Journal of Open Aviation Science (2023), Vol.**1**
doi:10.59490/joas.2023.7545

JOAS

EDITORIAL

Reviews and Responses for AI-Driven Identification of Contrail Sources: Integrating Satellite Observations and Air Traffic Data

Authors: Emmanuel Riggi–Carrolo, Thomas Dubot, Claire Sarrat, and Judicaël Bedouet

Reviewers: Ester Roosenbrand and Junzi Sun

Editor: Junzi Sun

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2023.7209

2. Review - round 1

2.1 Reviewer 1

Thank you for your nice paper, it presents a methodical approach to a complex problem.

1. Sentence in line 26 is confusing. Bracket end is missing and 'if several contrail formation models exist' should probably be 'As several contrail…'. Also what indications or literature do you have to suggest it is necessary to 'reduce their uncertainties' (so , citation needed).
2. I assume Huge in line 94 should be Hough?
3. To a first time reader, it is unclear why Ash RGB images are of such interest for contrail detection, what features of ash clouds are applicable to contrails and how does SO2 help in detection of contrails? Mention this reasoning in your introduction or section 2.2.
4. You use 'False colour image' and 'Ash RGB image' intermittently in the paper. I suggest sticking to one name in the text and in the figures.
5. In Figure 2 the axis font in the right subfigure is far too small, it needs to be made much bigger, so it is legible. Additionally, I question the use of displaying only the trajectories in the Hough transform subfigure and not the contrails as well. What I can only assume to be the ICAO24 numbers are illegible, and don't provide me as a reader relevant information. Perhaps this figure should be remade, but only showing 10 or 20 minutes' worth of data, this way the flight trajectories and contrails and both be shown. Limiting the geographical scope would also help. Finally, it would be nice to mention which location we are seeing in the caption (southern/Baja California).
6. For Figure 4 the text in the plot is too small, either make the font a legible size or consider leaving out some of the text. Do you need the titles of the plots, for example to illustrate the flow of the steps?
7. Line 183: 'A threshold of 70 is arbitrarily chosen'. I assume some other values were tried as well, in a trial-and-error type way? If so, mention this in the text.
8. Figure 6 could also use some editing. Perhaps don't include the data in the red box at the top of the figure, so it becomes more legible. Besides this, the colours of the boxes are arbitrary, right? They are not used elsewhere in the paper to identify the individual contrails?

9. In line 189, you write 'In our example, we observe that many flights have potentially formed 3 contrails.' Since you do not mean that multiple flights created three contrails each, I would rewrite this sentence for clarity. 'In our example, the three contrails could have been formed by multiple flights' for example.
10. In line 194, you state that 'we have information about the points that define the ends of the contrail'. How did you arrive at this information, how do you determine which end of the contrail is the start, and which is the tail? Maybe it was explained, but this reader missed it.
11. In Figure 7, perhaps only show the aircraft with teal circles, as this would make the figure more clear. Refer in the caption that the remaining or kept flights are still visible in Figure 8.
12. Figure 8, I like this figure, it is clear. However, I feel that the wind intensity indicating by the wind barb's isn't relayed well, since you have to look very carefully if the barb has 3 or 2.5 lines on it to indicate a difference. Also, the average reader will need to look up what these magnitudes mean in terms of m/s. I suggest just using arrows, indicating direction and the scale indicating the wind speed. In the caption, you can say that the wind speeds vary from 10 to 15 m/s.
13. In line 206, I think you mean 'In fact, even though 442 aircraft flew' instead of 'even if'.
14. Why was this location and date chosen? If a date (during the COVID-19 pandemic, for example) or location with a less busy airspace was chosen, would you be able to assign a contrail to a single flight? Would those types of flights be sufficient for building of your database?
15. Figure 8. I can image that some of the flights (pink, light blue, light green and orange) seen in the bottom right corner are flying at (very) similar cruise altitudes. If so, it would not necessarily help with narrowing down the number of possible contrail forming flights, but it would help narrowing down the altitude of contrail formation. Given the temporal/vertical resolution of your data, could that be an additional filter be sufficient to build a reliable database?

Additionally, how does this point interact with your assumption that 'all flights and contrails

are considered to be at the same altitude on the image. So, all aircraft trajectories are superimposed on the same horizontal plane.'?

