We sincerely thank the associate editor for their helpful feedback on this paper. We address all their comments and suggestions below.

Associate editor comments are in black italicised text.

Our responses are in blue, regular text.

Where practical/necessary, we provide a screenshot of the track-changes document to show the changes that we have made (in outlined boxes). In these, text that is removed is ~~struck through~~ and coloured red, while new text is underlined and coloured blue.

Ian Ashpole (on behalf of both authors).

Associate Editor Report

Dear authors,

Thank you for providing the revised version of your manuscript. While your response addresses most of the reviewer comments, I think some key points are currently not addressed sufficiently.

One major aim of your work is to demonstrate that a L3 "land-only" product (L3L) is better suited for studying CO VMRs and trends in coastal grid boxes than the official L3 product (L3O), which has issues because retrievals over land and water are often combined. You show in detail the difference between the L3L and L3O dataset. However, in order to show that a dataset is better, it is also necessary to show that the calculated VMRs and trends are more accurate. These points are brought up by reviewer #2, who questions the feasibility to compute "meaningful" mean values and trends from the dataset (Point 2.3 and 2.4). To consider your manuscript for publication, I think it is necessary that you consider the following points in more details:

(1) Please provide time series for the different coastal grid boxes to assess the temporal coverage for the different datasets.

A focused case study, including more detailed time series analysis, is now included at the start of the city focus section (new Section 3.4.1; screenshot below). This clearly demonstrates the stark improvement in temporal coverage of the L3L dataset, when compared to $L3O_L$ (the subset that users of L3O are encouraged to use) and the differences in detected temporal trends and mean VMRs. This also demonstrates the trend and mean VMR differences between L3L and $L3O_{LM}$, and how they are clearly linked to the inclusion of retrievals over water in $L3O_{LM}$ (by comparison with results from L3W).

705 **3.4. Illustrative examples comparing L3O and L3L: analysis of the most populous coastal cities**
706
707 In this section, we analyse time series from the 33 L3 coastal grid boxes that contain cities classified amongst
708 the 100 most populous in the world (derivation outlined in Sect. 2.5) to illustrate the differences between
709 mean values and trends obtained from the L3O and L3L datasets.~~potential consequences of working with the~~

710 ~~L3O dataset, compared to L3L.~~ We focus our comparison on $L3O_L$ and $L3O_{LM}$, as these are the L3O subsets
711 that data users would realistically choose to analyse if following the data filtering guidelines. For clarity, ~~we~~
712 from here~~on refer to~~ these grid boxes are referred to by the name of the city that they contain. A detailed case
713 study for the L3 grid box containing the city of Dubai is first presented, before considering results for all
714 cities analysed.
715
716
717 **3.4.1. Detailed case study: L3 grid box containing Dubai**
718
719 Summary stats derived from the L3O subsets, L3L, and L3W (included for comparison), for the L3 grid box
720 containing the city of Dubai, are given in Table 5. Figure 8 visualises the daily retrieved VMR time series
721 from L3L, with $L3O_L$ overlaid for comparison purposes.

Note, we have also made one additional change to the results presentation in Section 3.4 (specifically, this change is relevant to what is now Section 3.4.2 but spanned Sect. 3.4.1 and 3.4.2 in the original submission): Instead of presenting mean VMRs and temporal trends (with standard deviations and standard errors as error bars, respectively) for every city in figure form (Fig. 8 and 9, resp., in the original submission), we have replaced the figures with a single table (new Table 6). The benefit of this is that it enables us to present all of the information in one place in as clear a fashion as possible, and makes the sample sizes of each dataset compared more obvious to readers. It also enables us to easily include information on the relative land/water surface cover for each grid box, as requested in your second point (see below). We also bore in mind that Reviewer 1 had some difficulty interpreting the meaning of different line styles/symbols in the figures in their original review, and although they seemed satisfied by how we dealt with this (noting their recommendation of "publish as is" following review), the figure complexity increased once we addressed the need to make them more legible for readers with color deficiencies. In summary, we feel that this new presentation is much simpler for readers to follow. The text discussing the results in Sect. 3.4 is largely unaffected by this change, with the exception of minor superficial changes required to account for the replacement of Figures 8 and 9 with Table 6.

