Authors' Response to Reviews of

The ice supersaturation biases limiting contrail modelling are structured around extratropical depressions

Oliver G. A. Driver, Edward Gryspeerdt, Marc Stettler
Atmospheric Chemistry and Physics, `egusphere-2025-2737`

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

AR: We would like to thank the reviewers and our editor for their thoughtful and generous comments on our manuscript.

Please find detailed responses to individual comments below. As well as these changes, we've also also made some minor changes to wording.

A final set of technical corrections has been made to correct the editor's technical points noted following the revised submission. These are:

RC: ***Figure 1: it would help to have geographic coordinates (besides the coast lines)***

AR: Latitude/longitude grids has been added.

RC: ***For the composite figures, you might consider adding x and y labels (and ticks).***

AR: These have been added. Note that x and y axis labels are 'distance from the system centre' as these are only aligned with the cardinal directions at the origin, as has been described in the manuscript.

RC: ***line 18: typo "albeit while"***

AR: Fixed this and a second occurrence later in the document.

> **`line 18 (manuscript)`**
> Contrails (ice clouds that form from the exhaust of aircraft) cause a significant fraction of the warming radiative forcing due to aviation (Lee et al., 2021), ~~albeit while potentially having~~ although they potentially have significantly less impact on climate (Bickel et al., 2025).

> **`line 315 (manuscript)`**
> The structures observed in the composites are aligned with those seen in the high-altitude composite from the reanalysis model output in terms of their geography and sign, ~~albeit while suffering from~~ though they suffer far more noise due to the sparse measurements.

RC: ***line 509: "including with"***

> **`line 509`**
> Contrails have been repeatedly observed in subsaturated air, ~~including~~ with a peak in occurrence rates at the 90 % RHi value considered for ISSR diagnosis in this case (Li et al., 2023).

Subsequent line numbers refer to locations in the 'diff' file.

1. RC1: 'Comment on egusphere-2025-2737', Anonymous Referee #1, 25 Jul 2025

Overview:

RC: ***The authors of this paper had a clever idea, that is indeed a novel contribution to the complex problems of contrail prediction. The idea is that contrails and ice supersaturation often appear in certain dynamical regimes, and rarely or not at all in others. Early papers with such findings, like Kästner et al. (1999) and Immler et al. (2008) might be quoted here.***

AR: We have included reference to Kästner et al. (1999) as part of the introduction.

> **Lines 89–94**
> The dynamical regimes around these systems enable ISSRs to be formed and sustained, including after the formation of contrail ice. This has been demonstrated, for example, by Kästner et al. (1999), where the warm, slowly lifted air in warm conveyors, and the cold, rapidly moving air in jet streams created conditions that enhanced the rate at which persistent contrails were observed. Model biases (where an output quantity is systematically wrong) can ~~also reflect the structure of the atmosphere, when~~ stem from a failure to capture these underling mechanisms. Therefore, the ISSR biases ought to differ in different parts of atmospheric structures, where they are caused by different processes~~undergone in different parts of the system~~.

AR: We have also included Immler et al. (2008), which treats high pressure systems, as part of the discussion, providing some direction to future study.

> **Lines 494–499**
> . This similarity between the composite and counterfactual suggests that the ISSRs omitted from the composite are not found in similar regions to the lows, instead they differ in their seasonality or typical location within the study region. As a result, the lows detected here are not the only cause of ISSRs. Future work may consider the biases in other parts of the system with their own underlying ISSR mechanisms, for example, in extratropical high pressure systems, where contrails often form within ubiquitous very optically thin cirrus (Immler et al., 2008).

RC: ***The authors consider ice supersaturation in relation to lows over the atlantic and use ERA5 data to make a composite of thousands of lows, considering in particular the humidity structure and the winds around the composite low. In order to enhance the S/N ratio, another clever idea is to make a second composite with synoptic cases that happen one year after each composited low, thus they transform a random selection of weather patterns into a kind of featureless average weather. To extract the modifications that are due to the low, the two composites are substracted and a signal appears.***

RC: ***The appearance of ISSRs is then inspected in the difference plots (the signal), and the relation to air traffic patterns and observational data from IAGOS studied.***

RC: ***IAGOS data later serve as the truth to derive score values of ISSR prediction by comparing the ISSRs in the composites to the IAGOS data mapped to the composites. The result confirms that ISSR prediction is currently a big challenge. There is indication that lowering the RHi threshold to, e.g. 90% increases recall substantially, at only a minor loss to precision.***

RC: ***The paper closes with the statement that forecasting of ISSRs must be improved, and that regarding the structure of the underlying weather could help.***

RC: ***I agree fully with this statement.***

RC: ***The paper is well written and I could not spot any typographical error. I am impressed.***

RC: ***The paper is clearly worth of publication, in particular because of its novel ideas. But I think, a little more work should be invested to make a couple of issues more clear.***

AR: We thank the reviewer for reading our manuscript and their kind and constructive comments. We have made changes in response to their comments and have included the details and relevant extracts from the diff below.

Section 1:

RC: ***Abstract: The message transported in the abstract is not so clear to me. I understand that, if composites of depressions in ERA5 lead to a clear distinction in ISSR occurrence between the dry intrusion on the one hand and the warm conveyor on the other hand. But then, to my surprise, the authors state that the same ERA5 data have problems to capture the very infrequent occurrence of ISSRs in the dry intrusion. This sounds like a contradiction.***

AR: The abstract has been slightly reworded to simplify the message. We clarify that the structure is real and shown in both model and observation. This is said to show both that this tells us something about where the ISSRs are, but also that the model somehow captures the underlying causes. We hope this removes any scepticism about the prior point generated by the latter point.

Lines 1–20

Contrails are ice clouds ~~occasionally~~ formed along aircraft flight tracks, responsible for much of aviation's climate warming impact. ~~Contrails persist in ice-supersaturated~~ Ice-supersaturated regions (ISSRs) provide conditions where contrail ice crystals can persist, but meteorological models often mispredict their occurrence, limiting contrail modelling. This deficiency is often treated by applying local humidity corrections. However, model performance is also affected by synoptic conditions (such as extratropical depressions).

Here, composites of ERA5 reanalysis ~~model~~ data around North Atlantic extratropical depressions enable a link between their structure and ISSR modelling. ISSRs are ~~highly~~ structured by these systems: at flight levels, ~~the ERA5 ISSR rate is 91 % less~~ ISSRs occur less frequently in the dry intrusion – ~~in~~ descending upper-tropospheric air – than above warm conveyors – where air is lifted. ~~The contrast also occurs in composites of in~~ Both ERA5 reanalysis and in situ aircraft observations ~~, showing~~ show this contrast, demonstrating that the model reproduces the fundamental relationship. ~~However, performance in modelling individual ISSRs also differs across the structures. Compared to the warm conveyor belt, the infrequent ISSRs in the dry intrusion are less well captured by ERA5, with a~~ Individual-ISSR modelling performance (quantified using interpretable metrics) is also structured. Of the rare ISSRs diagnosed in the location associated with the dry intrusion, fewer are confirmed by in situ observations (20–25 % ~~drop in diagnosing confidence (precision) and~~ precision drop compared to the warm conveyor) and fewer of those observed were diagnosed (13–19 % ~~drop in comprehensiveness (recall~~ recall drop). Scaling humidity beyond ~~ISSR occurrence rate corrections is able to dramatically increase the recall with a small precision costand high specificity~~the occurrence rate bias dramatically increases the recall at low precision cost, demonstrating the potential value of scaling approaches designed with different intentions. However, the failure of scaling to improve precision, or the performance in the dry intrusion, implies that there is a need to account for the synoptic weather situation and structure in order to improve ISSR forecasts in support of mitigating aviation's climate impact.

RC: ***L 32: the meaning of the half sentence "insofar as the contrail population..." is not clear to me.***

AR: Changed to "based on the understood distribution of contrail properties from a modelled population"

> **Lines 36–38**
> Satellite monitoring and validation are still in the early stages and can detect at most 46 % of contrails, ~~insofar as the contrail population is understood based on contrail modelling~~ based on the understood distribution of contrail properties from a modelled population (Driver et al., 2025).

RC: ***L 47: what do you mean with "coupled to the wider meteorology"?***

AR: Clarified the intended meaning of those contrail models that are "not coupled" to be "offline" contrail modelling studies (when they are run separately from the meteorological model)"

> **Lines 52–54**
> ~~It is~~ These sources are commonly used for "offline" contrail modelling studies (when they are ~~not coupled to the wider meteorology~~ run separately from the meteorological model) (Teoh et al., 2022; Engberg et al., 2025).

RC: ***LL 59-61: Regarding the detrimental effect of saturation adjustment, see also Sperber and Gierens (2024).***

AR: Citations have been added in the introduction and the conclusions.

> **Lines 68–70**
> Modelling approaches designed to avoid such a parameterisation are uncommon, but are able to capture decaying ISSRs shortly after cirrus formation, and ISSRs in slowly-rising warm air that continue to be sustained even as clouds form (Sperber and Gierens, 2023).

> **Lines 97–99**
> For example, the saturation adjustment is applied in models when clouds form ~~(Thompson et al., 2024)~~ (Thompson et al., 2024; Sperber and Gierens, 2023), so a dry bias resulting from overactivity in this parameterisation would be structured.

> **Lines 576–579**
> Past work has associated these underestimations with over-active saturation adjustments when natural clouds form (Thompson et al., 2024)~~, this should be addressed~~. Nonetheless, ISSR precision remains low, including after applying humidity scaling; work where the saturation adjustment has been avoided shows ice-supersaturation behaviour that is more realistic (Sperber and Gierens, 2023). This should be further explored to improve ISSR predictive performance.

Section 2:

RC: ***Fig. 1: I find the colour bar not very helpful. If it is meant to clearly indicate RHi>100%, a clear change in colour would be good. Furthermore, regions with quite dry air have a similar blue to much moister regions.***

AR: We have adjusted the colour bar, including discretisation, which has made a significant improvement. This has meant that we have filled the regions identified as ISSRs, which improves the figure's clarity. Slight adjustments have been made to the figure caption as required.

Figure 1 caption
Charts showing the ERA5 reanalysis model meteorology around a sample of low pressure systems in the North Atlantic. ERA5 wind orientation and speed are indicated by arrows, and grey contours indicate the mean sea level pressure in hPa. ~~Coloured contours~~ Levels at 90 %, 97.79 % and 100 % RHi are ISSRs identified when applying successively less stringent thresholds~~to the RHi~~, equivalent to simple uniform humidity scalings. The ERA5 data is charted at two altitudes: a–c are representative of the mid-troposphere (around 470 hPa, 6 km altitude) and d–g the ~~upper-troposphere~~ upper troposphere (around 260 hPa, 11 km altitude). Empty regions (e.g. the upper right hand corner of panels a, b, d, and e) are 10 degrees beyond the region of interest and have been neglected from this analysis.

RC: ***L 108: "upper atmosphere" sounds to me as levels close to TOA.***

AR: Changed to "upper troposphere"

Lines 128–129
...to emphasising the 'dry intrusion' of downwelling air from the upper ~~atmosphere~~ troposphere (Browning, 1997).

Section 3:

RC: ***Why is this section called "Methodology" instead of "Methods"?***

AR: Changed to "Methods"

RC: ***Section 3.1, first paragraph: I have some questions about the compositing method. First, as a Low has a time-scale of, say, 2 days, it may appear in several subsequent 6-hour periods. How did you treat this? Is a low taken only once for the composite (e.g. in order to avoid multiple counting of the same system), or is every 6-hour slot taken as independent of each other?***

AR: Repeated detections of the same system are all included (as specified, "all the low pressure systems detected were kept, including where these are repeated detections of the same system"). We treat each 6 hour slot independently, though inherent to this work is that the systems share a similar structure, so do not see multiple inclusions of a system as a flaw, given the below point, that we accept the fact that we have systems at different phases of development. Our assumption is then that no one anomalous system overwhelms the composite. A comment on this has been added to the text (in the diff alongside the changes for the point below).

