Response to Reviewers – "Identifying episodic carbon monoxide emission events in the MOPITT measurement dataset" by Paul S. Jeffery et al.

We'd like to thank the Reviewers for their helpful comments. Here we address the comments of each Reviewer, with their comments in black and our responses indented in blue.

Reviewer 1

First, I would like to congratulate the authors for this excellent and tremendous amount of work.
Here are a few comments I would like to point out on this manuscript[.]

Thank you for the feedback. Please see below for our responses to your comments.

1. Along with the influence of ENSO, the other key climate drivers like, Indian Ocean Dipole (IOD) and Madden-Julian Oscillation (MJO, especially in sub-seasonal time scale in southeast Asia) could have an impact on the trends in CO. Have you already considered or analyzed these effects?. It would be worth updating the MLR based on these trends from the above mentioned climate driver impacts.

Following this suggestion, we tested a series of other key climate drivers in our regression model; namely the Indian Ocean Dipole (IOD), Quasi Biennial Oscillation (QBO), North Atlantic Oscillation (NAO), Madden-Julian Oscillation (MJO), and the Solar Flux (10.7 cm). As with the MEI term used in the study, the regressor terms used for this all came from the National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory (PSL; `https://psl.noaa.gov/data/climateindices/list/`, last access: December 13, 2023). The tests were performed using an extension of the regression model used in the study to include an additional term for each independent climate regressor; thus no two parameters were tested simultaneously. This resulted in the following model:

$$\mathrm{CO}(t) = a_0 + a_t t + a_{MEI}\mathrm{MEI}(t) + a_{CD}\mathrm{CD}(t). \tag{1}$$

As before, the regressed deseasonalized time series of a bin is expressed as $\mathrm{CO}(t)$ for a given timestep t. The a coefficients correspond to the regression components of the model, with the first two corresponding to the offset, a_0, and linear trend, a_t. The third coefficient corresponds to a model parameterization of the Multivariate ENSO Index (MEI; $\mathrm{MEI}(t)$), and the fourth coefficient corresponds to the parameterized climate driver regressor being tested (CD; $\mathrm{CD}(t)$). It should be noted that the MJO regressor is composed of a pair of orthogonal terms and as such actually has a pair of regression coefficients corresponding to the fit rather than just one.

We tested these additional parameters on a subset of the global 0.5 ° by 0.5 ° gridded daily-mean MOPITT data, chosen to provide approximately global coverage. A total of 580 test bins were used, shown in Fig. 1 The results of this testing focused on two main metrics, the threshold values for each bin, and the number of anomaly events detected in each bin, with specific attention paid to how these two metrics change with the inclusion of these additional terms. These changes are summarized in Table 1. The geographic distributions of the bins where the threshold changed by 10 % or more with the addition of these additional climate drivers are shown Figs. 2 through 6.

From the results of this cursory assessment, it is clear that the addition of any of these terms serves to increase/decrease the anomaly threshold by 5 % (10 %) or more in approximately 255 bins (225 bins) on average. While this change is not insignificant, from the geographic distribution of these affected bins it is found that the inclusion of the additional climate drivers affects the threshold values globally, as opposed to the more localized changes expected from some of these climate drivers, such as the NAO and IOD, which are expected to strongly affect only a limited region. The same holds true for the affected event counts. Furthermore, as all of the additional terms lead to similar numbers of affected bins, there is strong evidence that the additional terms contribute similarly to the fitting process. Altogether, this indicates that the addition of any of these extra climate drivers likely leads to over-fitting the MOPITT CO data, obscuring the anomaly events that are sought in this study. While the potential exists for regional-scale applications of this technique that use these additional parameters, in this study the global-focus renders their usage spurious, and thus we have kept with the use of the regression model used in the paper. That being said, this was a very interesting suggestion that might serve as a potential avenue of future research.

