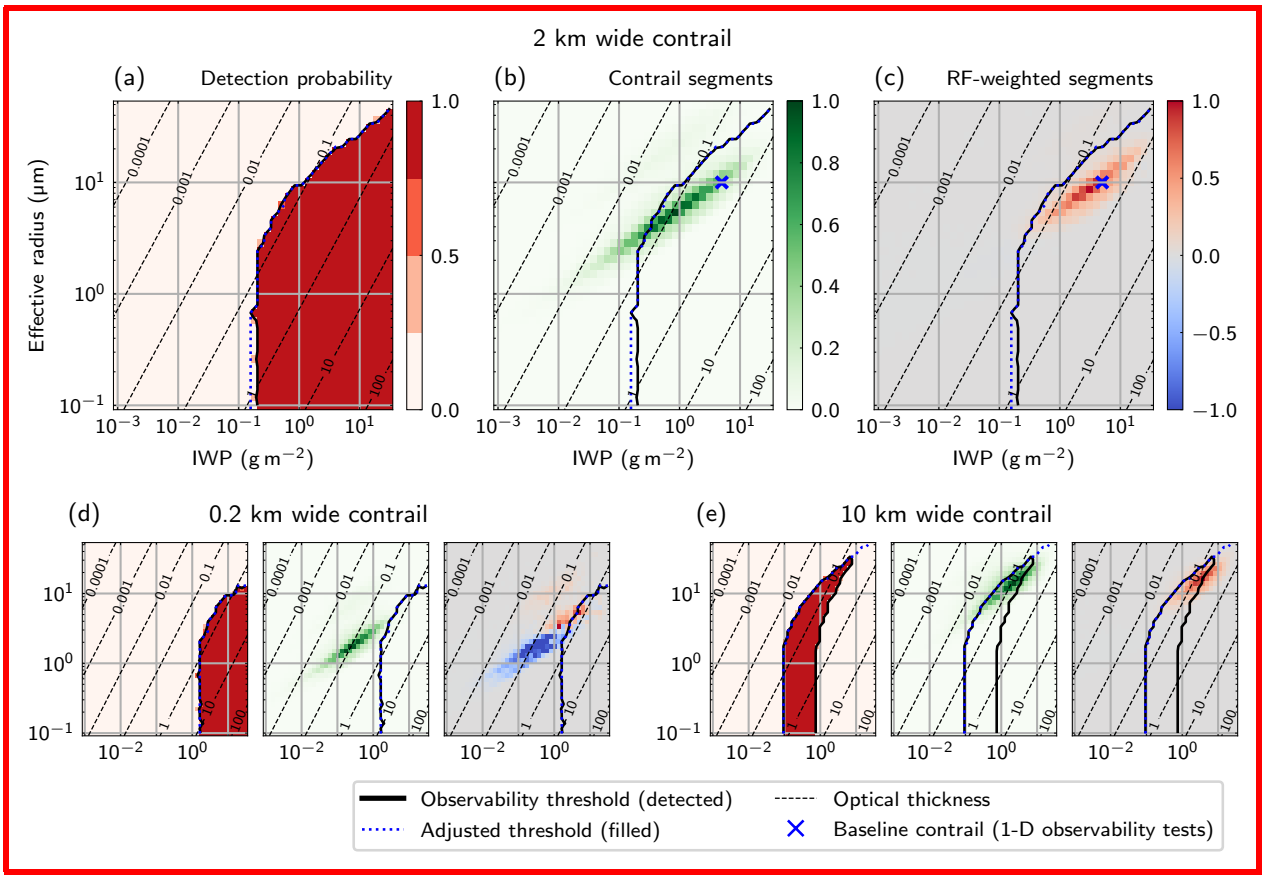
RC: ***Technical corrections***
L 53: micro-
L 207: width mentioned twice.
Fig. 4: contrails as width ?
L 285: on doubled
Fig. 9: Perisisting –> Persisting
L 336: Incomplete sentence?
L 389: was also found is plotted ?
L 418: dominated by soot while current fuels ?

