We sincerely thank both reviewers for their helpful feedback on this paper, and the associate editor for their time with this. We address all the reviewer comments and suggestions below.

Reviewer 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.

Note that the revised submission also contains a change to one of the tables, following a notification that we received from the Editorial Support Team at the previous submission stage. The change is detailed at the end of this document, along with our response to one other request for information/clarification from them regarding figure copyrights.

Ian Ashpole (on behalf of both authors).

REVIEWER REPORT #1

(Anonymous referee #3)

Suggestions for revision or reasons for rejection

The study analyzes the difference, impact and vertical sensitivity of coastal pixels (land, water or mixed) from the L2 MOPITT v8 surface CO retrievals used and averaged for L3 products. The study is of interest to the scientific community using satellite retrievals of CO. Comments can be found below.

• The manuscript is well written but can be difficult to follow with all the annotations used. I would recommend using a table in the methodology describing all datasets in this study (L3OL, L3L, L3O, ..).

Thank you for this suggestion. We have included this in the Data and Methods section as a new Table 1, and referenced it where appropriate when outlining the datasets used in the study.

Table 1. List of dataset short names used in the main article text, and their corresponding full descriptive name.

Dataset short name	Full descriptive dataset name
L3O	Original, "as downloaded" Level 3 (L3) dataset
$L3O_L$	Subset of L3O only containing L3 retrievals with a surface index of land
$L3O_M$	Subset of L3O only containing L3 retrievals with a surface index of mixed
$L3O_{LM}$	Subset of L3O only containing L3 retrievals with a surface index of land OR mixed
$L3O_W$	Subset of L3O only containing L3 retrievals with a surface index of water
$L3O_{NF}$	The L3O dataset with no filtering by surface index ($L3O_{NF}$ is identical to L3O)
L3L	A new L3 "land-only" dataset, created only from Level 2 retrievals performed over land (creation method outlined in Sect. 2.4)
L3W	A new L3 "water-only" dataset, created only from Level 2 retrievals performed over water (creation method outlined in Sect. 2.4)

In addition, to aid readers in following the paper we have included a list of short names and abbreviations used throughout the text, along with their full descriptive name, purpose for use and section introduced as an Appendix at the end of the paper (Appendix 1).

We have also made small changes to the text throughout the paper to aid with clarity and flow where appropriate, and have thoroughly edited the abstract, introduction, and conclusion with a focus on stating the paper aim more clearly. The focus of the writing is now more clearly on the surface level in Section 3.1.

• *Ln.195. Further details on surface index used in MOPITT should be provided. How is defined the surface index in the L2 retrievals. What are the uncertainties associated with this surface index?*

This is a great question, but one that, unfortunately, we are unable to answer. We have searched all available Product User Guides, ATBD's, Data Quality Summaries, and product validation / algorithm description papers listen on the NCAR MOPITT website (https://www2.acom.ucar.edu/mopitt/) to get an answer for this (including looking at the documentation for older product versions), but to the best of our knowledge there is no information publicly available explaining where the surface index for L2 retrievals is actually derived from. The lack of transparent information on this matter, and the fact that it is not mentioned in any publication by the MOPITT Science team, suggests to us that whatever method is used to define the surface index in L2 retrievals may be standard practise, with little uncertainty about the L2 surface indexes prescribed.

Without reliable information on this matter, we feel uncomfortable speculating, and have therefore not modified the paper in response to this comment. We hope this is understandable/acceptable, given the lack of further information.

• *Ln. 230. It is suggested that these inland false coastal grid boxes would be linked to flood and appear some of the time. Have you looked at this assumption (using cumulative precipitation or soil moisture data)? What do you mean by "some of the time"? Have you looked by season, or month? If one of the cause is due to surface ice cover or flood, this should be observed depending on seasons.*

We have not tested this assumption by analysing the seasonal distribution of "false coastal" grid boxes, or using different data sources (e.g. precipitation or soil moisture, as suggested) to verify their occurrence due to e.g. flooding or ice coverage. Our reasoning is that the aim of this step is to identify coastal grid boxes to study ("2.3. Coastal grid box classification for this study"), as opposed to analysing the distribution of grid boxes with a surface index of "mixed", and potential reasons why they get this flag. Our discussion of "false coastal" grid boxes stems from the fact that some inland areas are also classified as "mixed", which meant that we could not simply use the presence of a surface index of "mixed" as a coastal identifier (Fig.