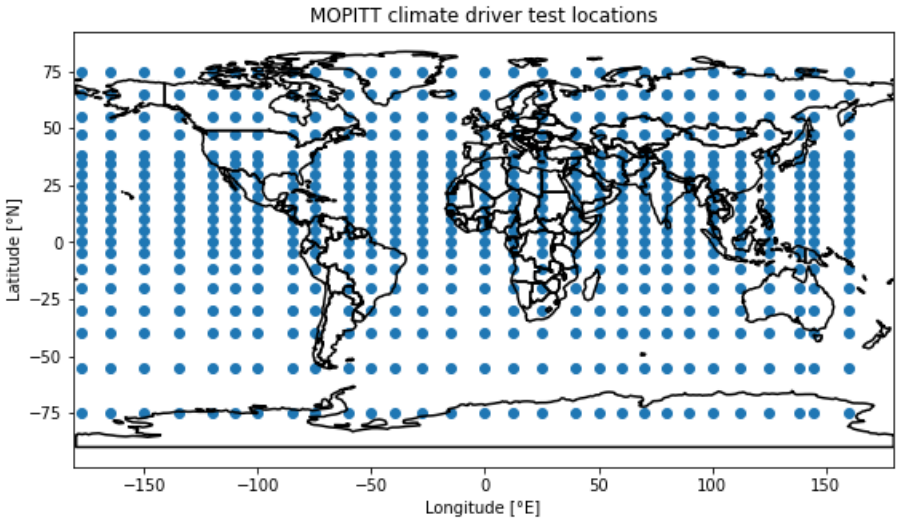

Figure 1: Locations used to test the inclusion of additional climate drivers in the regression model employed in this study.

Table 1: Number of bins with affected threshold values and anomaly event counts found when incorporating additional climate drivers in the regression model. For threshold values, changes of 5+ or 10+ % to the threshold are shown, and for the anomaly event counts, changes of 3+ and 10+ are shown.

Additional climate driver	Affected threshold values		Affected event counts	
	5+ % change	10+ % change	3+ event difference	10+ event difference
MJO	278	238	220	134
IOD	253	223	195	119
QBO	250	218	197	115
NAO	253	221	200	125
Solar Flux	256	227	203	118

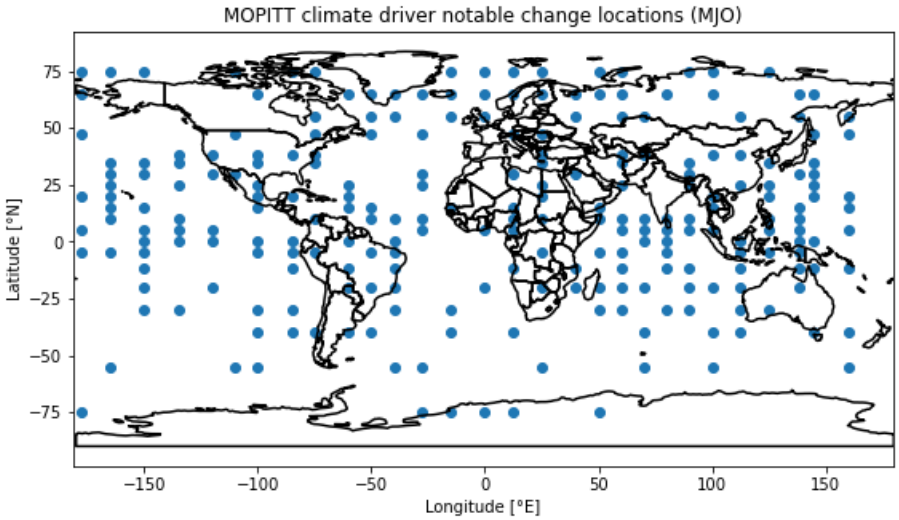

Figure 2: Daily-mean CO bins whose anomaly event threshold is perturbed by at least 10 % by the inclusion of a parametrization for the MJO in the regression model.