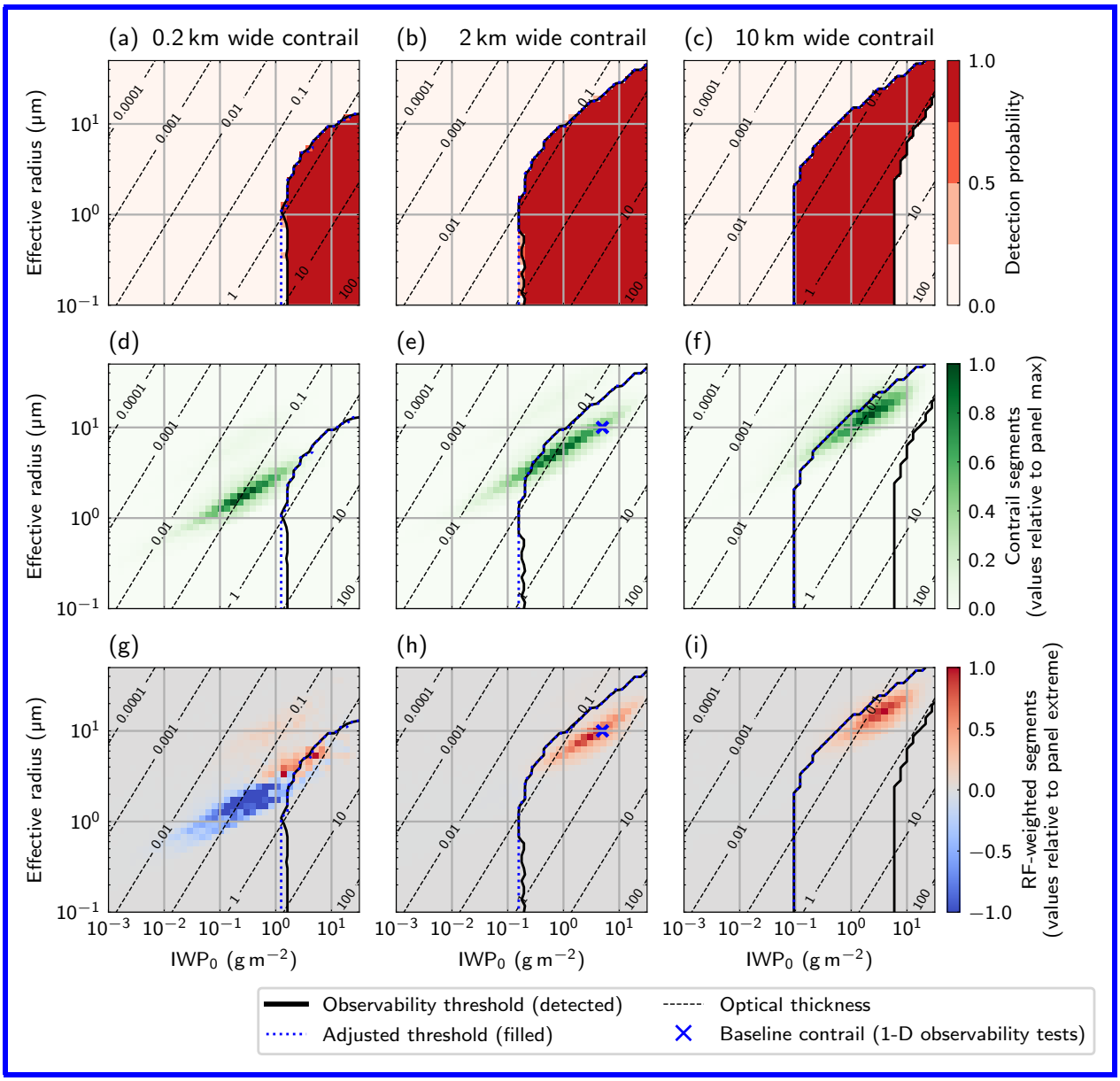
AR: Thank you for spotting these errors. Each has been revised.

Significantly altered figures

Figure 5

The strucutre of this plot has been simplified to comprise evenly-sized panels. Additionally, after making change **C.1**, the interpretation of this plot has been made simpler (although the relative values in the RF-weighted histogram have not changed—because each panel has a fixed contrail width).