1. Related to the previous point, in line 226, you claim that you overcome challenges related to altitude in your approach, but don't you superimpose all trajectories on the same horizontal plane? Maybe I misunderstood this assumption, but I don't understand how you can make this claim.
2. In the conclusion in line 235, scenarios over the Atlantic Ocean are mentioned. This is the first time such a scenario is mentioned. Why would the limited number of trajectories be disadvantageous for your methodology (See comment 14)? Elaborate more on what happened in the Atlantic Ocean scenario. Or do you mean that OpenSky coverage is not sufficient and you are missing trajectories? If so, make that clearer in the text.
3. If the improvements mentioned in the last bit of the conclusion are implemented, do you think you would be able to improve the method in such a way that in a majority of cases be able to assign just a single flight to a single contrail? What kind of data or other things would you need to improve the assignments further?

2.2 Reviewer 2

The paper details the process of identifying contrails from GOES satellites and correlating them with specific flights based on OpenSky trajectories. The dynamics of the wind are taken into account to enhance identification accuracy. It employs Hough transforms for both contrails and flight trajectories and proposes a method for correlating them in Hough space. The paper clearly describes the entire process, from data gathering and methods to analysis.

Following are my comments for improving the paper:

1. Please remove unnecessary abbreviations in Line 13

2. Line 44: The acronym 'CDA' is used only once; it may not be necessary to abbreviate it.
3. Line 87: Could you please specify which channels are used to generate the differences?
4. Table 1: 'Percent correct' and 'probability of detection' are metrics not commonly used in machine learning. Could you elaborate on how they are calculated?
5. Figure 2: The text is too small to read. Consider removing the labels associated with the dots in Hough space. Additionally, the red dots are not easily visible and should be plotted above the blue dots for clarity.
6. Equation 1 and 2: The multiplication symbol should be '
cdot' in LaTeX for clarity.
7. Figure 4: The text within the figure is quite small; increasing its size would enhance readability.
8. Line 165 and Figure 4 mention a 'similarity filter,' but this concept is not thoroughly explained in the paper. Please provide additional details.
9. At the same places, a 'decision filter' is mentioned. It appears to be referred to as 'final filter' in the paper (Line 191). It would be beneficial to unify these terms if they are the same; otherwise, further explanation of the 'decision filter' is needed.
10. Line 191: The phrase 'exclude flights headed towards a contrail' is ambiguous. Does it refer to flights not aligned parallel to the contrail, or does it imply something else?
11. Please also improve the readability of Figures 6 and 7, as the font is very small in the figures. Figure 8 has both very small and very large fonts; please unify them for better visualization.
12. Line 208: You state that 13 out of 442 flights were identified for the 4 recognized contrails. Is there a method to determine the probability of association between flights and each contrail?
13. Finally, the reproducibility sections (including data and code) are missing from the paper.
14. Conclusion: The study appears to be based on a limited data set from a single GOES image. Have additional tests been conducted with more examples to assess the efficacy of your approach?

2.3 Reviewer 3

Review text

3. Response - round 1

3.1 Response to reviewer 1

1. Sentence in line 26 is confusing. Bracket end is missing and 'if several contrail formation models exist' should probably be 'As several contrail…'. Also what indications or literature do you have to suggest it is necessary to 'reduce their uncertainties' (so, citation needed). 2. I assume Huge in line 94 should be Hough?

Response

These sentences have been corrected.

3. To a first time reader, it is unclear why Ash RGB images are of such interest for contrail detection, what features of ash clouds are applicable to contrails and how does SO2 help in detection of contrails? Mention this reasoning in your introduction or section 2.2

Response

The sentence in the paper has been completed to give some precision about the ash RGB, as we use a pre-processing based on the combination of three longwave GOES-16 brightness temperature to create an "ash" false color image (previously used to detect volcanic ashes and SO2 gas in satellite images).

The purpose of this pre-processing transformation is to assist the algorithm in its learning as it allows a better identification of condensation trails in satellite images.

4. You use 'False color image' and 'Ash RGB image' intermittently in the paper. I suggest sticking to one name in the text and in the figures.

Response

The text has been corrected referring to false color image instead of ash RGB.

5. In Figure 2 the axis font in the right subfigure is far too small, it needs to be made much bigger, so it is legible. Additionally, I question the use of displaying only the trajectories in the Hough transform subfigure and not the contrails as well. What I can only assumeto be the ICAO24 numbers are illegible, and don't provide me as a reader relevant information. Perhaps this figure should be remade, but only showing 10 or 20 minutes' worth of data, this way the flight trajectories and contrails and both be shown. Limiting the geographical scope would also help. Finally, it would be nice to mention which location we are seeing in the caption (southern/Baja California)

Response

The figure is redone according to the instructions and the caption is improved.