2a in the paper). Additional analysis around this subject is beyond the scope of the paper, and would add unnecessary complication and additional length. We hope this is acceptable.

Reflecting on the text in this section (2.3), we wonder if it is the use of the term “false coastal grid boxes” that has prompted this question, since the word “false” implies that something is in error – which is not the case. These grid boxes are simply classified as mixed “some of the time” (see below for refinement of that terminology). There is no error. To help prevent such questions from arising, we have replaced the term “false coastal” with “inland_mixed”, since this better reflects the actual grid box classification. We have also made minor changes to some of the wording in this section, to greater aid clarity.

Regarding what is meant by “some of the time”: we have replaced this with the wording “at least once during the study period” to remove the vagueness.

> true for coastal grid boxes. However, analysis of the global distribution of L3 grid boxes featuring a surface index of mixed revealed that, in addition to actual coastlines, a large proportion of inland grid boxes that are clearly not coastal ~~(“false coastal”)~~ are given the surface index of mixed at least ~~some of the time~~once during the study period (“inland_mixed”; Fig. 2a). The reason for this is unclear, but it could be for real physical reasons, such as land grid boxes sporadically flooding, or due to issues in the retrieval schemes caused by e.g. cloud screening problems or the presence of surface ice cover. One characteristic of these ~~false coastal~~inland_mixed grid boxes is that, compared to the total number of days with L3, the relative frequency with which they are flagged as land is very high (expressed as the ratio “n_days($L3O_L$/L3O)”~~)”~~, plotted in Fig. 2b; a list of short names and abbreviations referred to in the text can be found in Appendix A for reference). This relative frequency is much lower for “true” coastal grid boxes, to be expected given prior

• *The study works at global scale and for a 20 years period. It is difficult to learn if the sensitivity of AVK for instance in section 3.1.1 is similar for every regions across the globe. Additionally, what would have been the results if analyzed by seasons and by regions/latitudes (depending on ice cover, swamps, ..)? This is an information missing which can be of interest for the users.*

This is a fair comment, and we have now included a brief analysis of latitudinal and seasonal variability in the land-water retrieval sensitivity contrast in the Supplementary Material (SM3). The analysis shows that there is indeed some seasonal and latitudinal variance in the magnitude of the land-water sensitivity contrast: there is a tendency for greater land-water retrieval sensitivity differences in the Northern Hemisphere than Southern Hemisphere when averaged across the year, with some nuances by season, although no clear and obvious patterns. A land-water sensitivity gradient is generally evident irrespective of latitude or season.

Given that this analysis does not alter the overall results of the paper, and that discussing the nuances would add words and figures to an already quite long and complex paper (as noted in reviewer report #2), we feel that the Supp. Mat. is the best location for it.

> An analysis of latitudinal and seasonal variability in the land-water surface level retrieval sensitivity contrast is provided in the Supp. Mat. (SM3). Briefly, this shows a tendency for greater land-water retrieval sensitivity differences in the Northern Hemisphere than Southern Hemisphere when averaged across the year. The land-water AK rowsum differences tend to vary least by season in the tropical regions (between 30° South and 30° North) and show the greatest contrast in the midlatitudes (30° – 60°) in the respective hemisphere's spring and summer months, with smallest differences in the winter months. Overall, a land-water sensitivity contrast is evident irrespective of latitude or season.

• Ln. 456. Which additional physical factors could play a role?
You have analyzed the impact of wind on the land-water difference observed in the retrieved VMRs at different locations around the world. But one area of interest you do not mention is the transport of CO concentrations from African fires to the coastal lands of South America. During Africa's fire season, CO concentrations can be found higher over ocean than on land along the African or Brazilian coasts. Have you examined and considered specific cases like that?