The caption has been changed as follows:

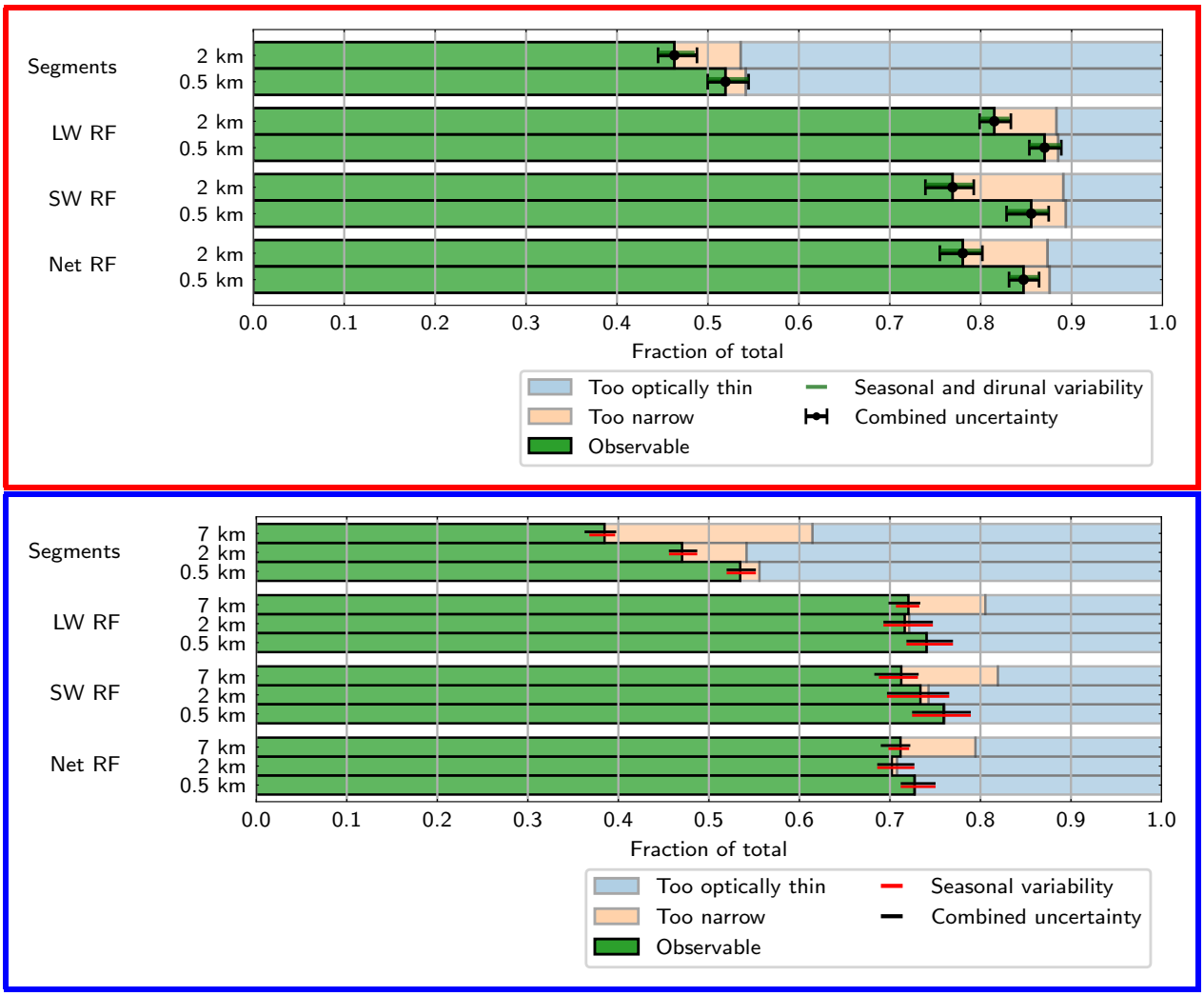
Figure 5 caption

Illustrative slices from a ~~3-D~~ contrail observability test ~~, including~~for three contrail-width bins (chosen from 30 total bins): ~~observability quantified using effective width~~ 0.2 km- (a, d, g), ~~and histograms of CoCiP populations weighted by occurrence~~ 2 km- (b~~)~~, e, h) and ~~net radiative forcing~~ 10 km-wide (c~~)~~, ~~for 2 km wide contrail segments~~f, i) contrails. ~~Similar~~ The observability threshold is plotted over the derived detection probability (a–c), and histograms ~~are shown for 0.2 km wide contrails~~ of CoCiP population 1 (~~d~~d–f) and ~~15 km wide~~ of the same population weighted by the mean net RF of the

contrails in each bin (eg–i). The histogram values (d–i) are relative to the magnitude of the extreme value in each plot. The 'adjusted threshold' represents the observability threshold adapted to include all theoretically observable contrails, without the high-width algorithm deficiency. Contours of the contrail optical thickness (estimated based on IWP_0 and r_{eff}) are also shown.