RC: ***Second, as the 6-hour snapshots find the lows in different phases of their development, doesn't this smooth the features that you intend to find?***

AR: We agree that this is a trade-off, leading to smoothing, but enabling the inclusion of more data (especially important for the spatially and temporally sparse IAGOS observations).

AR: We have added some commentary to acknowledge the previous two comments more explicitly.

Lines 192–195
To achieve this, all the low pressure systems detected were kept, including where these are repeated detections of the same system ~~.~~ — we assume that no individual anomalous system overwhelms the composites produced. This approach is inherently a trade-off, no attempt has been made to isolate a particular phase of system development, instead prioritising the inclusion of as many systems as

possible. Producing composites with meaningful structure acts as the test of validity for this approach.

RC: ***Third, the actual pressure on a model level is determined by the surface pressure. As this varies across the system, the surface pressure structure is somehow imprinted on the upper model levels, that is, there is a certain pressure variation as well across the system. How large is it and how do you think it impacts your analyses?***

AR: Thank you for raising this, it is an interesting question, though ultimately we believe the impact of these variations to be small, particularly at the altitudes relevant for cruise aircraft. We believe that the observation of composite structures which are independent from the pressure structure validate this approach. Nonetheless, we below explore the size of this effect:
The pressure level conversion is

$$\frac{\mathrm{ph}}{\mathrm{hPa}} = a + b\frac{sp}{\mathrm{hPa}},$$

giving the 'half-level pressure' (ph) as a function of the 'surface pressure' (sp) and tabulated parameters a and b. Hence, within a model level, $b = \mathrm{d}p/\mathrm{dsh}$. At the 'high level', where most analysis is performed, model level 80, $b = 0.06$. For typical variations on the order of 30 hPa, there are around 2 hPa variations across the system. At the mid level, the sensitivity is greater $b = 0.3$, so variation of around 10 hPa might be typical. These are approximately 60 m and 150 m vertical variations at high and mid levels respectively (assuming a hydrostatic atmosphere). These variations are small with respect to the vertical levels on which compositing has been performed, but are appreciable. We have included this in the manuscript's discussion.

Lines 480–485
The use of ERA5 model levels means there is some dependence of the surface pressure that may be imprinted on the system. Based on the model level definition (ECMWF, 2023) and approximating the troposphere to be hydrostatic, surface pressure variations of approximately 30 hPa would lead to pressure variations less than 12 hPa and altitude variations less than 170 m (at the lowest level, model level 70, around 600 hPa). At typical flight level, as mapped in Fig 1(d–f), and forming the main part of the analysis, these variations are 1.8 hPa, or around 60 m (two flight levels). These variations are deemed small, so have been neglected during this analysis.

AR: There's also an impact due to the geometry of the underlying surface. Greenland has a particular impact, due to its altitude and position, featuring in many composites. Greenland rises to around 1500 m above sea level (the variation on model level 0). In a single system, this leads to variations of 80 hPa at the 'mid-level' and 20 hPa at the 'high level', these are larger in terms of altitude (on the order of 1 km within a model level). However, the position of Greenland is not held fixed (unlike the pressure variations associated with the low)—which means the impact is not responsible for the structure in the composites—and these variations are shared in both the composite and counterfactual—which means that they will be removed when the 'signal' is isolated. We are further reassured by the fact that none of the composites show any features which seem to be artefacts of this effect. Therefore, we make no changes to account for this, aside from also mentioning this in the manuscript.

Lines 485–489
The model levels are also impacted by surface geometry, for example, many of the systems occur in proximity to Greenland. This can lead to significantly larger pressure variations within a model level, however, any impact of this effect is removed when the 'signal' is calculated, because these features are shared in both the low-centred regions and the counterfactual composite. We are further reassured by

the fact that none of the composites examined suffer noticeable artefacts due to this effect.

RC: ***Further questions to methods: ISSRs are frequently connected to cirrus clouds. How is this treated in your method? As you don't mention cloud clearing, I suppose clouds are part of the ISSRs shown and analysed later. If so, are they considered or at least mentioned in the analysis?***

AR: We do not explicitly consider any data on cirrus cloud occurrence, or 'natural clouds' present in the model in general. It is important that cirrus-related ISSRs are captured in models as well as those without cirrus, and those regions which are humid but may not be cold enough for some relevant processes, such as homogeneous freezing. All have relevance for contrail modelling. Elsewhere in the analysis, we use the 'mid-level' ISSRs as a proxy for underlying cloud in order to understand the signals in the contrail forcing in different parts of the system. We have emphasised the lack of any distinction being made between between ISSRs with or without other cloud in the later analysis of the relevant composites.

Lines 246–260

~~The signal~~ This includes all ISSRs diagnosed against the threshold, including where cloud is found and where the temperature may not be low enough for a contrail to form. The signal (panel b) then highlights the change in this rate due to the presence of the low – the structure given to the system because a low pressure system occurs here.

This is in broad agreement with past work, such as Catto et al. (2010). Specifically, a dry intrusion behind the cyclone at the high level, and the comparatively-humid region above the warm conveyor belt are visible. A clear comma-shaped structure is seen at the mid-level altitude, which corresponds to the location of the cloud head (despite this being a higher level than might be typical for those features). Humid air here has been lofted in the warm conveyor belt. ~~In the upper-level~~

To be explicit, the mid-level ISSRs should not be conflated with regions at risk of contrail occurrence, and temperatures at this level are mostly too warm for contrails to form. Instead, the mid-level composite is presented to provide context for the vertical structure, the location of lower-level clouds, and for comparison with past compositing work. The conditions for contrails, including temperature and air traffic are considered in the analysis below, where CoCiP-modelled contrails are used as a single proxy including all these factors. Although ISSR occurrence remains the useful metric in the context of this article, which characterises the ISSR occurrence rate bias, conserved quantities offer a useful point of comparison to other atmospheric physics work and provides an intuition as to the flows in the system, therefore, we provide composites of the water vapour mass mixing ratio in Appendix A.

RC: ***L 145: I suggest to delete "empirical". The M&K paper has a lot of theory involved to derive the formula, but, as far as I remember, no new measurements, so the word "empirical" is not appropriate here.***

AR: The word "empirical" has been deleted.

Line 176

...according to the Murphy and Koop (2005) ~~empirical~~ formulae for the saturation vapour pressure ...

RC: ***LL 164 ff: Do I understand that correctly: You take the 2009 air-traffic for each day at 5, 11, 17, 23 UTC and the corresponding contrails over the next hour?***

AR: This is not the case. We take the existent contrails each day at 6, 12, 18, 24 UTC, i.e. those modelled to persist up to this time from flights up to 12 hours (the maximum CoCiP lifetime). Instantaneous populations

of contrails are shown.
Because flight distance is inherently an integrated quantity, we need to integrate the routes flown over some time period to estimate the flight activity. We composite "flight distance in the previous hour". Ultimately the choice of 1 hour is made somewhat arbitrarily, but is of the order of magnitude expected for the lifetime for many persistent contrails, while remaining specific to the current synoptic conditions, and interfaces well with the available dataset. The flight distance is then interpreted as a “rate”, (flight density per hour), an estimate of the flight activity which leads to the set of contrails. The data description, particularly the term “corresponding” was misleading in the original manuscript. We have revised it to clarify.

> **Lines 198–203**
> Other data was composited over subsets of the time of the ERA5 composites. ~~Air traffic (Teoh et al., 2024b) over the hour previous to the composite timestep was added for timesteps in 2019, with~~ Flight activity (the number of flown kilometres, per square kilometre horizontal area, over the hour previous to the composite timestep; Teoh e was added to the composites for 2019. Flight data from 2020 and 2021 ~~data~~ was excluded to ensure normal behaviour of air traffic. ~~Corresponding contrails modelled over 2019~~ Contrails modelled by Teoh et al. (2024a) using CoCiP (the Contrail Cirrus Prediction model Schumann, 2012), using the pycontrails Python library’s ‘trajectory’ implementation (Shapiro et al., 2023), were also included, based on all the flight activity in 2019 and ERA5 meteorological data with the humidity scaled.

RC: ***Section 3.2, L 187: There are "non-local" approaches that I would like to bring to your attention, for instance by Duda and Minnis (2009a) and Duda and Minnis (2009b) and recently by Wang et al. (2025).***

AR: Duda and Minnis (2009a) and Duda and Minnis (2009b) focus on the prediction of persistent contrail occurrence by the application of statistical models, rather than applying a classical thermodynamic approach. While we appreciate that this work accounts for many more variables than an approach where the humidity was simply scaled, as far as we can tell the models produced are still what we mean by local: they account only for the meteorology at the location where the model is applied, and don’t account for the dynamical state of the system, beyond the local velocities. Nonetheless, we have added reference to this work, as this is an interesting approach which goes beyond a humidity scaling (similar to the Wolf et al. (2025) and Teoh et al. (2024a) approaches).
Wang et al. (2025) is an interesting and novel approach. Seeing such an approach succeed when applied to contrail modelling would be transformative, and would validate the need to account for dynamical factors. We have included discussion of this both within the introduction and the current section of our manuscript.

> **Lines 71–82**
> The frequency of ISSRs in model data, and the distribution of humidity within them, can be adjusted by scaling the RHi to best match the population statistics of observed ISSRs (discussed further in Section 3.2, Teoh et al., 2022; Teoh et al., 2024a; Wolf et al., 2025) . This approach may be enough to model a population of contrails for climate impact assessment or for statistical analysis of contrail properties. However, accurate ISSR occurrence rates without capturing their spatial structure are insufficient to model individual contrails more accurately~~(such as part of a contrail avoidance programme; Molloy et al., 2022)~~ , which is required for model validation or to isolate the impact of individual aircraft (as may be required for a contrail mitigation p
> .
>
> One similarity between all of ~~these~~ the aforementioned ‘humidity corrections’ is that they are applied locally, based only on the conditions at the point of their application~~.~~, this neglects the influence that dynamics has on the processes leading to ISSRs. The need to account for this structure is demonstrated by the relative success of the novel humidity correction approach taken by (Wang et al., 2025), where

humidity is corrected based on both the current and past conditions both local to and in the immediate proximity to each point.

Lines 223-227
~~Ultimately, these corrections~~ Other approaches, such as the statistical modelling of persistent contrails as in the work of (Duda and Minnis, 2009a; Duda and Minnis, 2009b), have taken into account an even broader array of characteristics. But ultimately, the aforementioned approaches are all local, so don't fully account for the weather's structure. Only an approach that accounts for the dynamics, as in the recent work by Wang et al. (2025), has the potential to account for the causes of the bias.

Section 4:

RC: ***Figure 3: I come back to my question whether you have applied cloud-clearing. I am a bit surprised on the quite strong ISSR frequency at mid-level north of the centre of the low. Is ist possible that we simply see water or mixed-phase clouds there? Inside a water cloud, the water vapour is clearly supersaturated, but this is usually not considered in studies of ice supersaturation.***

AR: We have not applied cloud clearing. We agree that it is likely that the strong mid-level ISSR frequency is almost entirely water or mixed-phase clouds – indeed, they occur at temperatures likely too warm for the homogeneous freezing, are likely to be too warm for the Schmidt–Appleman criteria to be met, and are therefore not altogether relevant for contrail predictability. However, the inclusion of the mid-level composite is not to serve these aims, but to add additional context for the rest of the analysis. Specifically, it offers a point of comparison to past compositing work, such as Catto et al. (2010) (where 500 hPa RH and 400 hPa system-relative winds were composited). The relevant changes (identical to those presented above) are repeated here for convenience:

Lines 246-260
~~The signal~~ This includes all ISSRs diagnosed against the threshold, including where cloud is found and where the temperature may not be low enough for a contrail to form. The signal (panel b) then highlights the change in this rate due to the presence of the low – the structure given to the system because a low pressure system occurs here.