Regarding concerns from original Reviewer #2 about the "meaningful(ness)" of trends presented in this paper and how this is affected by patchy temporal coverage: this is ultimately limited by the temporal coverage of the MOPITT instrument itself, and is something that we cannot affect. However, it does not prevent scientists from presenting trends based on these data in publications (e.g. Buchholz et al. 2022 https://doi.org/10.1038/s41467-022-29623-8). The central point that we demonstrate in our paper is that our L3L dataset is favorable over the original L3 data in coastal areas due to 1) demonstrably more days with land-only data due to less discarded retrievals; and 2) demonstrably greater information content in the retrievals being regressed when L3O filtering is allowed to include retrievals over water (the $L3O_{LM}$ subset that we analyse). Both of these points increase the likelihood that mean VMRs and trends in L3L are more "meaningful" than in the original L3 data in these coastal areas, although it is obviously still constrained by limitations in instrument coverage. However, this does not mean that the results have no scientific significance.

(2) Please analysis the spatial coverage of L2 pixels in the L3 grid boxes to understand how the different datasets are affected by sparse sampling in the grid boxes.

This was already done in our original submission, but on reflection we can understand how it was not clear. For greater clarity, we have added an explanation to the methods section (Sect. 2.4) about how the spatial coverage of the L2 pixels in L3 grid boxes is analysed (screenshot below). We introduce the metric ratio(land/water) to quantify the proportion of the grid box surface that is covered by land vs water, and refer to this in the following sections when variable surface coverage is discussed as an explanation for results: Sect. 3.2.2 (screenshot below), Sect. 3.4.1, Sect. 3.4.2.

293 L3L and L3W retrievals. Additionally, the number of L2 retrievals that are used for calculating the area
294 averages when creating L3L and L3W ("n_ret$_L$" and "n_ret$_W$", respectively) is recorded. The ratio
295 n_ret$_L$/n_ret$_W$ (herein referred to as "ratio(land/water)" for simplicity) is used to indicate the proportion of
296 the L3 grid box that is covered by land vs water: a ratio of 1 indicates an even split of these surface types in
297 the grid box; a ratio < 1 indicates that a greater proportion of its surface is water covered; and a ratio > 1
298 indicates that the grid box is land-dominated.

562 • The grid boxes of BOTH tend to have a greater proportion of their surface covered by water than land
563 when compared to L3L_L3W_ONLY. This is determined by analysis of ratio(land/water) values for
564 each grid box (derivation of this metric is outlined in Sect. 2.4). ~~quantified by comparing the mean~~
565 ~~number of L2 retrievals over land and water that are averaged together to make L3L and L3W each~~
566 ~~day ("n_ret(L3L)" and "n_ret(L3W)"), for each coastal grid box compared.~~ A mean ~~n_ret(L3L/L3W)~~
567 ~~ratio~~ratio(land/water) of 0.87 for BOTH indicates a greater water influence on $L3O_M$ than for the grid
568 boxes of L3L_L3W_ONLY, for which a mean ratio(land/water)~~n_ret(L3L/L3W) ratio~~ of 1.00
569 indicates a more even land/water split. Thus, $L3O_M$ more closely resembles L3W – which is
570 significantly different to L3L – in BOTH than in L3L_L3W_ONLY.

(3) Please conduct a suitable statistical analysis for computing means and trends, which can show that the L3L dataset increases accuracy of the analysis.

The Theil-Sen slope estimator has been used to verify that the trends presented in the paper, calculated using Weighted Least Squares (WLS) regression, do not change depending on the method used (Theil-Sen is nonparametric and therefore less sensitive to outliers). These are included in the Supp. Mat. (SM7) for reference (referred to in the text).
Beyond this, we are unclear about what a "suitable statistical analysis for computing means and trend" constitutes. We asked for clarification on this from the Associate Editor in a personal communication and were directed that a different method for calculating trends might be necessary (email dated August 15 2022) – hence, our inclusion of Theil Sen analysis.

(4) Please validate your mean values and trends with independent measurements for some cases to proof that such a product would be useful. I am aware that it might be impossible to find suitable measurements for such an analysis.

As acknowledged in this request, we have been unable to find surface-level CO measurements for any of the coastal cities analysed to validate the mean values and trends compared. CO data *are* available for some of these cities, but these are restricted to total column measurements as part of the TCCON network. We do not consider the CO total column in this paper. Despite the lack of verification, which we feel is beyond the scope of this paper, the results are still of scientific significance in that they 1) demonstrate that the scientific value of a publicly available satellite dataset (MOPTIT L3O) is lessened in certain situations

(coastal grid boxes) as a direct consequence of how it is created from finer resolution parent data (MOPITT L2), and 2) present a solution to this. Note that our L3L dataset is now published, as recommended below.