Figure 7

AR: Subject to change **C.1**, the RF-weighted detection efficiencies have changed, and the variability has been limited to only seasonality (change **C.2**). Additionally, following change **C.4**, the observability using a 7 km imager is shown. Finally, the errorbars have been separated to make them clearer.



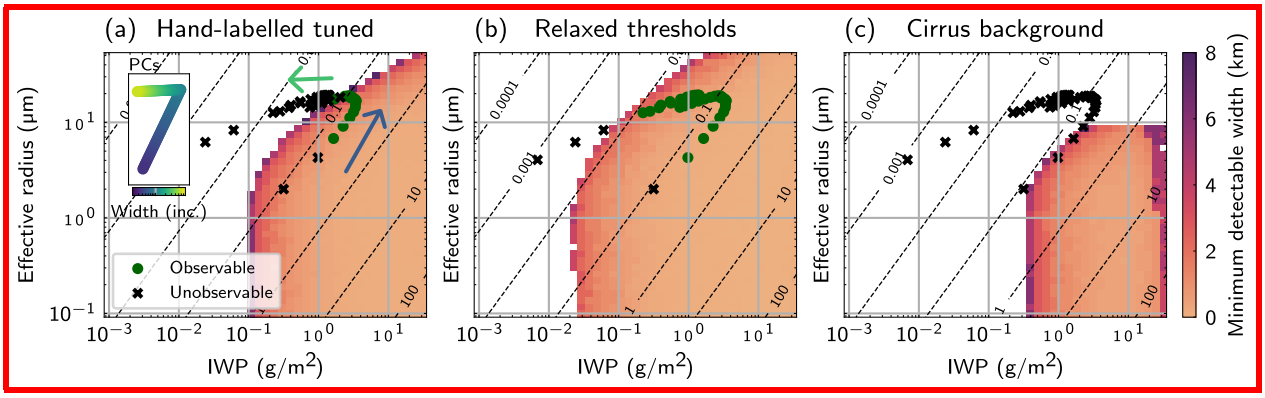
The caption has been changed as follows:

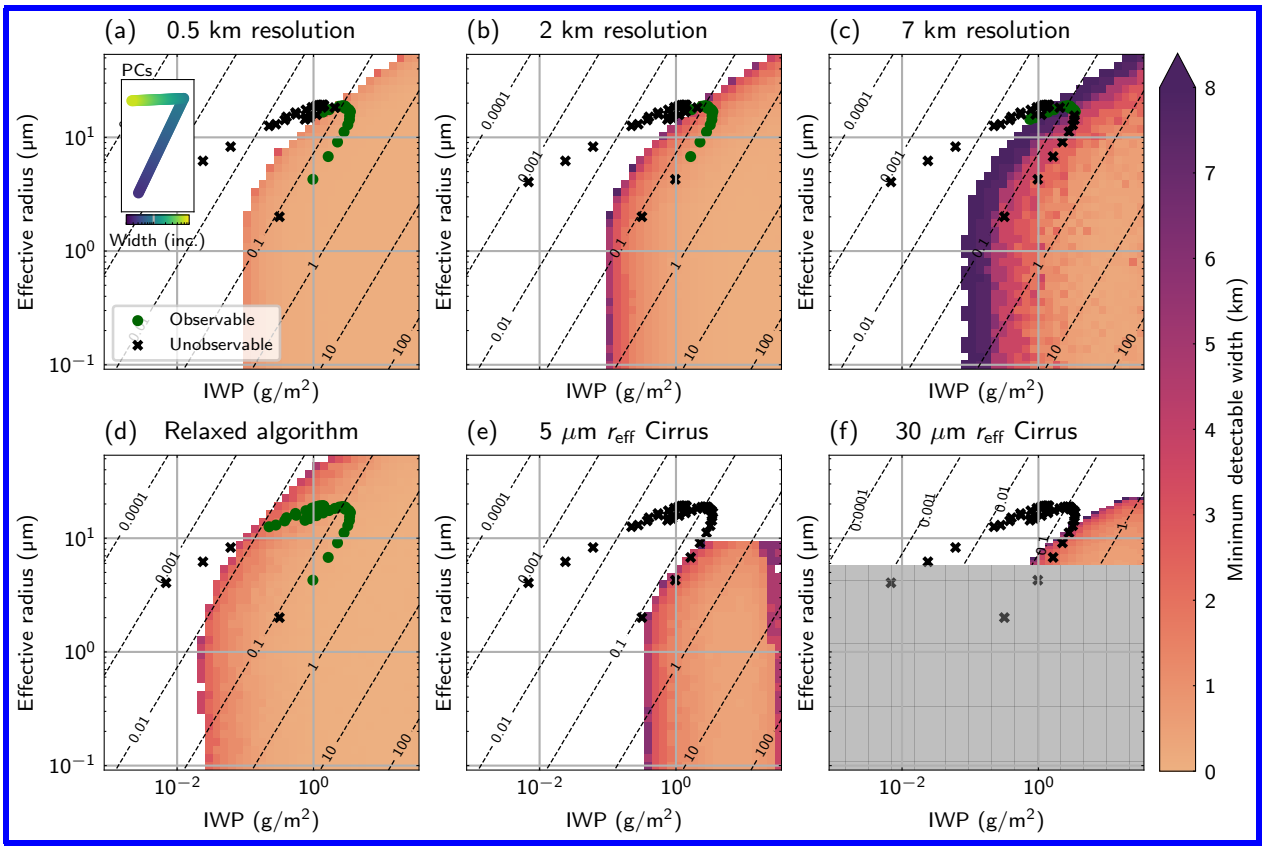
Figure 7 caption

The fraction of contrail segments that are theoretically observable ~~fraction~~ using imagers with 7, 2, and 0.5 km spatial resolution. Shown as a proportion of contrail segments, and weighted by their instantaneous ~~SW,~~ LW, SW, and absolute net ~~radiative forcing, based~~ RF (per unit length). Based on the distribution of properties modelled in CoCiP ~~populations~~ population 1 using Jet-A1 fuel. Unobservable contrails are categorised as either too narrow or too optically thin to be observable. Error bars are a combination of the variability in contrail properties due to seasonal ~~and diurnal~~ effects (also shown independently), and uncertainty in the derived threshold based on the observable proportion using 0.25 and 0.75 ~~P_{obs}~~ p_{obs} thresholds.

Figure 8

For clarity, the minimum width thresholds for the three resolutions is shown, as well as that in the less-conservative and two background cirrus cases (following change **C.3**).





The caption has been changed as follows:

Figure 8 caption

The theoretically observable fraction of contrail segments and instantaneous radiative forcing (as Fig. 7), based on ~~the~~ CoCiP ~~populations~~ population 1, assuming ~~a~~ fractional adoption of SAF biofuel leads to reduced effective emission of ice. ~~Observability has been tested for a simulated~~ A 2 km spatial resolution imager ~~for contrails in an otherwise clear sky~~is used.~~Error bars indicate the observable proportion using 0.25 and 0.75 P_{obs} thresholds.~~

Figure 12

Similar to Fig. 7, this has been updated with the revised RF weighting following change **C.1**, and errorbars following change **C.2**.