This is in broad agreement with past work, such as Catto et al. (2010). Specifically, a dry intrusion behind the cyclone at the high level, and the comparatively-humid region above the warm conveyor belt are visible. A clear comma-shaped structure is seen at the mid-level altitude, which corresponds to the location of the cloud head (despite this being a higher level than might be typical for those features). Humid air here has been lofted in the warm conveyor belt. ~~In the upper-level~~

To be explicit, the mid-level ISSRs should not be conflated with regions at risk of contrail occurrence, and temperatures at this level are mostly too warm for contrails to form. Instead, the mid-level composite is presented to provide context for the vertical structure, the location of lower-level clouds, and for comparison with past compositing work. The conditions for contrails, including temperature and air traffic are considered in the analysis below, where CoCiP-modelled contrails are used as a single proxy including all these factors. Although ISSR occurrence remains the useful metric in the context of this article, which characterises the ISSR occurrence rate bias, conserved quantities offer a useful point of comparison to other atmospheric physics work and provides an intuition as to the flows in the system, therefore, we provide composites of the water vapour mass mixing ratio in Appendix A.

AR: The fact that this humidity structure reflects the expected structure of underlying cloudiness means it is useful as a proxy to contextualise the difference between contrail occurrence and contrail forcing in the comparison of Figs. 5 and 6. Given all these factors, we believe that the argument to include the mid-level composites of Fig. 3a–c, without applying cloud clearing, is strong. It could be said that ISSR occurrence is not a relevant quantity with which to examine humid structures at this level, given that the conditions for ice clouds are lacking, but is an appropriate compromise against complexity. We've added composites of the mass mixing ratio of water in the appendix:

Appendix A (Lines 610–619)

The structure in humidity that is induced by low pressure systems is presented in the body of the manuscript using ISSRs (Fig. 3). While ISSR occurrence is the natural quantity with which to explore the consequences for contrail science, ice processes are less relevant at lower levels, including the 'mid-level', around 470 hPa, as is occasionally examined in this work. Indeed, temperatures at this altitude are often too high for homogeneous freezing of water droplets, or for the Schmidt–Appleman contrail formation criteria (Schumann, 1996) to be met.

In Fig. A1, we present the structure induced in the water vapour mixing ratio. This quantity is conserved within an air parcel, aside from any condensation or evaporation. It exhibits similar structure to that of Fig. 3, though the latitudinal dependence is different. The differences between this and the ISSR composites occur because of structure in temperature, which in turn affects the relevant saturation vapour pressure of ice.

AR: We note that the CoCiP composites of Fig. 5 enables the interpretation of the structure of contrail formation, and account for the conditions that are actually relevant for contrail formation and persistence, while Fig. 6 enables the interpretation of the relative impact of typical contrails, including accounting for underlying natural clouds.

RC: ***At the higher level, I am suprised to have the signal circling around the low and that nothing is to be seen circling into the anticyclonal direction. However, this is what I would expect from past work on ISSRs in relation to atmospheric dynamics (e.g. Spichtinger et al., 2005; Wilhelm et al., 2022).***

AR: The mentioned turning into the anticyclonal direction is indeed seen in the horizontal velocity composites at both mid and high level (the arrows in Fig. 3a, d), as in the signal at the high level (Fig. 3e). Furthermore, despite anticyclonic flow not always being evident from the wind arrows, the nature of the motion is anticyclonic. This is evident in composites of the vorticity, which we have included in an appendix. This accounts for any correlation between ISSR occurrence and vorticity. More qualitatively, the ISSR occurrence rate in the region where ISSRs would be 'circling into the anticyclonal direction' do occur in the specific case studies shown (e.g. Fig. 1d–f), and the signal is likely to exhibit smoothing resulting from the inclusion of systems at different phases of development (as has already been acknowledged in the manuscript). Ultimately, these are composites and the general occurrence of circling humidity should not be interpreted as a humid region, rather as a part of the system wherein the structures leading to modelled ISSRs can occur.

Lines 267–272

It is counterintuitive that the upper-level composite shows a sweeping arc of humidity in a manner that appears cyclonic, in a part of the system where a transition to anticyclonic flow might be expected, coincident with the outflow. In fact, the nature of the flow in this part of the system is indeed anticyclonic, as demonstrated in composites of the potential vorticity, which are presented in Appendix B. This transition to anticyclonic flow is shown in outflow regions in the individual case studies (Fig. 1d–f).

Ultimately, the composite winds do suffer some smoothing as a result of the trade-off made to include systems at multiple phases of their evolution, in order to consider as much data as possible.

Appendix B (Lines 620-629)
The direction of composited winds (overlaid on Figs 3–7) can give the impression that the wind around these systems is largely cyclonic, including in the regions of enhanced upper level humidity. This would conflict with the interpretation that these ISSRs occur in the outflow from the warm conveyor belt, where winds tend to turn anticyclonically. This is not the case, as is demonstrated by Fig. B1, wherein the vertical component of relative vorticity (i.e. based on horizontal variations of the horizontal ERA5 winds) is composited. The same sweeping region associated with enhanced ISSR occurrence rate is associated with negative relative vorticity (so anticyclonic winds).

This sweeping region is a result of the trade-offs accepted in order to include as much data as possible (outlined in Section 3.1). Therefore, it is important that these composites are not interpreted as a hypothetical state of the weather, but as a lens through which to understand the structure of these systems.

RC: ***Figure 4 and sect. 4.2.1 Have you tried to separate flights in east- and westbound flights? As you explain, the signal southeast of the low is probably mainly due to eastbound flights. Why is there no signal from westbound flights?***

AR: Our data does not allow us to isolate eastbound and westbound flights, we only have gridded flight distance available currently. To investigate this effect, we have expanded the analysis of flight distance, specifically by subsetting based on the longitude of flight activity. This addresses this point and the point below, so is expanded on in the response below.

RC: ***Figure 4 and 5: The air traffic and contrail maxima at the western and eastern edges of these figures are explained by traffic close to America and Europe. Now, in a composite there is no real America and Europe because these continents get smeared out by the composition. Your red box in figure 2 extends far into America such that I wonder why there is a structure at all in the counterfactuals of figs 4 and 5. Perhaps an indication of where your 4000km * 4000km box is on average (longitude) would help to clarify this issue.***

AR: While there is blurring as described in the comment, an effect due to the geography of the lows is expected, and forms the basis for the inclusion of counterfactuals (as alluded to in the comment). We don't believe that specifying a single "average location" helps elucidate this, rather, this would imply that the composites might be considered a hypothetical situation in a specific place, which they should not. Instead, we have expanded on the analysis as an appendix, by including plots of the contribution to flight distance based on three longitude bands (West of 65W, between 65 and 30W, and East of 30W). These plots document the origin of the features in the composites.

Lines 302-305
The structure present in the composites of air traffic density suggest different behaviour depending on the longitudes, strong enough to cause peaks in the total composite. The counterfactual methodology has enabled the general signal due to a low to be isolated independently, but it is useful to break down the air traffic by longitude in order to provide supporting evidence. This is presented in Appendix C.

Appendix C (Lines 630-644)

The composite structure of flight distance drawn out by the composites in the body of the article (Fig. 4) shows peaks in air traffic on the east and west edge of composites. This is inferred to be structure resulting from the distribution of routes, specifically, air traffic associated with termini in Europe and North America. In order to break down this signal, composites have been produced of the density of air traffic, and the structure around the low for three longitudinal bands: east of 30W, between 65W and 30W, and west of 65W. This draws out a different set of behaviours. Corresponding composites are shown in Fig. C1

European air traffic is concentrated south east of the location of low pressure systems; around half of it redistributes closer to the lows in order to take advantage of winds assisting eastbound motion. Over the North Atlantic, air traffic is generally seen to concentrate more in the west than the east, consistent with the behaviour of air traffic, which approaches a more diverse range of latitudes in Europe than when approaching North America; redistribution here is weak compared to the typical air traffic density, but there is a preference for travelling near the lows (where perhaps routes with constructive winds can be chosen). Finally, in the western band, the structure is less clear. A strong peak occurs on the eastern edge of the composite (where a limited number of systems will contribute to the signal); there seems to be a slight preference to fly further south, clear of any disruption from the presence of a low pressure system.

RC: ***Figure 6 and sect. 4.2.2: To my opinion, the discussion is incomplete here. It is difficult to believe that a contrail located in the dry intrusion is something like a big hit. Furthermore, the signal in fig. 6 may mainly show noise, since panels a and b show similar forcings. I suggest to use the background (i.e. panel c) to estimate a standard deviation of the signal and to compare that with the size of the signal. You may also try t-tests or similar (if a t-test is inappropriate) to test the Null-hypothesis that there is no statistical difference between 6a and 6c.***

AR: In preparing our response to this comment, we uncovered two errors in the analysis:
The lengths of contrails were being incorrectly averaged over the vertical levels, but was compared to flight distance data had been summed. This meant that the contrail lengths presented were incorrect in magnitude (Fig. 5), which led to errors in per-length forcing. The relevant figures have been revised (Fig. 5 and Fig. 6).
We have also addressed the underlying cause of artefacts which appeared on the top edge of CoCiP and IAGOS composites, which has enabled us to revise the colorbar. These artefacts occurred because of the way that contrail segments and IAGOS data were included in the composite, which previously included a filter being applied to latitude based on the the North East and North West corners of the region around the low (which are not the Northenmost points in the region, given the projection). This, combined with the low air traffic density at the relevant latitudes, led to noise near the Northern edge of the composite, and an inappropriate colorbar.
We have recomposited the relevant data and revised subsequent figures in the manuscript.
Finally, we considered the relevant component of the forcing. Diurnal and latitudinal effects are key drivers of variability in net forcing, due to the varying incident short wave – this behaviour, shared between the composite and counterfactual case. This shared variability adds noise, which means that the composite and counterfactual are harder to differentiate, particularly for the sensitivity test applied below. Therefore, we have opted to reframe this result in a presentation of the LW forcing, rather than the net.

Lines 316-320

Fig. 6 shows the spatial distribution of the warming long wave component of instantaneous radiative forcing per unit length of persistent contrail that is modelled at a given location around a low. We present the long wave forcing, rather than the net forcing, because variability in incident solar radiation introduces a large component of variability shared between the counterfactual and composite, making the underlying distribution of values difficult to differentiate.

AR: After making these changes, it was clear that there is a meaningful, structured signal due to the presence of the low, and in particular, that there is reduced impact over the regions which are humid at high level, and systematically less impact in the part of the system where one might expect to see underlying cloud.
We believe that the presence of structure in the signal, with a physical explanation, form the more important basis for motivating the interpretation of composites, and are conscious that the construction of significance tests can lead to a reduction, rather than a gain in transparency. Therefore, we have opted to omit significance tests from the body of the manuscript. Nonetheless, we have performed Mann–Whitney U tests, and have produced plots of the signals overlaid with the regions of significance though present these as part of an appendix and referenced as part of the discussion.