6. For Figure 4 the text in the plot is too small, either make the font a legible size or consider leaving out some of the text. Do you need the titles of the plots, for example to illustrate the flow of the steps?

Response

We assume that the titles of the plots help the reader to understand the flow of the chart and the successive transformations that are done. Nevertheless, the Figure is now improved with better titles.

7. Line 183: 'A threshold of 70 is arbitrarily chosen'. I assume some other values were tried as well, in a trial-and-error type way? If so, mention this in the text.

Response

The sentence is reformulated for a better understanding of our choices.

8. Figure 6 could also use some editing. Perhaps don't include the data in the red box at the top of the figure, so it becomes more legible.

Response

The figure is redone.

9. In line 189, you write 'In our example, we observe that many flights have potentially formed 3 contrails.' Since you do not mean that multiple flights created three contrails each, I would rewrite this sentence for clarity. 'In our example, the three contrails could have been formed by multiple flights' for example.

Response

In this example, we can identify 4 domains for 4 contrails in the area. This allows to determine which aircraft are suspect for each contrail formed. Here, three contrails could have been formed by multiple flights.

10. In line 194, you state that 'we have information about the points that define the ends of the contrail'. How did you arrive at this information, how do you determine which end of the contrail is the start, and which is the tail? Maybe it was explained, but this reader missed it.

Response

The sentence has been reformulated to make it clearer.

11. In Figure 7, perhaps only show the aircraft with teal circles, as this would make the figure clearer. Refer in the caption that the remaining or kept flights are still visible in Figure 8.

Response

In order to make the Figure 7 more clear, the caption is improved.

12. Figure 8, I like this figure, it is clear. However, I feel that the wind intensity indicating by the wind barbs isn't relayed well, since you have to look very carefully if the barb has 3 or 2.5 lines on it to indicate a difference. Also, the average reader will need to look up what these magnitudes mean in terms of m/s. I suggest just using arrows, indicating direction and the scale indicating the wind speed. In the caption, you can say that the wind speeds vary from 10 to 15 m/s.

Response

The wind barbs are used in order to give a precise information to the reader about the wind strength. That is the reason why the authors would prefer to keep these bards instead of the arrow.

13. In line 206, I think you mean 'In fact, even though 442 aircraft flew' instead of 'even if'.

Response

The sentence is corrected.

14. Why was this location and date chosen? If a date (during the COVID-19 pandemic, for example) or location with a less busy airspace was chosen, would you be able to assign a contrail to a single flight? Would those types of flights be sufficient for building of your database?

Response

This study case was chosen in order to illustrate the complexity of such an exercise. Of course, a date and location with less flights should be able to resolve, i.e. to attribute a contrail to one or more trajectory, but it wouldn't have been sufficient to demonstrate the need of these successive algorithms to associate aircrafts and contrails.

15. Figure 8. I can image that some of the flights (pink, light blue, light green and orange) seen in the bottom right corner are flying at (very) similar cruise altitudes. If so, it would not necessarily help

with narrowing down the number of possible contrail forming flights, but it would help narrowing down the altitude of contrail formation. Given the temporal/vertical resolution of your data, could that be an additional filter be sufficient to build a reliable database?

Response

The altitude of contrail formation is the main remaining issue for association with an aircraft. Unfortunately, this information is not accessible, even though we developed the Schmidt Appleman filter to estimate the possible level of flight favorable to contrail formation (using the ERA5 met data). The altitude information could be available using LIDAR data for example and would of of a great help to build a reliable database.

Additionally, how does this point interact with your assumption that 'all flights and contrails are considered to be at the same altitude on the image. So, all aircraft trajectories are superimposed on the same horizontal plane.'?

1. Related to the previous point, in line 226, you claim that you overcome challenges related to altitude in your approach, but don't you superimpose all trajectories on the same horizontal plane? Maybe I misunderstood this assumption, but I don't understand how you can make this claim.

Response

Your observation is valid. The Hough transform is not directly related to altitude considerations. We will modify the text accordingly to eliminate this dependency.