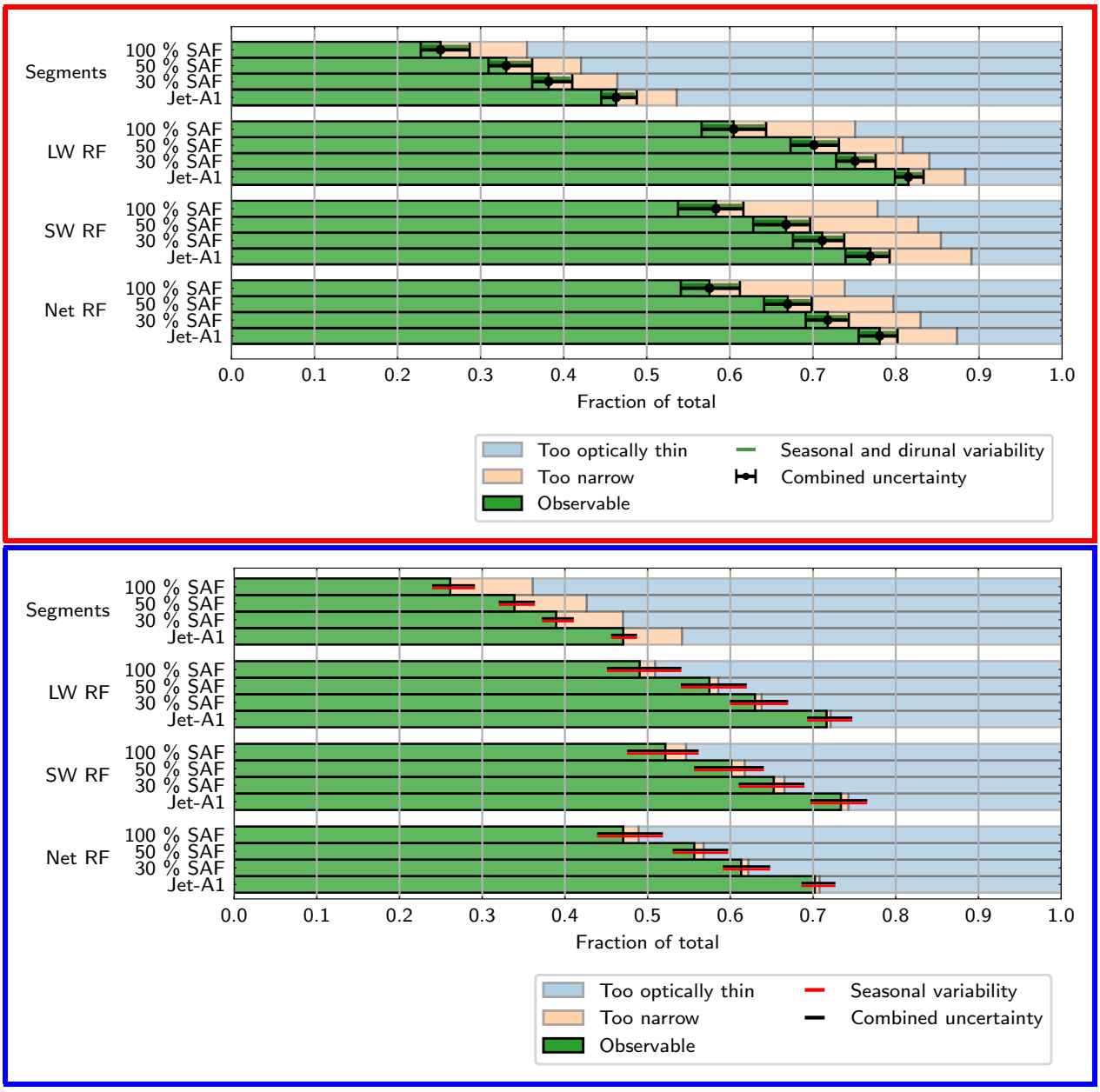


Figure 12 caption

The theoretically observable fraction of contrail segments and instantaneous radiative forcing (as Fig. 7), based on ~~the~~ CoCiP ~~populations~~ population 1, assuming ~~a~~ fractional adoption of SAF biofuel leads to reduced effective emission of ice. ~~Observability has been tested for a simulated~~ A 2 km spatial resolution imager ~~for contrails in an otherwise clear sky~~is used.~~Error bars indicate the observable proportion using 0.25 and 0.75 P_{obs} thresholds.~~

References

Agarwal, A. et al. (Jan. 2022). "Reanalysis-Driven Simulations May Overestimate Persistent Contrail Formation by 100%–250%". In: *Environmental Research Letters* 17.1, p. 014045. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac38d9. (Visited on 04/05/2024).

Bodas-Salcedo, A. et al. (Aug. 2011). "COSP: Satellite Simulation Software for Model Assessment". In: *Bulletin of the American Meteorological Society* 92.8, pp. 1023–1043. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/2011BAMS2856.1. (Visited on 06/03/2024).

Chevallier, R. et al. (July 2023). "Linear Contrails Detection, Tracking and Matching with Aircraft Using Geostationary Satellite and Air Traffic Data". In: *Aerospace* 10.7, p. 578. ISSN: 2226-4310. DOI: 10.3390/aerospace10070578. (Visited on 08/07/2023).

Emde, C. et al. (May 2016). "The libRadtran Software Package for Radiative Transfer Calculations (Version 2.0.1)". In: *Geoscientific Model Development* 9.5, pp. 1647–1672. ISSN: 1991-959X. DOI: 10.5194/gmd-9-1647-2016. (Visited on 02/17/2023).

Gasteiger, J. et al. (Nov. 2014). "Representative Wavelengths Absorption Parameterization Applied to Satellite Channels and Spectral Bands". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 148, pp. 99–115. ISSN: 0022-4073. DOI: 10.1016/j.jqsrt.2014.06.024. (Visited on 09/08/2023).

Geraedts, S. et al. (Dec. 2023). *A Scalable System to Measure Contrail Formation on a Per-Flight Basis*. DOI: 10.48550/arXiv.2308.02707. arXiv: 2308.02707 [physics]. (Visited on 04/08/2024).