Lines 475-479
The emergence of structure in composites of many systems is sufficient to justify the interpretation of signals. However, there are cases in the composites produced where data is limited, and noise makes the structures less clear. In order to provide additional confidence, we have performed Mann–Whitney U tests of the significance of the difference between composite and counterfactuals. We find that all the structures that are interpreted are statistically significant over extended parts of the system at the 5 % significance level. We present the particular regions of significance in Appendix D.

Appendix D (Lines 645-656)
In the article, we produce low-centred composites, and counterfactual data sharing the geography and seasonality of lows, which are agnostic to the presence of low pressure systems. Then, we present and interpret the difference between these two composites as a "signal due to the presence of the low". The structures observed within these signal plots, which extend across regions, are inherently evidence that the comparison, and interpreted signal, is meaningful.

Nonetheless, the opportunity exists to test the statistical significance of difference between the distributions of composited values at corresponding points in the 'composite' and 'counterfactual' populations. We perform Mann–Whitney U tests (Mann and Whitney, 1947) at each location in the signal, testing the difference between the composite and counterfactual distribution at the 5 % significance level – extended regions of significance reflect independently-validated significance over the structure and therefore provide particularly strong evidence in support of differences due to the presence of the low. Regions of significance are presented for all the signals presented in the body of the article in Fig. D1. All the structures interpreted in this analysis are found to be significant.

AR: In the case of the revised Fig. 6, several clusters of significance are found. Given that each (coarsened) grid box is tested independently, clusters offer particularly strong support to the significance of the results. The most clear takeaway in this case is the reduced impact of contrails which occur over the cloud head, lending support to the comparison made in the manuscript with the mid-level humidity signal. We acknowledge that the region of increased impact in the likely cloud-free region is not wholly identified to be statistically significant so unambiguously. Therefore it is an overstatement to say that contrails here *are* big-hitting, rather that they are a return to the norm, where the big-hitting contrails *can* occur. We have rephrased the presentation of these results to reflect that.

Lines 320-327
Contrails tend to have a reduced impact in a region centred north east of the low (Fig. 6b). This region of reduced impact is coincident with the 'cloud head' (~~c. f.~~ cf. the mid-level humidity signal, a proxy for underlying clouds, Fig. 3b). Natural cirrus is likely to form here, which means that contrails are unable to contribute additional radiative effect. Conversely, ~~contrails are comparatively 'big hitting' on average (with greater forcing~~ there is a return to the typical contrail forcing behaviour (similar to the counterfactual) behind the low pressure system, partially co-located with the dry intrusion. ~~Although contrails here are infrequent (based on the meteorological data currently used), the ones that form here can be particularly impactful, so being able to diagnose the infrequent ISSRs in this region is crucial~~It is in these parts of the system where contrails have the opportunity to exert stronger forcing when there is not a competing natural cloud radiative effect.

AR: We've also reframed the statement in the conclusions and made it less strong of an assertion.

Lines 568-571
Even though contrail occurrence is enhanced above the warm conveyor, the contrails that exist here are systematically less strongly forcing, due to coincident natural cloud. Therefore, the contrails which have larger radiative impact may occur in less well constrained parts of the systems (Fig. 6).

RC: ***Sect. 4.3: Please indicate whether the IAGOS flights are vertically summed up or whether only flight paths at about 260 hPa are selected.***

AR: The IAGOS measurements at all levels are included. The composite reflects an average humidity at any levels where the plane flies within a grid box. In much the same way as the CoCiP composites, our expectations are set based on the 260 hPa level because this is typical for the planes making the measurements. We've clarified this in the manuscript, as well as being explicit in each of the captions.

Lines 329-335
ISSRs detected by IAGOS in situ ISSR measurements are composited in Fig. 7. To maximise the number of measurements, we include IAGOS observations taken at all levels. It remains appropriate to compare against the 'high' altitude composite of ERA5, because this approximately corresponds to typical flight altitude (Fig. 3d–f). The structures observed in the composites are aligned with those seen in the high-altitude composite from the reanalysis model output in terms of their geography and sign~~(Fig 3d–f)~~, albeit while suffering from far more noise due to the sparse measurements. The same sweeping region with a greater tendency for ice-supersaturation, associated with the warm conveyor, is seen. A prominent dry intrusion is also seen, as it was in the ERA5 composite.

Section 5:

RC: ***Eq. 4 and adjacent text: The use of specificity (or recall on non-ISSR cases) is not convincing. It has a similar flaw as accuracy has. Assume that I predict always non-ISSR and assume that TN is about 0.9 (this number does not actually matter here). As I always predict non-ISSRs, FP=0, since I never predict a positive outcome. That is, TN/(TN+FP)=1. In Figure 8f) the results are >0.9, so the real prediction of non-ISSRs is not far away from the grotesque prediction of always non-ISSR. I don't believe that this makes sense.***

AR: It is true that an examination of specificity alone would be incomplete, but we believe that the argument for including it is very strong, specifically for its physical interpretability that accuracy lacks.

The discussed flaw is also a weakness of recall (which could deem a model that placed ISSRs everywhere flawless if it was considered as a sole metric–even though this is an even less realistic model than assuming subsaturation everywhere).
Like accuracy, these metrics are affected by population imbalance. This has been acknowledged where the manuscript discusses the fact that the direction of change to both specificity and recall when using different thresholds and between the different regions are a result of this (section 5.2, beginning "Gains should be expected based on the definition of recall..." – we have clarified and strengthened this statement – and "In contrast to recall, specificity can only decrease").

Lines 440-442
Gains should be expected ~~based on~~ as a consequence of the definition of recall (eq. 3) as a lower threshold on ERA5 data ~~means~~ changes the population imbalance: FN regions become TP (as well as TN regions becoming FP).

AR: However, the nature of the specificity value is directly interpretable. As we state in the manuscript, it quantifies the models ability to "leave confidently dry regions unencumbered". Its high value reflects not only that the true underlying population is unbalanced, but the model accurately produces an unbalanced population too, and its small changes with the scaling show that the near-doubling of the recall does not cause a false alarm problem.
An important distinction is that the crude "always non-ISSR" model given would give neither perfect nor imperfect accuracy – it would only show the imbalance. However, when precision, recall, and specificity are all considered in the case of the crude model, its nature becomes clear: yes, 100 % specificity, but alongside 0 % recall (precision is undefined, which would be another indicator of a flawed model). Similarly, an all-ISSR predictor would give perfect recall; the precision would reflect the imbalance of the popultions (this quantity's own flaw) and a 0 % precision. For ISSRs specifically, the weakness of this model would be clear from the low precision (because population has more non-ISSR cases), but for applications with a different imbalance, only examining the specificity might reveal the crude nature of the model.
The limitation that the specificity and recall are influenced by the changing population imbalanes is already noted in the discussion of Section 5.2. We have also added the below summary of these limitations to Section 5.1.

Lines 400-411
Finally, it is important ~~that~~ to retain the ability to identify *unsaturated* regions~~is not significantly impacted by applying scaling~~. Any changes made to capture more ISSRs should be weighed against the risk of incorrectly flagging large subsaturated regions as hazards. Precision quantifies the impact that these FP diagnoses have on the diagnosed population, but does not give an indication of the burden induced in the regions which are truly subsaturated. This point was raised by Thompson et al. (2024) who addressed this using the Matthews correlation coefficient (MCC) over the F1 score. ~~The~~ 'However, like the F1 score, the MCC is not directly interpretable, instead providing 'at a glance' comparability. Instead, in this analysis the 'specificity' is ~~now~~ considered, which is equivalent to a recall for diagnosing unsaturated regions,

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \tag{1}$$

~~It indicates the extent to which predicted ISSRs are restricted to where they actually occur~~High specificity values, stable under attempts to improve ISSR diagnosis, are critical to validating that any false alarm problem is avoided. Unlike precision and recall, specificity doesn't speak to the quality of the predicted ISSR population, but instead to the ability to leave confidently-dry regions unencumbered.

RC: ***Section 5.2: This is a very interesting section, but it can be improved. I assume that the discussion is exclusively for the composite weather around this average low. But I am not sure because Fig. 8 also shows bars for the counterfactual. Perhaps, these two artificial situations, one with a composite low and one with an average weather without a low, can be separated more clearly here. A question that arises is whether the artifical average weather is something that has anything to do with any real weather that might occur.***

AR: The presented values are calculated over individual 'measurements', occurring in their actual meteorology (i.e. the comparison is not between the artificial average but in real systems). As in the comparison of composites, the counterfactual is given as a point of comparison, enabling an answer to the hypothetical question 'how much of this signal is due to the presence of the low, rather than the nature of the population?'. This is particularly insightful when looking at recall changes with scaling – the jumps are similar whether the low is there or not, but performance falters in e.g. the dry intrusion in a way not solved by scaling. This interpretation is referred to in the text "The precision and recall are markedly worse for ISSRs encountered in the dry intrusion than in the warm conveyor, and than the composite **and counterfactual** regions as a whole." This comparison is similar to comparisons to the 'all points' estimates, but accounts for the particular geography of the subregions. It is worth highlighting, as is done in the text, that the counterfactual is not 'without a low', but is instead agnostic to the presence of lows – it will include the variability due to lows, but not the structure.
The comment also aludes to the fact that, in many systems the warm conveyor or dry intrusion won't be found in the location it emerges in the smoothed average system. This is reflected by the statement "In fact, some of the ISSRs that are detected may reflect lows with particular morphologies that differ from the composite structure."
We have added some description to clarify this in Section 5.1 – the changes are inluded with those made in response to the next comment.

RC: ***Another issue with this section is the use of relative quantities. I suppose that ISSRs in dry intrusions are close to impossible, that is, if there are any at all they should be very rare. In this case, ratios result from division by quite small numbers and the result is inflated therefore. I suggest to consider the absolute values as well. Perhaps a large relative error in the dry intrusion is negligible against a small error over the WCB.***

AR: We are not entirely clear whether the relative quantities referred to are the metrics themselves (relative to e.g. the number of ISSR occurrences observed or modelled), or are referring to the comparisons between metric values. We hope that the below response aleviates any concerns with either intended meaning.
Comparative relative quanties are comparisions between different precision and recall values, each raw value being given in Fig. 8 – the 'relative quantities' are only saying how much bigger/smaller the bars are to one another. Within each metric, the values occupy a similar size, so no such inflation occurs. We do not only use relative quantites and also state the absolute values of some changes where appropriate, for example, the drop in precision is stated to be three percentage-points (an absolute difference).
We acknowledge that the significance of variations between quantities is not clear without an estimate on the uncertainty on the metrics, which is a result of the number of observations. We have calculated bootstrapped uncertainties, demonstrating that each comparison made is significant.

Lines 378–382

To assess ISSR measurement performance, we calculate binary classification metrics in these two regions, and across the composited data as a whole. We also calculate in the corresponding parts of the counterfactual – as before, this provides a point of comparison that is agnostic to the presence of a low, but shares the particular geography and seasonality of low pressure systems (and the slight differences for the location of the region). We calculate error bars by bootstrapping across 1000 samples (Efron, 1979).

AR: The bootstrapped errors on the metrics also demonstrate that these values are meaningful, and haven't been skewed by the low occurrence rates. It's true that the lower recall in the dry intrusion doesn't amount to many omitted ISSRs, and may be 'negligible' for ISSR accounting. However, examining the humidity metrics is meaningful due to the clues provided as to the cause of model limitations. For example, humidity scaling doesn't enable the dry intrusion performance to "catch up to" the counterfactual, which is evidence for a kind of ISSR modelling error distinct from just the saturation adjustment.

Section 6:

RC: ***LL 396 ff: I am not sure, what you want to say. I understand that 53 and 54% of ISSRs are found in the 50% of the study area where the lows occur. Do you want to say, that accordingly 47 and 46% of ISSRs are in regions whithout a low in the middle? And that therefore lows are not the only cause of ISSRs? It would be nice if that were stated more clearly.***

AR: We have revised this section to be more explicit about the implication of these proportions.