Good point. We have not studied this specific case, but do already mention in this section that ocean air is not necessarily "pristine" compared to land-based air (see the quotation below). We have added a sentence more explicitly acknowledging that in some specific cases CO VMRs can be greater over water than land (see below), but err away from including additional case studies and detail at this stage. Note that whether CO VMRs are higher over land or water is not the central point in question in this paper; it is the magnitude of the difference, which we argue is too great to be explainable by physical factors alone and therefore brings the land-water contrast in retrieval sensitivity into question. This issue is navigated by our use of absolute land-water CO VMR differences (i.e. ignoring whether the CO VMR is greater over land or water), when relating them to land-water differences in retrieval sensitivity parameters.

relatively small distance covered by a L3 grid box. Given the relatively long-lived, well-mixed nature of atmospheric CO, VMRs retrieved at a given location are a function of both local emissions *and* transport, and the portion of coastal L3 grid boxes situated over water therefore do not represent pristine conditions in

22

comparison to the adjacent land-based portion of the grid boxes. This is verified by comparing a priori VMRs (also shown in Fig.~~ure~~ 4), which suggest the land-water difference in CO concentrations should be negligible (mean L3L – L3W a priori VMR difference = 0.69 ppbv, compared to a mean retrieved VMR difference of 10.29 ppbv). Indeed, in some specific cases – e.g. uninhabited coastal areas downwind of large trans-oceanic pollution sources – VMRs may be higher over the water portion of coastal gridboxes than the adjacent land portion (note that Fig. 4 does show that this is the case in some grid boxes). The above reasoning can also be

• *Ln. 569. Could you detail what further information about the "true" VMR is needed?*

Apologies if this sentence caused confusion. What we mean by this is that the actual ("true") VMR amount that is being measured would need to be known if an assessment about the accuracy of a priori is to be made. Without this knowledge we can only speculate, but in the context of this section it is possible to make an informed guess when the information about retrieval sensitivity (given by averaging kernels and proxied by VMR ret-apr values) is available.

We have modified the text to make it more clear what we mean by "information about the "true" VMR":

corresponding L3L retrieval. Interestingly, it is also notable that retrieved and a priori VMRs are lower in BOTH$_{VMRs}$ than in L3L_L3W_ONLY$_{VMRs}$, and that retrieved minus a priori VMR values are greater in BOTH$_{VMRs}$ than in L3L_L3W_ONLY$_{VMRs}$. This could imply that the a priori VMRs are closer to reality (i.e.

28

the a priori CO amount is closer in value to the actual ("true") CO amount that is being measured) for the grid boxes of L3L_L3W_ONLY$_{VMRs}$ than those of BOTH$_{VMRs}$, however ~~further information on "true" VMRs is required~~ to properly assess this it would be necessary to know what the actual "true" VMR values are that are being measured.

• Table 5 and Ln. 736. "L3OL trend being significantly stronger than the L3L […] given the far superior temporal coverage of L3L, this is the more reliable result". L3OL is the only dataset with less than 90% days with data and with the larger standard deviation (+/- 56 ppbv). While L3L has a superior temporal coverage, its standard deviation is important (+/- 44 ppbv) in comparison to the other datasets. How do you explain this large standard deviation with L3L, not observed with L3OLM and L3W?

Good question. The greater standard deviation in L3L (and L3O_L) than in L3O_{LM} and L3W is a function of the greater sensitivity in retrievals over land than over water. The more sensitive retrievals over land can deviate more from their a priori VMRs – and therefore reach higher (and lower) retrieved VMR values – than the retrievals over water which, due to their lower sensitivity, are tied more closely to their a priori VMRs, resulting in lower variance and standard deviation. The observation of reduced variability in retrievals over water than land is discussed in detail in Section 3.2.1 ("Retrieved VMR comparison between L3L and L3W") – specifically, this sentence makes the point clearly: "This may be explained as follows: when sensitivity over water is especially low…the retrieved VMR will be heavily weighted by the a priori and unable to match the variation present in the more sensitive retrieval over land. As sensitivity over water increases, this a priori weighting weakens and the retrieved VMR will more closely track the retrieval over land." (This quote is from L487 – 492 in the revised manuscript).