– (Jan. 2024). "A Scalable System to Measure Contrail Formation on a Per-Flight Basis". In: *Environmental Research Communications* 6.1, p. 015008. ISSN: 2515-7620. DOI: 10.1088/2515-7620/ad11ab. (Visited on 10/15/2024).

Gierens, K. et al. (Dec. 2020). "How Well Can Persistent Contrails Be Predicted?" In: *Aerospace* 7.12, p. 169. ISSN: 2226-4310. DOI: 10.3390/aerospace7120169. (Visited on 01/19/2023).

Gryspeerdt, E. et al. (2024). *Operational Differences Lead to Longer Lifetimes of Satellite Detectable Contrails from More Fuel Efficient Aircraft*.

Kärcher, B. et al. (Aug. 2009). "Factors Controlling Contrail Cirrus Optical Depth". In: *Atmospheric Chemistry and Physics* 9.16, pp. 6229–6254. ISSN: 1680-7316. DOI: 10.5194/acp-9-6229-2009. (Visited on 06/03/2024).

Lee, T. F. (Sept. 1989). "Jet Contrail Identification Using the AVI-IRR Infrared Split Window". In: *Journal of Applied Meteorology and Climatology* 28.9, pp. 993–995. ISSN: 1520-0450, 0894-8763. DOI: 10.1175/1520-0450(1989)028<0993:JCIUTA>2.0.CO;2. (Visited on 04/05/2023).

Mannstein, H. et al. (Jan. 1999). "Operational Detection of Contrails from NOAA-AVHRR-data". In: *International Journal of Remote Sensing* 20.8, pp. 1641–1660. ISSN: 0143-1161, 1366-5901. DOI: 10.1080/014311699212650. (Visited on 10/04/2022).

McCloskey, K. et al. (July 2021). "A Human-Labeled Landsat-8 Contrails Dataset". In: *Climate Change AI*. https://www.climatechange.ai/papers/icml2021/2. Climate Change AI. (Visited on 04/03/2024).

Meijer, V. R. et al. (Mar. 2022). "Contrail Coverage over the United States before and during the COVID-19 Pandemic". In: *Environmental Research Letters* 17.3, p. 034039. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac26f0. (Visited on 10/04/2022).

Meyer, R. et al. (May 2002). "Regional Radiative Forcing by Line-Shaped Contrails Derived from Satellite Data: CONTRAIL RADIATIVE FORCING FROM SATELLITE DATA". In: *Journal of Geophysical Research: Atmospheres* 107.D10, ACL 17-1-ACL 17–15. ISSN: 01480227. DOI: 10.1029/2001JD000426. (Visited on 11/23/2022).

Meyer, R. et al. (May 2007). "Contrail Observations over Southern and Eastern Asia in NOAA/AVHRR Data and Comparisons to Contrail Simulations in a GCM". In: *International Journal of Remote Sensing*

28.9, pp. 2049–2069. ISSN: 0143-1161, 1366-5901. DOI: 10.1080/01431160600641707. (Visited on 11/23/2022).

Minnis, P. et al. (Aug. 2005). "Contrail Properties over the Eastern North Pacific from AVHRR Data". In: *Meteorologische Zeitschrift* 14, pp. 515–523. DOI: 10.1127/0941-2948/2005/0056.

Molloy, J. et al. (July 2022). "Design Principles for a Contrail-Minimizing Trial in the North Atlantic". In: *Aerospace* 9.7, p. 375. ISSN: 2226-4310. DOI: 10.3390/aerospace9070375. (Visited on 04/08/2024).

Ng, J. Y.-H. et al. (Apr. 2023). *OpenContrails: Benchmarking Contrail Detection on GOES-16 ABI*. http://arxiv.org/abs/2304.02122. arXiv: 2304.02122 [cs]. (Visited on 04/20/2023).

Palikonda, R. et al. (Sept. 2005). "Contrail Coverage Derived from 2001 AVHRR Data over the Continental United States of America and Surrounding Areas". In: *Meteorologische Zeitschrift* 14.4, pp. 525–536. ISSN: 0941-2948. DOI: 10.1127/0941-2948/2005/0051. (Visited on 06/03/2024).

Sausen, R. et al. (Jan. 2023). "Can We Successfully Avoid Persistent Contrails by Small Altitude Adjustments of Flights in the Real World?" In: *Meteorologische Zeitschrift*. ISSN: , DOI: 10.1127/metz/2023/1157. (Visited on 04/06/2024).