Lines 490–499

These weather systems provide structure to the humidity, but they are not the only source of humidity variation and ISSRs in the region. Around 50 % of the area in the entire study region is included in the composites, with 53% of ISSRs (at 260 hPa) – this means that the other 47 % of ISSRs occur outside the 4000 by 4000 km square regions around lows. The counterfactual includes a similar proportion: 54 % of ISSRs~~, with the comparable ISSR fractions showing the importance of the geography and seasonality of the lowsin overall ISSR occurrence, rather than the presence of a low itself.~~. This similarity between the composite and counterfactual suggests that the ISSRs omitted from the composite are not found in similar regions to the lows, instead they differ in their seasonality or typical location within the study region. As a result, the lows detected here are not the only cause of ISSRs. Future work may consider the biases in other parts of the system with their own underlying ISSR mechanisms, for example, in extratropical high pressure systems, where contrails often form within ubiquitous very optically thin cirrus (Immler et al., 2008).

2. RC2: ‘Comment on egusphere-2025-2737’, Anonymous Referee #2, 28 Aug 2025

General Comments

RC: ***The manuscript reports a novel approach for improving the predictability of ice-supersaturated regions in the upper troposphere. The approach investigates ice-supersaturation in relation to low pressure systems over the North Atlantic and use ERA5 data to create a composite of thousands of lows, with consideration of humidity and wind structures around the composite low. To remove the underlying background signal, a second composite is created for the same position at the low but using synoptic cases that have occurred one year after each composite low. The difference between the two composites produces the modifications due to the low as output signal. The appearance of ISSRs is finally examined in the difference plots, and their relationship with air traffic patterns and IAGOS observational data is studied.***

RC: ***The authors determined score values of ISSR predictability against IAGOS in-situ observations of relative humidity with respect to ice (RHi). The study confirms that ISSR prediction is currently a big challenge. There is indication that lowering the RHi threshold to, e.g. 90% increases recall substantially, at only a minor loss to precision.***

RC: ***The study builds on a novel idea and the presentations of the approach and of the results are well structured and clearly written. The manuscript fits very well into the scope of the journal and deserves publication after a few minor revisions haven been implemented.***

RC: ***The following revisions or adaptions are requested:***

RC: ***1. In the methodology section on the scaling and morphology of ISSRs (lines 189–200), the authors describe that their composite approach of diagnosing ISSRs works best for a threshold of RHi > 90%, and this threshold was used in their entire manuscript. There is no further justification added for applying this threshold.***
However, there are several observation-based justifications for using a threshold lower than 100% RHi to identify regions of cirrus and contrail-cirrus occurrence:
For contrail-cirrus, Li et al. (2023) showed from an analysis of the ML-Cirrus aircraft campaign, that contrail-cirrus occur most frequently at slight subsaturation with highest probability at 90% RHice (Figures 4 and 7). In the same analysis, PDF of RHice of natural cirrus also has its maximum at slight subsaturation (Figure 7). Similar results were reported from an earlier flight campaign by Kübbeler et al. (2011). These arguments may help justifying the application of this lower threshold.
Furthermore, the statement on line 430 that there is no suggestion that using a 90% RHI threshold produces a more-physically representative population of contrails should be adapted. Indeed, there is observational evidence that such a lower threshold is physically justified.

AR: We agree that 100 % RHi is not the only factor determining contrail persistence, and should not be applied as the only test when modelling persistent contrails. Therefore, perhaps these studies provide evidence that the 90 % threshold is suitable for considering cases with a “persistent contrail risk”. We would caution that the ISSR composites in this study should not be interpreted as the regions with a persistent contrail risk, because no test of the Schmidt–Appleman criteria has been made. This particularly affects ISSRs at the “mid-level”, where temperatures are not generally cold enough for contrail formation. This has been discussed in response to a comment above from the other reviewer and clarifications have been made in the text, including the addtion of an appendix containing composites of the water vapour mixing ratio.
However, we believe that these the observation of contrails in observed subsaturation does not serve as direct evidence for diagnosing *actual* ISSRs against the lower threshold, such as for the comparison with IAGOS-observed ISSRs. Nonetheless, we acknowledge the relevance of these results, and have included

citations and a brief description of this argument in the text. We also agree that this is potentially a line of evidence that the diagnosed contrails would be more physically realistic, at least in some regards, so have made adaptations here to say that this work does not suggest that the ISSRs are always contrail-relevant.

> **Lines 529–540**
> ISSRs are not the only important factor in determining contrail persistence. Contrails have been repeatedly observed in subsaturated air, including with a peak in occurrence rates at the 90 % RHi value considered for ISSR diagnosis in this case (Li et al., 2023). To be explicit: the current study is about specifying the meteorological foundation required to validate contrail models. It is then up to the contrail models to ensure that contrail persistence (including the contrails that may be sustained despite being in subsaturated air) are well characterised. Only if more precise meteorology is established, can the ability to model this underlying contrail physics be tested.
>
> Although the extreme scaling of humidity to use a 90 % RHi threshold dramatically increases the recall with little precision cost, ~~there is no suggestion that it produces~~ this work does not mean to suggest that scaling the humidity to this degree would produce a more-physically representative ~~population of contrails~~set of contrail properties. Some applications (e.g. global impact assessments, or study of contrail evolution and properties) are not concerned with reproducing the individual contrails that form in reality, unlike the applications primarily discussed above (i.e. flight impact assessment, mitigation actions). Instead, they seek to find the contrail occurrence rate and characterise contrail property distributions.

RC: ***2. The study presented in the manuscript targets a better predictability of ISSRs in the context of flight rerouting for reducing the climate impact of aviation, see, e.g., abstract and lines 37–42. In the current discussion, uncertainty in predicting ISSR is identified as one major bottleneck for contrail mitigation concepts, which is targeted by the study presented in the manuscript. Another more fundamental criticism of contrail avoidance strategies is based on the question how to assess climate impact 'saved' by contrail avoidance against climate impact generated by additional CO2 emissions. It might be worthwhile mentioning this conflict of objectives, particularly since a recent study has shown that the increase of surface temperature in K per W m-2 of generated effective radiative forcing, the s-called 'climate sensitivity', is smaller for contrail-cirrus than for CO2 (Bickel et al., 2025). Furthermore, the timescales of the climate impacts of contrail-cirrus and CO2 differ significantly (Borella et al., 2024). A short paragraph on these arguments may be added to the introduction section.***

AR: We agree that this is an important uncertainty for mitigation strategies. This particular work is motivated chiefly by the bottleneck on contrail model validation, and does not seek to comment on the effective impact of contrail mitigation. The section referenced only identifies that accurate ISSRs are required to demonstrate contrail avoidance, and we do not believe this is a cogent point here. We have added reference to the climate sensitivity in the opening paragraph.

> **Line 23**
> Contrails (ice clouds that form from the exhaust of aircraft) cause a significant fraction of the warming radiative forcing due to aviation (Lee et al., 2021), albeit while potentially having significantly less impact on climate (Bickel et al., 2025).

AR: Furthermore, the relevance of the different timescales does not have a clear relevance to this point. It motivates the importance of action on aviation's CO2 impact as being more urgent than action on contrails, but doesn't provide evidence that contrail action is unneccesary and doesn't imply any additional uncertainty or bottleneck.

RC: ***3. The connection between warm conveyor belts and ice clouds has been investigated in-depth by Wernli et al. (2016), who presented climatological distributions of in situ and liquid-origin ice clouds over the North Atlantic. This study may make a substantial contribution to the scientific background of Section 2 on low pressure systems.***
Also for section 2, the influence of surface inhomogeneities on the distribution of water vapour an RHi is demonstrated for the regions over eastern North America, the North Atlantic and Europe by Petzold et al. (2020). The annual distribution patterns of water vapour an RHi shown there underpin the selection of an ocean area for the study presented here.

AR: We have included a discussion of (Wernli et al., 2016), particularly to play a role demonstrating distinct air parcel histories depending on the part of these systems.

> **Lines 151–157**
> The different parts of the systems not only have different character, but have distinct histories, which mean that their state is influenced by different model processes. Wernli et al. (2016) examine the history of air parcels to classify different cirrus in reanalysis data, demonstrating the different processes leading to distinct character in different parts of the system. For example, the ice clouds associated with warm conveyor belt outflow tend to have evolved from historic liquid clouds, which is a fundamentally different process to the ice clouds which form in situ, entirely at high altitudes. The diversity of these systems, where contrasting model processes will have led to the current state, make this a compelling situation through which to examine and contrast ISSR modelling performance.

AR: We have also included reference to (Petzold et al., 2020) to aid in justification of the choice of an ocean region.

> **Lines 119–123**
> Additionally, using an ocean region for this study minimises the influence of surface inhomogeneities on weather systems – to enable the structure of weather systems to be resolved as clearly as possible. Petzold et al. (2020) demonstrate enhanced ISSR occurrence in this region compared to land regions, a result of the structure that the ocean region enables to develop. In this study, a slightly different region definition to that of Teoh et al. (2022) and Teoh et al. (2024a) was used: −100–0 °E and 20–70 °N.

SPECIFIC COMMENTS

RC: ***1 | Line 62: Another very successful approach to adjusting humidity distributions to improve ISSR forecasts has recently been published, see Arriolabengoa et al. (2025).***

AR: This is an interesting piece of work, but is ultimately a change to the cloud scheme, rather than a humidity correction, so does not serve to be included precisely where suggested.
The suggested work is also implementing some capacity for ice supersaturation modelling in a model where its representation was completely inhibited – it is therefore difficult to draw a comparison with the IFS model (used for ERA5), which is not limited as strongly as the preexisting ARPEGE scheme. The revised model certainly seems successful in comparisons to this preexisting scheme, but any improvement over IFS is not demonstrated unambiguously, and it seems likely that this revised scheme remains limited in some way. In light of these nuances, and the work still being a preprint, we don't believe the inclusion of this citation to be valuable or necessary in motivating this piece of work, but appreciate it being pointed out and will follow its

development keenly.

RC: ***2 | Figures 4 to 7: The height level of the composite plots should be mentioned in the figures or figure captions. The reference to Figure 3 does not help because there, the composites are shown for two height levels.***

AR: Height levels were not mentioned directly in Figs. 4–7 as these aggregate all data (not limited to a single level). The altitude characterised by the data is instead limited by the vertical distribution of air traffic (specifically IAGOS traffic for Fig. 7). We have specified this explicitly in the figure captions.

RC: ***3 | The adequate reference to the performance characteristics of the IAGOS humidity sensor is Neis et al. (2015).***

AR: We have included this reference in the discussion of the approach taken to compare IAGOS and ERA5 measurements. Having reflected upon the approach informed by this paper our treatment (following (Wolf et al., 2025)) seems sufficient and we make no further changes.

> **Lines 366–369**
> IAGOS measurements are frequent, approximately every 4 s during a flight, yielding an apparent spatial resolution of approximately 1 km. However, limitations of the humidity sensor at the cold temperatures expected (Neis et al., 2015) mean that ISSRs are unlikely to be resolved at this scale (Wolf et al., 2024), furthermore, this is significantly higher-resolution than ERA5.

RC: ***4 | In the conclusions, the authors discuss the approach of balancing the under-representation of ice-supersaturation in ERA5 via an over-active saturation adjustment when natural clouds form. Here, the work presented by Arriolabengoa et al. (2025) might also be mentioned.***

AR: Under-representation of ice supersaturation and overactive saturation adjustment are not ‘balanced’, rather it is likely that one is a mechanism that causes the other. As where its inclusion was suggested above, we don’t believe that this citation adds much value or additional clarity, so haven’t made changes.