2. In the conclusion in line 235, scenarios over the Atlantic Ocean are mentioned. This is the first time such a scenario is mentioned. Why would the limited number of trajectories be disadvantageous for your methodology (See comment 14)? Elaborate more on what happened in the Atlantic Ocean scenario. Or do you mean that OpenSky coverage is not sufficient and you are missing trajectories? If so, make that clearer in the text.

Response

The developed methodology is based on ADS-B data provided by OSN. As the oceanic signal is not recovered by the terrestrial antenna, the OSN ADS-B data miss aircraft trajectories over the Atlantic Ocean. That is the reason why we see a limitation to apply this method to such a case. A sentence has been added in the text to answer that comment.

3. If the improvements mentioned in the last bit of the conclusion are implemented, do you think you would be able to improve the method in such a way that in a majority of cases be able to assign just a single flight to a single contrail? What kind of data or other things would you need to improve the assignments further?

Response

The improvements described in the conclusion should be of a great help to associate contrails with one or several aircraft, as induced-cirrus are often produced by successive trajectories crossing ISS Regions. As stated before, using satellite images the main issue remains the uncertainties of contrail altitude.

3.2 Response to reviewer 2

Please remove unnecessary abbreviations in Line 13

Line 44: The acronym 'CDA' is used only once; it may not be necessary to abbreviate it.

Response

The corresponding sentences have been corrected.

Line 87: Could you please specify which channels are used to generate the differences?

Response

The details are now given in the text: the transformation concerns the RGB channels following the Equation 1

Table 1: 'Percent correct' and 'probability of detection' are metrics not commonly used in machine learning. Could you elaborate on how they are calculated?

Response

These metrics are coming from Hoofman et al., 2023 and used here in order to compare their score to our results. The Percent Correct (PC) is A+D/(A+B+C+D) and the Probability Of Detection (POD) is A/(A+C) where A is a True Positive ("hit"), B a False Positive ("false alarm"), C a False Negative ("miss") and D a True Negative ("correct negative"). These details are added in the text.

Figure 2: The text is too small to read. Consider removing the labels associated with the dots in Hough space. Additionally, the red dots are not easily visible and should be plotted above the blue dots for clarity.

Response

The Figure is redone.

Equation 1 and 2: The multiplication symbol should be '
cdot' in LaTeX for clarity.

Response

The correction is done.

Figure 4: The text within the figure is quite small; increasing its size would enhance readability.

Response

The Figure is redone.

Line 165 and Figure 4 mention a 'similarity filter,' but this concept is not thoroughly explained in the paper. Please provide additional details.

Response

In order to clarify the paper, the ‘similarity filter has been renamed in ‘Advection filter’ as it refers to the contrail movement under given wind conditions. A specific paragraph starting with “The advection filter” gives specific details.

It appears to be referred to as ’final filter’ in the paper (Line 191). It would be beneficial to unify these terms if they are the same; otherwise, further explanation of the ’decision filter’ is needed.

Response

The filter names have been homogenized in text.

Line 191: The phrase ’exclude flights headed towards a contrail’ is ambiguous. Does it refer to flights not aligned parallel to the contrail, or does it imply something else?

Response

The sentence meant to say that the aircraft shouldn’t precede the contrail to make it. The sentence has been rewritten to make it clearer.

Please also improve the readability of Figures 6 and 7, as the font is very small in the figures. Figure 8 has both very small and very large fonts; please unify them for better visualization.

Response

The Figure is redone.

Line 208: You state that 13 out of 442 flights were identified for the 4 recognized contrails. Is there a method to determine the probability of association between flights and each contrail?

Response

We assume you are referring to the probability of associating remaining final candidate flights with contrails. In this context, we haven’t delved into methods for classifying these remaining candidates. In this paper, our aim was to demonstrate the likely difficulty of associating a contrail with a single flight using such satellite images, especially considering our strong assumptions such as the time frame considered for traffic history. It would indeed be interesting to explore a method to assign a probability to each final candidate flight.

Finally, the reproducibility sections (including data and code) are missing from the paper.

Response

The authors will provide data and code with the reviewed paper.

Conclusion: The study appears to be based on a limited data set from a single GOES image. Have additional tests been conducted with more examples to assess the efficacy of your approach?

Response

Yes, text modified before the conclusion to indicate we have been exploring the methodology on several GOES-16 images, but kept this one, as illustrative and complex enough.

4. Review - round 2

The reviewers confirm the comments have been addressed.