Schumann, U. (May 2012). "A Contrail Cirrus Prediction Model". In: *Geoscientific Model Development* 5.3, pp. 543–580. ISSN: 1991-9603. DOI: 10.5194/gmd-5-543-2012. (Visited on 10/04/2022).

Schumann, U. et al. (July 2012). "A Parametric Radiative Forcing Model for Contrail Cirrus". In: *Journal of Applied Meteorology and Climatology* 51.7, pp. 1391–1406. ISSN: 1558-8424, 1558-8432. DOI: 10.1175/JAMC-D-11-0242.1. (Visited on 12/07/2023).

Schumann, U. (Mar. 1996). "On Conditions for Contrail Formation from Aircraft Exhausts". In: *Meteorologische Zeitschrift* 5.1, pp. 4–23. ISSN: 0941-2948. DOI: 10.1127/metz/5/1996/4. (Visited on 10/04/2022).

Schumann, U. and A. J. Heymsfield (Jan. 2017). "On the Life Cycle of Individual Contrails and Contrail Cirrus". In: *Meteorological Monographs* 58.1, pp. 3.1–3.24. DOI: 10.1175/AMSMONOGRAPHS-D-16-0005.1. (Visited on 10/06/2022).

Schumann, U. et al. (Jan. 2017). "Properties of Individual Contrails: A Compilation of Observations and Some Comparisons". In: *Atmospheric Chemistry and Physics* 17.1, pp. 403–438. ISSN: 1680-7316. DOI: 10.5194/acp-17-403-2017. (Visited on 01/27/2023).

Shapiro, M. et al. (Nov. 2023). *Pycontrails: Python Library for Modeling Aviation Climate Impacts*. Zenodo. DOI: 10.5281/zenodo.10182539. (Visited on 04/03/2024).

Teoh, R. et al. (Aug. 2022a). "Aviation Contrail Climate Effects in the North Atlantic from 2016 to 2021". In: *Atmospheric Chemistry and Physics* 22.16, pp. 10919–10935. ISSN: 1680-7324. DOI: 10.5194/acp-22-10919-2022. (Visited on 10/04/2022).

Teoh, R. et al. (Dec. 2022b). "Targeted Use of Sustainable Aviation Fuel to Maximize Climate Benefits". In: *Environmental Science & Technology* 56.23, pp. 17246–17255. ISSN: 0013-936X. DOI: 10.1021/acs.est.2c05781. (Visited on 03/14/2023).

Teoh, R. et al. (May 2024). "Global Aviation Contrail Climate Effects from 2019 to 2021". In: *Atmospheric Chemistry and Physics* 24.10, pp. 6071–6093. ISSN: 1680-7316. DOI: 10.5194/acp-24-6071-2024. (Visited on 06/10/2024).

Vazquez-Navarro, M. et al. (Aug. 2010). "An Automatic Contrail Tracking Algorithm". In: *Atmospheric Measurement Techniques* 3.4, pp. 1089–1101. ISSN: 1867-1381. DOI: 10.5194/amt-3-1089-2010. (Visited on 10/16/2024).

Vázquez-Navarro, M. et al. (Aug. 2015). "Contrail Life Cycle and Properties from 1 Year of MSG/SEVIRI Rapid-Scan Images". In: *Atmospheric Chemistry and Physics* 15.15, pp. 8739–8749. ISSN: 1680-7316. DOI: 10.5194/acp-15-8739-2015. (Visited on 12/03/2022).

Voigt, C. et al. (Feb. 2017). “ML-CIRRUS: The Airborne Experiment on Natural Cirrus and Contrail Cirrus with the High-Altitude Long-Range Research Aircraft HALO”. In: *Bulletin of the American Meteorological Society* 98.2, pp. 271–288. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-15-00213.1. (Visited on 02/24/2023).

Wolf, K. et al. (Nov. 2023). “Sensitivity of Cirrus and Contrail Radiative Effect on Cloud Microphysical and Environmental Parameters”. In: *Atmospheric Chemistry and Physics* 23.21, pp. 14003–14037. ISSN: 1680-7316. DOI: 10.5194/acp-23-14003-2023. (Visited on 10/30/2024).

Yang, P. et al. (Apr. 2010). “Contrails and Induced Cirrus: Optics and Radiation”. In: *Bulletin of the American Meteorological Society* 91.4, pp. 473–478. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/2009BAMS2837.1. (Visited on 03/03/2023).