RC: ***5 | Since IAGOS data are heavily used for the verification of the method, it might be appropriate adding the IAGOS Data Policy statement to the acknowledgements. This statement is available here https://www.iagos.org/data-policy/.***

AR: Thank you for pointing out this omission, we have included the relevant statement. As the manuscript was developed the use of IAGOS data expanded, and on reflection we agree that the IAGOS data are heavily used, so have also taken appropriate steps to ensure compliance with the data policy more broadly.

MINOR ISSUES:

RC: ***Line 400: There is one ’the’ too much, please remove.***

AR: We have removed this repeated word.

3. Revised Figures

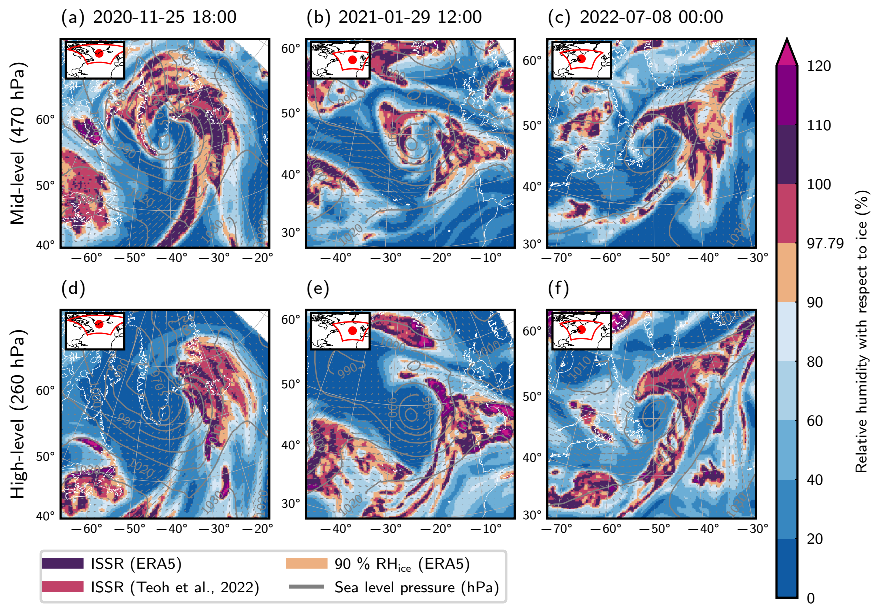

Figure 1: Charts showing the ERA5 reanalysis model meteorology around a sample of low pressure systems in the North Atlantic. ERA5 wind orientation and speed are indicated by arrows, and grey contours indicate the mean sea level pressure in hPa. ~~Coloured contours~~ Levels at 90 %, 97.79 % and 100 % RHi are ISSRs identified when applying successively less stringent thresholds~~to the RHi~~, equivalent to simple uniform humidity scalings. The ERA5 data is charted at two altitudes: a–c are representative of the mid-troposphere (around 470 hPa, 6 km altitude) and d–g the ~~upper-troposphere~~ upper troposphere (around 260 hPa, 11 km altitude). Empty regions (e.g. the upper right hand corner of panels a, b, d, and e) are 10 degrees beyond the region of interest and have been neglected from this analysis.

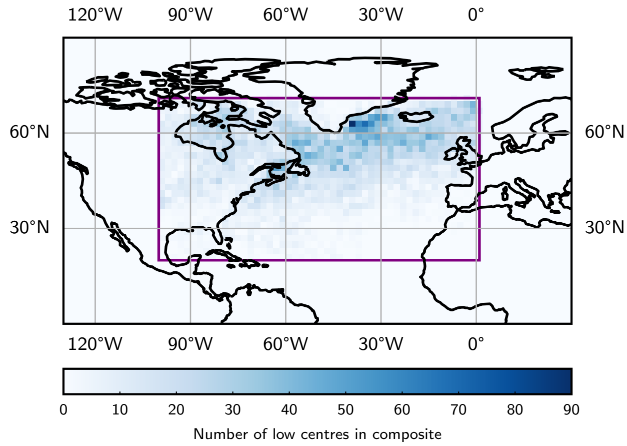

Figure 2: The distribution of low centres included in the composite.

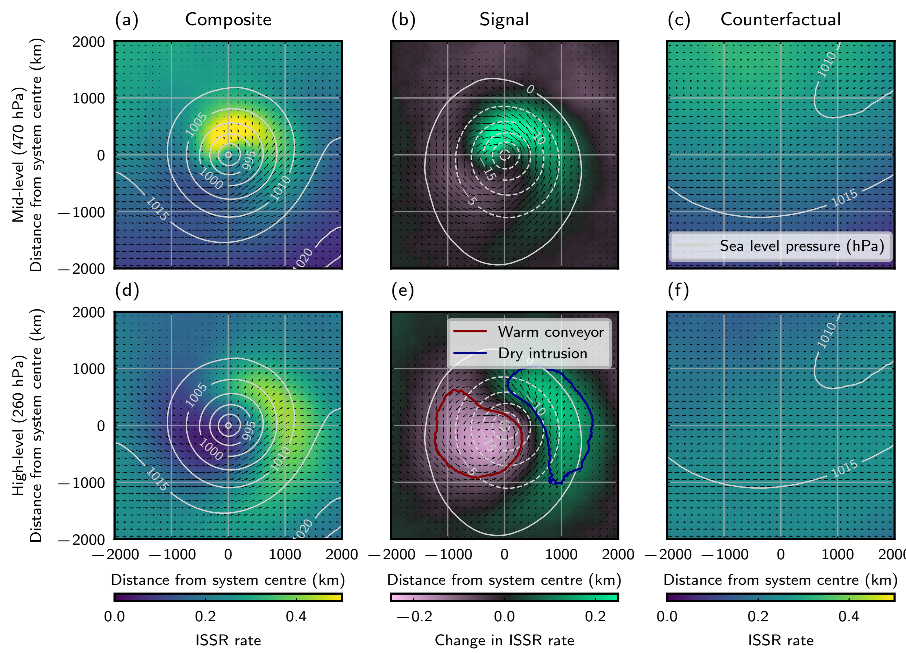

Figure 3: The mid- (a–c) and upper- (d–f) troposphere composite structure of ERA5 ISSR occurrence around locations of North Atlantic low pressure systems. The composites (a,d), show the rate at which ice-supersaturation occurs at locations relative to the centre of a low, while the counterfactual data (c, f), show the same ISSR rate occurring 365 days after a low was detected (i.e. a mean state of the meteorology, sharing the geography and seasonality of low occurrence). The 'signals' (b, e), show the difference between the composite and counterfactual fields. The signal can be interpreted as the impact due to the presence of the low pressure system, or the system-relative quantity. A RHi greater than 90 % has been used to indicate modelled ISSR. The red and blue contours of panel e denote contrasting regions used for analysing ERA5 accuracy in Sect. 4.3.

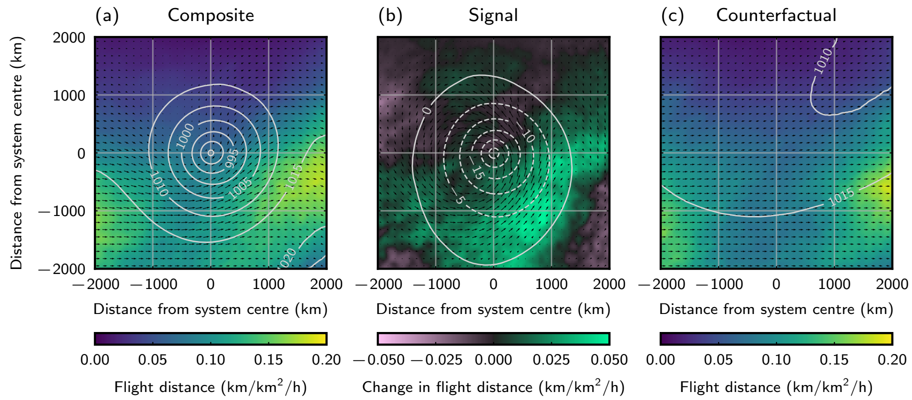

Figure 4: Composite, signal, and counterfactual (as Fig. 3) of distances flown around lows by aircraft in the GAIA dataset per unit area, per hour, based on traffic in the hour preceding the composite time (Teoh et al., 2024b). Data is integrated across the composited vertical levels.

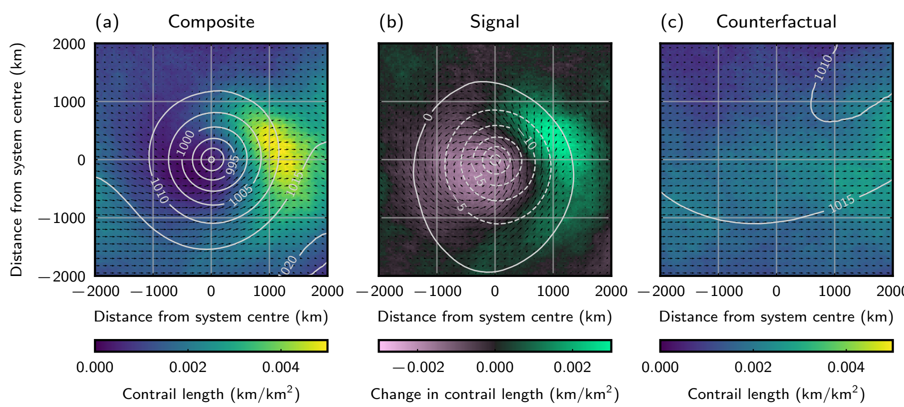

Figure 5: Composite, signal, and counterfactual (as Fig. 3) of the persisting length of contrails existing at the time of the composite per unit area. Data is integrated across the composited vertical levels so reflects contrails which exist at any altitude.

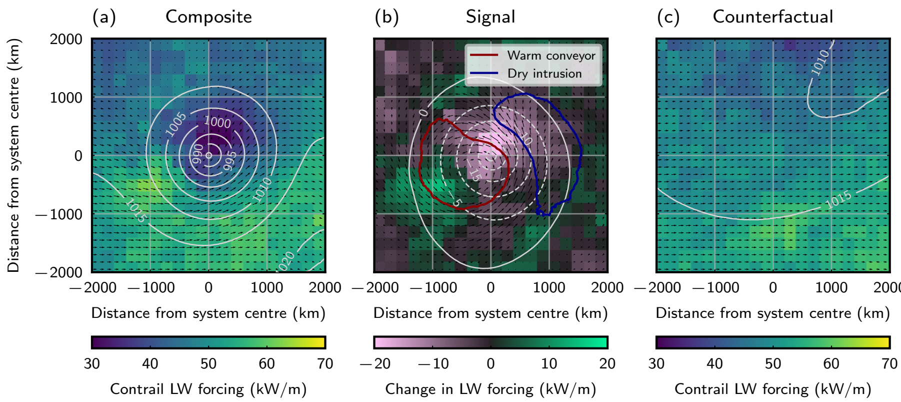

Figure 6: Composite, signal, and counterfactual (as Fig. 3) of the mean estimated instantaneous ~~net~~ long wave radiative forcing of contrails per unit length (i.e. the relative importance of the contrails at this location for climate forcing). Data is integrated across the composited vertical levels so reflects contrails which exist at any altitude.