Please also consider the following suggestion: In your manuscript, you conclude that a L3 "land-only" product would be beneficial to the research community. I noticed that your study already creates this product. Since AMT strongly encourages the publication of underlying data, I suggest that you publish your "land-only product for coastal grid boxes" in a public data repository (e.g., zenodo.org) and revise your manuscript changing the focus to describe your new dataset, analyzing the difference to the official L3 product and show that it is better suited for computing means and trends. This small change would enhance the scientific significance of your manuscript. Note that this would still require addressing the points above.

Thank you for this suggestion. The L3L dataset (and L3W, for comparison) have been uploaded to a public data repository and are available for download from the following link: https://doi.org/10.5683/SP3/ERCG2H. We have modified the text in the Abstract, Introduction (screenshot below), Data and Methods, and Conclusion (screenshot below) sections of the paper to reflect the fact that these are now publicly available datasets, with download instructions given in the Data Availability section. Please note that this has not significantly changed the focus of the paper, which remains demonstrating L3O shortcomings over coastal grid boxes.
The full citation information for the published L3L and L3W datasets is:
"Ashpole, I., and Wiacek, A.: Land- and water-only Level 3 products from MOPITT TIR-NIR Version 8 CO retrievals, https://doi.org/10.5683/SP3/ERCG2H, Borealis, V1, 2022".

especially to less-expert users, who may lack the expertise required to scrutinize the data for potential a priori bias. Secondly, many of the world's largest agglomerations are situated within a coastal L3 grid box (5 of the top 10 and 33 of the top 100 largest agglomerations by population; derivation outlined in Sect. 2.5), making these likely targets for analyses of air quality indicators, especially their changes over time.

This paper presents a comparison of results from analyses performed using original, "as downloaded" L3 data products, and a new ~~separate~~ land-only ~~and water-only~~ L3 product ("L3L", Ashpole and Wiacek (2022) – outlined in Sect. 2.4) that has been created ~~area averages~~ from L2 products, for all MOPITT L3 grid boxes that overlay coastlines (a water-only L3 product "L3W" has also been created for comparison purposes) ~~("L3L" and "L3W" respectively – derivation outlined in Sect. 2.4)~~. Section 2 describes the datasets and methods used, including outlining the creation of the new L3L and L3W data products analysed in this paper. Section 3.1 demonstrates the magnitude of the sensitivity difference for retrievals over land and water,

892 **4. Summary and Conclusions**
893
894 Motivated by the work of Ashpole and Wiacek (2020) which demonstrated, for the MOPITT L3 grid box
895 containing the coastal city of Halifax, Canada, that mean VMR statistics and temporal trends differ depending
896 on whether L2 or L3 data are analysed, this paper has examined what proportion of all coastal L3 grid boxes
897 also see differences between results from analyses performed with L2 and L3 data. While it is recommended
898 to MOPITT data users that analyses are restricted to retrievals performed over land owing to known
899 sensitivity issues over water (MOPITT Algorithm Development Team, 2018; Deeter et al., 2015), such
900 recommendations cannot practically be followed by users of L3 data for coastal grid boxes owing to the way
901 the data are created from their bounded L2 retrievals. In short, this study has sought to answer the question:
902 "does it matter"? Analysis has focussed on comparing the original, "as-downloaded" L3 dataset ("L3O")
903 with new land-only and water-only L3 products ("L3L" and "L3W" respectively) that have been created from
904 the L2 retrievals. The main results are summarised below.

Please feel free to contact me, if you have any questions.

Sincere thanks for your help and attention with this.

Kind regards

Gerrit Kuhlmann

PS: Many of your figures are difficult to access for readers with color vision deficiencies. Please have a look at the guidelines on figures and tables on the AMT website:

https://www.atmospheric-measurement-techniques.net/submission.html#figurestables

Apologies for overlooking this in our original submission. We have modified the following Figures: Fig. 1, Fig. 5, Fig. 6, Fig. 7. We have also replaced Figs 8 and 9 from the original submission with a Table (Table 6), which we feel is a much more straightforward presentation of the results.