We have added a sentence highlighting the greater standard deviation in retrievals over land, which is an expected finding given the discussion in earlier sections of the paper:

> mean differences are significant ($p < 0.1$). Consistent with the results shown in Sect. 3.2.2 when identifying factors that determine whether the averaging of L2 retrievals over land and water to create L3O_M can yield statistically significantly different retrievals to L3L, this L3 grid box is water-dominated, with a mean ratio(land/water) of 0.60. It is also notable that the standard deviation of long-term mean retrieved VMR in L3L (and L3O_L) is roughly twice as large as that in L3O_{LM} and L3W, which is to be expected given that retrievals over water are more greatly tied to their a priori than retrievals over land due to their comparatively lower sensitivity (as discussed in Sect. 3.2.1).

• *Legend Fig. 1. NASA of nasa blue Marble should be in capital letters.*

Thank you for pointing this out. The change has been made.

> classification is outlined in Sect. 2.3. The coastal L3 grid box visualized here contains the city of Dubai (~centre = 55.296° E, 25.277° N), which features in the case study analysis of Sect 3.4. Faint background shading is from ~~Nasa~~ NASA Blue Marble imagery.

REVIEWER REPORT #2

(Anonymous referee #4)

Suggestions for revisions or reasons for rejection

Paper conclusion and abstract need to be summarized in a more clear way. The objective of the paper needs to be clarified.

We apologise for the lack of clarity. The abstract, introduction, and conclusion have been edited with a focus on stating the paper aim more clearly. We have also made some minor additional changes to the text throughout the paper to aid with clarity and flow where appropriate.

Notes from Report:

The authors presents a comparison of results from analyses performed using original L3 MOPITT data products, and a new land-only L3 product ("L3L") and a water-only L3 product ("L3W) that have been created from L2 products, for all MOPITT L3 grid boxes that overlay coastlines. Comparing the full L3O dataset to L3L, it is shown that if L3O is filtered so that only retrievals over land (L3OL) are analyzed, there is a huge loss of days within the data. This is because L2 retrievals over land are routinely discarded during the L3O creation process. Even by retaining L3OM (mixed) retrievals, the resulting L3O "land or mixed" (L3OLM) subset still has less data days than L3L for 61 % of coastal grid boxes. The loss of data influenced the results where it is shown that, the mean VMRs in L3OL and L3L differ significantly for 11 of the 27 grid boxes that can be compared. They concluded that a L3 product based only on L2 retrievals over land – the L3L product analyzed in this paper, could be of benefit to MOPITT data users, given the significant differences in mean CO VMRs and trends that can be obtained for coastal grid boxes using L2 products.

The paper can be published with minor changes.

The main challenge in the paper is the length and the clarity. Paper needs to be shortened and rewritten in a clear way. It is very hard for regular user to follow up the flow of the paper.

As noted above, the abstract, introduction, and conclusion have been edited with a focus on stating the paper aim more clearly. We have also made some minor additional changes to the text throughout the paper to aid with clarity and flow where appropriate (e.g., the focus of the writing is now more clearly on the surface level in Section 3.1). In addition, to aid readers in following the paper we have included a list of short names and abbreviations used throughout the text, along with their full descriptive name, purpose for use and section introduced as an Appendix at the end of the paper(Appendix 1).

Please note that we have been unable to make any notable reduction to the length of the paper: everything that is not essential to the paper aims has already been placed in the Supplementary Material.

RESPONSE TO NOTIFCATION FROM EDITORIAL SUPPORT

Notification to the authors:

Checking your paper, I noticed that your table 6 contains coloured cells. Please note that this will not be possible in the final revised version of the paper due to HTML conversion of the paper. When revising the final version, you can use footnotes or italic/bold font. For now, the process will continue, but please note that the final version cannot be published by using coloured tables. Please check if your Figures S5 to S10 (maps/aerials) require a copyright statement and add it to the figure captions. If you are the originator, you can just inform us.

Thank you for informing us of these requirements. Table 7 (was table 6 in previous submission) has been edited to remove coloured shading from cells, which has been replaced with the use of bold and/or italics where appropriate.

Please also note that we do not need a copyright statement for any figures in supp matt; these were all produced by the authors.