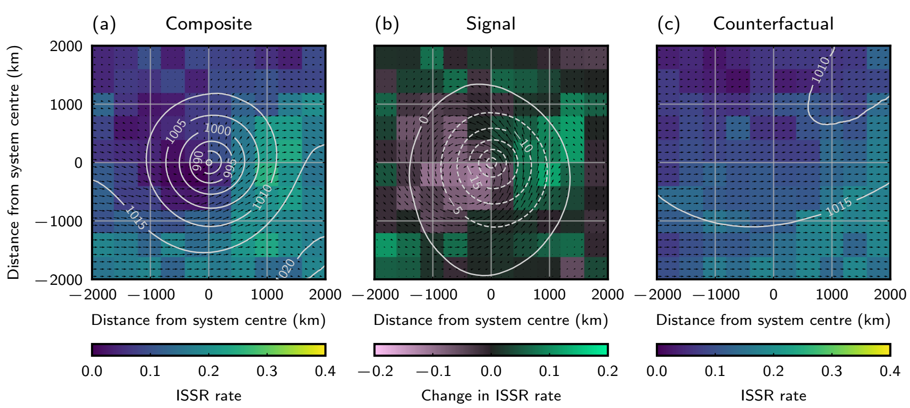

Figure 7: Composite, signal, and counterfactual (as Fig. 3) of the IAGOS-detected ISSR rate. The scales are chosen to match Fig. 3. Data is averaged across the composited vertical levels where IAGOS aircraft fly, so the altitude of detected ISSRs reflects that typical of operational aircraft.

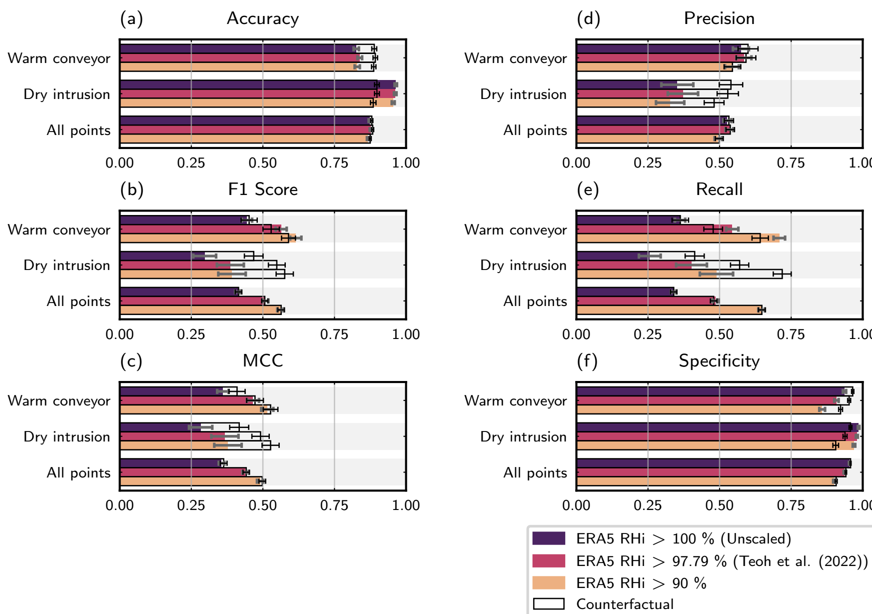

Figure 8: The ability to diagnose ISSRs detected by IAGOS using ERA5 reanalysis meteorology data in different regions around the low, quantified using accuracy, precision, recall, and specificity (a, b, c,and d respectively). Each metric is calculated using three different thresholds to define ISSRs, equivalent to applying simple scalings to the RHi. Error bars result from bootstrapping the metrics across 1000 samples with replacement.

4. New Figures

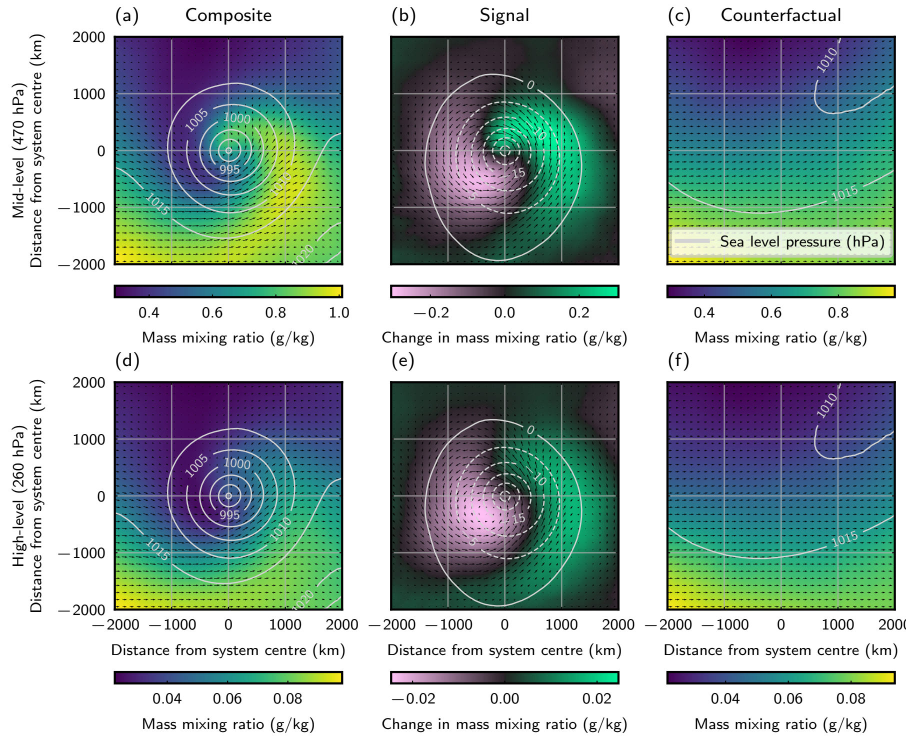

Figure A1: Composite, signal, and counterfactual (as Fig. 3) of the ERA5 water vapour mixing ratio. Data is shown for the two vertical levels presented in Figs 1 and 3.

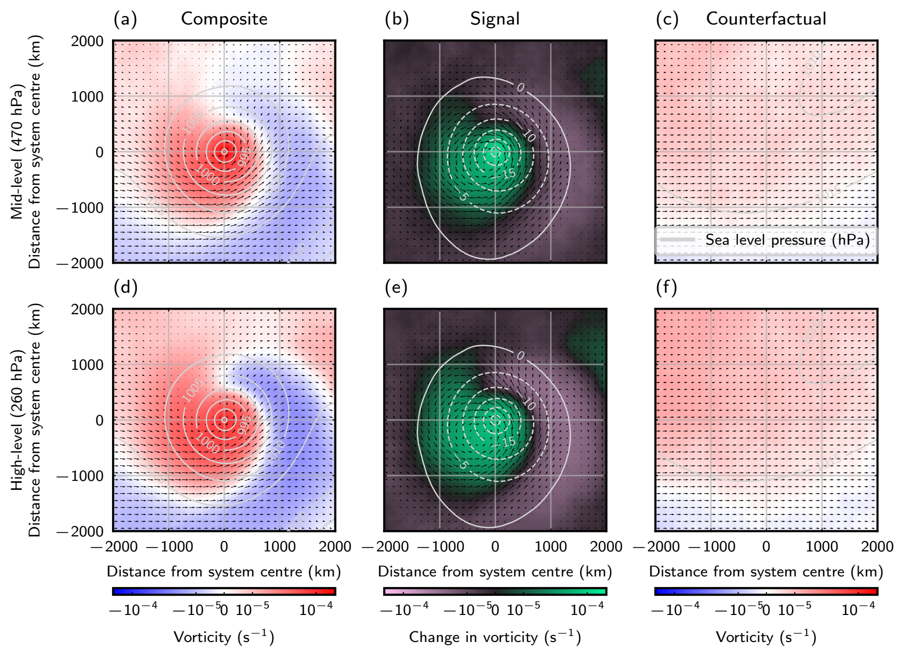

Figure B1: Composite, signal, and counterfactual (as Fig. 3) of the relative vorticity. Data is shown for the two vertical levels presented in Figs 1 and 3.

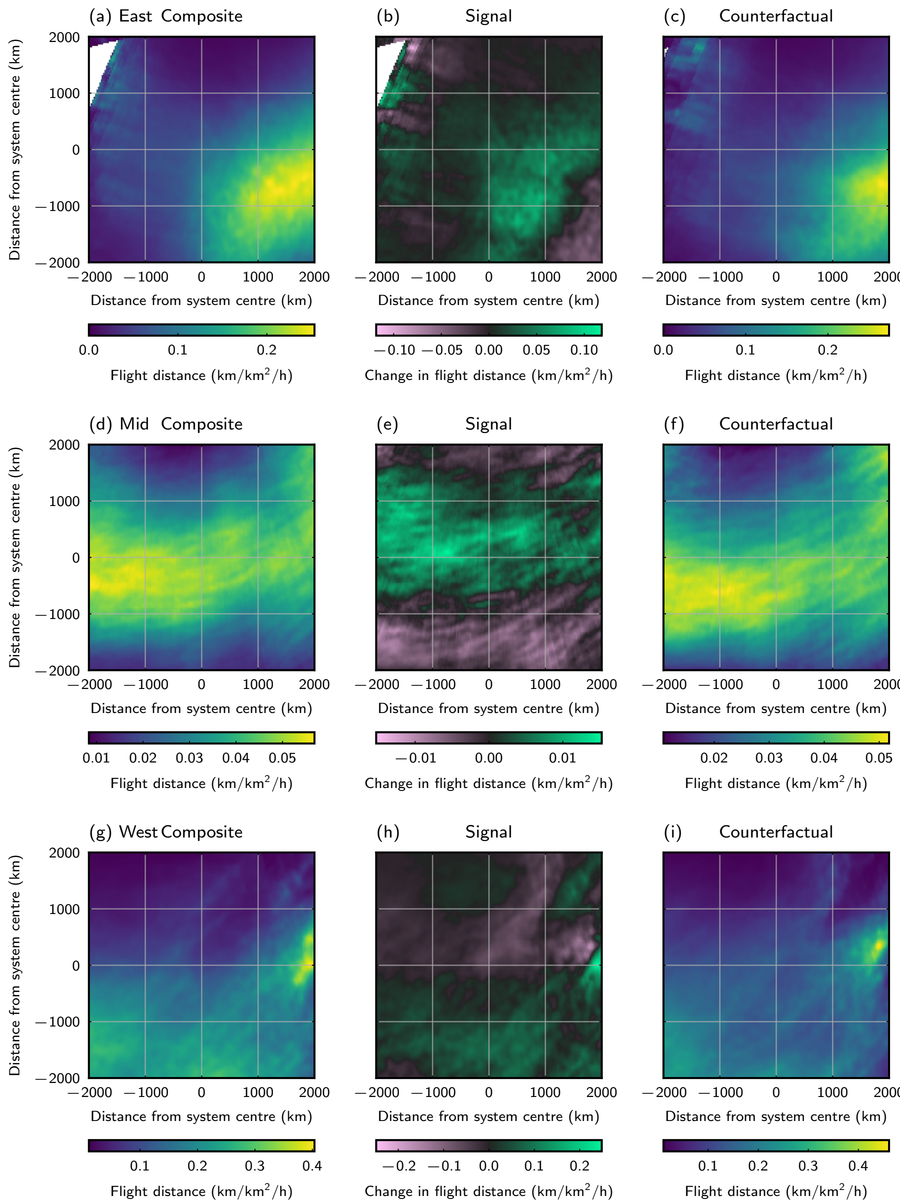

Figure C1: Composite, signal, and counterfactual (as Fig. 3) of flight distance, as in Fig. 4, filtered to only show flights occurring east of 30W (a–c), between 65W and 30W (d–f), and west of 65W (g–i).

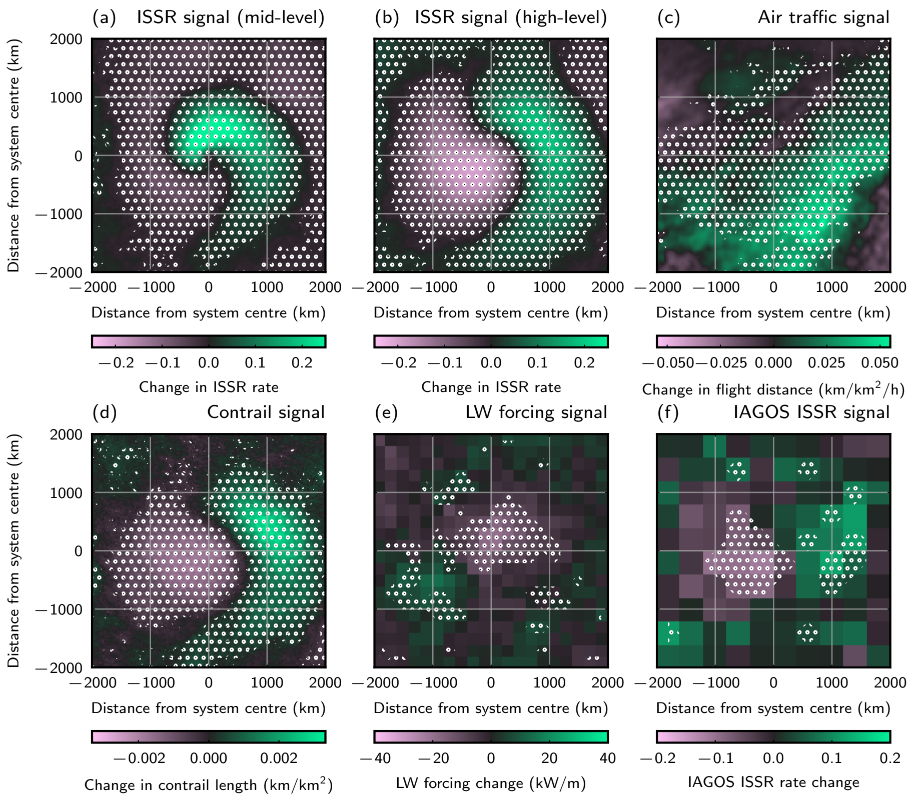

Figure D1: Signal plots as presented in the body of the manuscript, overlaid with dots in regions where the difference between distribution of data in low-centred and counterfactual composites are significant.

References

Arriolabengoa, S. et al. (May 2025). “Modeling and verifying ice supersaturated regions in the ARPEGE model for persistent contrail forecast”. In: *EGUsphere*, pp. 1–34. DOI: 10.5194/egusphere-2025-1499. (Visited on 09/10/2025).

Bickel, M. et al. (Apr. 2025). “Contrail Cirrus Climate Impact: From Radiative Forcing to Surface Temperature Change”. In: *Journal of Climate* 38.8, pp. 1895–1912. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-24-0245.1. (Visited on 06/17/2025).

Borella, A. et al. (Sept. 2024). “The importance of an informed choice of CO_2-equivalence metrics for contrail avoidance”. In: *Atmospheric Chemistry and Physics* 24.16, pp. 9401–9417. ISSN: 1680-7316. DOI: 10.5194/acp-24-9401-2024. (Visited on 09/15/2025).

Browning, K. A. (1997). “The dry intrusion perspective of extra-tropical cyclone development”. In: *Meteorological Applications* 4.4, pp. 317–324. ISSN: 1469-8080. DOI: 10.1017/S1350482797000613. (Visited on 02/05/2025).

Catto, J. L. et al. (Apr. 2010). “Can Climate Models Capture the Structure of Extratropical Cyclones?” In: *Journal of Climate* 23.7, pp. 1621–1635. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/2009JCLI3318.1. (Visited on 06/10/2024).

Driver, O. G. A. et al. (Mar. 2025). “Factors limiting contrail detection in satellite imagery”. In: *Atmospheric Measurement Techniques* 18.5, pp. 1115–1134. ISSN: 1867-1381. DOI: 10.5194/amt-18-1115-2025. (Visited on 04/09/2025).

Duda, D. P. and P. Minnis (2009a). “Basic Diagnosis and Prediction of Persistent Contrail Occurrence Using High-Resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part I: Effects of Random Error”. In: *Journal of Applied Meteorology and Climatology* 48.9, pp. 1780–1789. ISSN: 1558-8424. JSTOR: 26173662. URL: https://www.jstor.org/stable/26173662 (visited on 09/15/2025).

– (2009b). “Basic Diagnosis and Prediction of Persistent Contrail Occurrence Using High-Resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part II: Evaluation of Sample Models”. In: *Journal of Applied Meteorology and Climatology* 48.9, pp. 1790–1802. ISSN: 1558-8424. JSTOR: 26173663. URL: https://www.jstor.org/stable/26173663 (visited on 09/15/2025).

ECMWF (June 2023). “IFS Documentation CY48R1 - Part III: Dynamics and Numerical Procedures”. In: *IFS Documentation CY48R1*. IFS Documentation 3. DOI: 10.21957/26f0ad3473.

Efron, B. (Jan. 1979). “Bootstrap Methods: Another Look at the Jackknife”. In: *The Annals of Statistics* 7.1, pp. 1–26. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/aos/1176344552. (Visited on 09/15/2025).

Engberg, Z. et al. (Jan. 2025). “Forecasting contrail climate forcing for flight planning and air traffic management applications: the CocipGrid model in pycontrails 0.51.0”. In: *Geoscientific Model Development* 18.2, pp. 253–286. ISSN: 1991-959X. DOI: 10.5194/gmd-18-253-2025. (Visited on 02/03/2025).

Immler, F. et al. (Mar. 2008). “Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes”. In: *Atmospheric Chemistry and Physics* 8.6, pp. 1689–1699. ISSN: 1680-7316. DOI: 10.5194/acp-8-1689-2008. (Visited on 09/08/2025).

Kästner, M. et al. (1999). “Influence of weather conditions on the distribution of persistent contrails”. In: *Meteorological Applications* 6.3, pp. 261–271. ISSN: 1469-8080. DOI: 10.1017/S1350482799001231. (Visited on 09/08/2025).

Kübbeler, M. et al. (June 2011). “Thin and subvisible cirrus and contrails in a subsaturated environment”. In: *Atmospheric Chemistry and Physics* 11.12, pp. 5853–5865. ISSN: 1680-7316. DOI: 10.5194/acp-11-5853-2011. (Visited on 09/15/2025).

Lee, D. et al. (Jan. 2021). “The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018”. In: *Atmospheric Environment* 244, p. 117834. ISSN: 13522310. DOI: 10.1016/j.atmosenv.2020.117834. (Visited on 10/04/2022).

Li, Y. et al. (Feb. 2023). “Upper-tropospheric slightly ice-subsaturated regions: frequency of occurrence and statistical evidence for the appearance of contrail cirrus”. In: *Atmospheric Chemistry and Physics* 23.3, pp. 2251–2271. ISSN: 1680-7316. DOI: 10.5194/acp-23-2251-2023. (Visited on 09/12/2025).

Mann, H. B. and D. R. Whitney (Mar. 1947). “On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other”. In: *The Annals of Mathematical Statistics* 18.1, pp. 50–60. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177730491. (Visited on 09/24/2025).

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).

Murphy, D. M. and T. Koop (2005). “Review of the vapour pressures of ice and supercooled water for atmospheric applications”. In: *Quarterly Journal of the Royal Meteorological Society* 131.608, pp. 1539–1565. ISSN: 1477-870X. DOI: 10.1256/qj.04.94. (Visited on 02/12/2025).

Neis, P. et al. (Mar. 2015). “Evaluation of the MOZAIC Capacitive Hygrometer during the airborne field study CIRRUS-III”. In: *Atmospheric Measurement Techniques* 8.3, pp. 1233–1243. ISSN: 1867-1381. DOI: 10.5194/amt-8-1233-2015. (Visited on 09/10/2025).

Petzold, A. et al. (July 2020). “Ice-supersaturated air masses in the northern mid-latitudes from regular in situ observations by passenger aircraft: vertical distribution, seasonality and tropospheric fingerprint”. In: *Atmospheric Chemistry and Physics* 20.13, pp. 8157–8179. ISSN: 1680-7316. DOI: 10.5194/acp-20-8157-2020. (Visited on 09/10/2025).

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. (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).

Shapiro, M. et al. (Nov. 2023). *pycontrails: Python library for modeling aviation climate impacts*. Zenodo [code]. DOI: 10.5281/zenodo.10182539. (Visited on 04/03/2024).

Sperber, D. and K. Gierens (Dec. 2023). “Towards a more reliable forecast of ice supersaturation: concept of a one-moment ice-cloud scheme that avoids saturation adjustment”. In: *Atmospheric Chemistry and Physics* 23.24, pp. 15609–15627. ISSN: 1680-7316. DOI: 10.5194/acp-23-15609-2023. (Visited on 09/08/2025).

Spichtinger, P. et al. (Mar. 2005). “A case study on the formation and evolution of ice supersaturation in the vicinity of a warm conveyor belt’s outflow region”. In: *Atmospheric Chemistry and Physics* 5.4, pp. 973–987. ISSN: 1680-7316. DOI: 10.5194/acp-5-973-2005. (Visited on 09/15/2025).

Teoh, R. et al. (Aug. 2022). “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. (May 2024a). “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).

Teoh, R. et al. (Jan. 2024b). “The high-resolution Global Aviation emissions Inventory based on ADS-B (GAIA) for 2019–2021”. In: *Atmospheric Chemistry and Physics* 24.1, pp. 725–744. ISSN: 1680-7316. DOI: 10.5194/acp-24-725-2024. (Visited on 04/08/2024).

Thompson, G. et al. (Dec. 2024). “On the fidelity of high-resolution numerical weather forecasts of contrail-favorable conditions”. In: *Atmospheric Research* 311, p. 107663. ISSN: 0169-8095. DOI: 10.1016/j.atmosres.2024.107663. (Visited on 01/22/2025).

Wang, Z. et al. (Mar. 2025). “Machine learning for improvement of upper-tropospheric relative humidity in ERA5 weather model data”. In: *Atmospheric Chemistry and Physics* 25.5, pp. 2845–2861. ISSN: 1680-7316. DOI: 10.5194/acp-25-2845-2025. (Visited on 09/12/2025).
Wernli, H. et al. (2016). “A trajectory-based classification of ERA-Interim ice clouds in the region of the North Atlantic storm track”. In: *Geophysical Research Letters* 43.12, pp. 6657–6664. ISSN: 1944-8007. DOI: 10.1002/2016GL068922. (Visited on 09/10/2025).
Wilhelm, L. et al. (Jan. 2022). “Meteorological Conditions That Promote Persistent Contrails”. In: *Applied Sciences* 12.9, p. 4450. ISSN: 2076-3417. DOI: 10.3390/app12094450. (Visited on 09/15/2025).
Wolf, K. et al. (Apr. 2024). “Distribution and morphology of non-persistent contrail and persistent contrail formation areas in ERA5”. In: *Atmospheric Chemistry and Physics* 24.8, pp. 5009–5024. ISSN: 1680-7316. DOI: 10.5194/acp-24-5009-2024. (Visited on 06/14/2024).
Wolf, K. et al. (Jan. 2025). “Correction of ERA5 temperature and relative humidity biases by bivariate quantile mapping for contrail formation analysis”. In: *Atmospheric Chemistry and Physics* 25.1, pp. 157–181. ISSN: 1680-7316. DOI: 10.5194/acp-25-157-2025. (Visited on 01/22/2025).