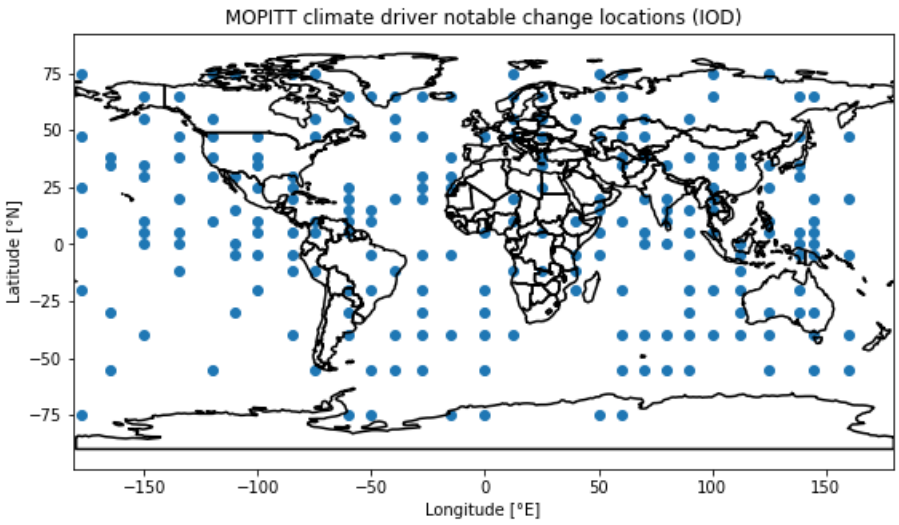

Figure 3: Same as Fig. 2, but for the IOD.

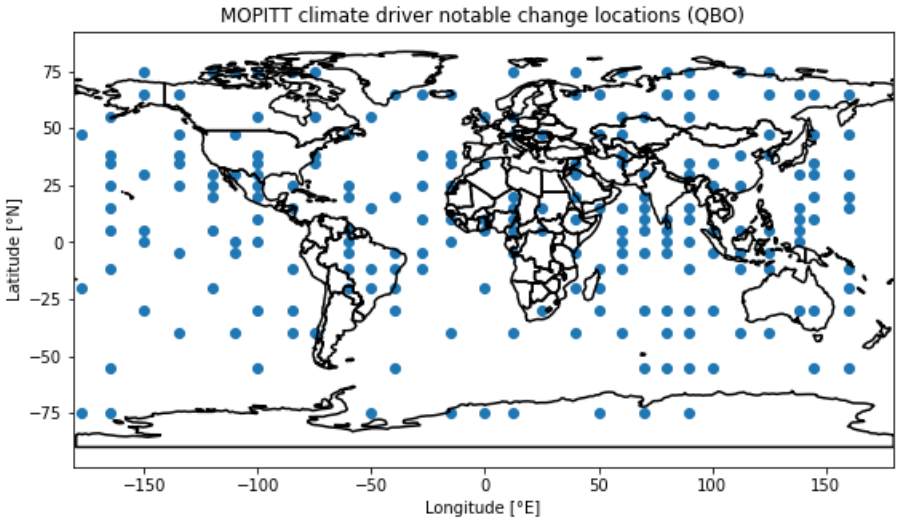

Figure 4: Same as Fig. 2, but for the QBO.

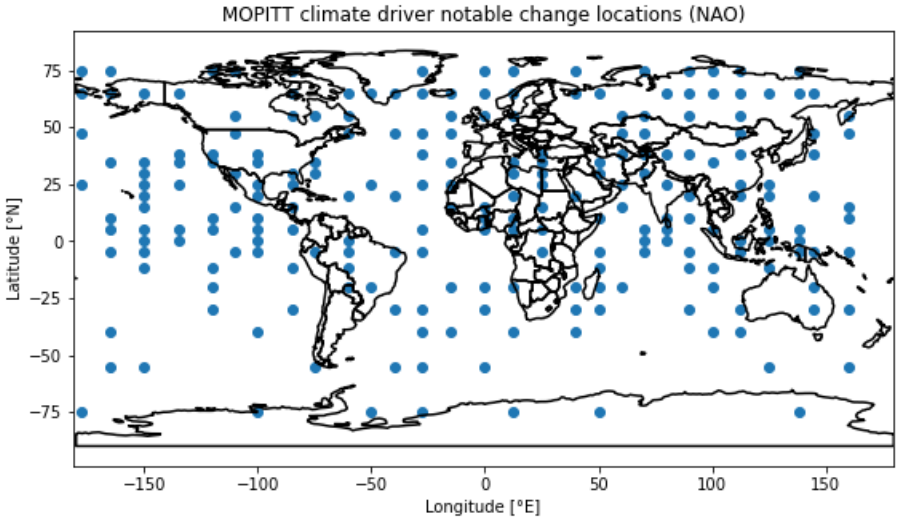

Figure 5: Same as Fig. 2, but for the NAO.

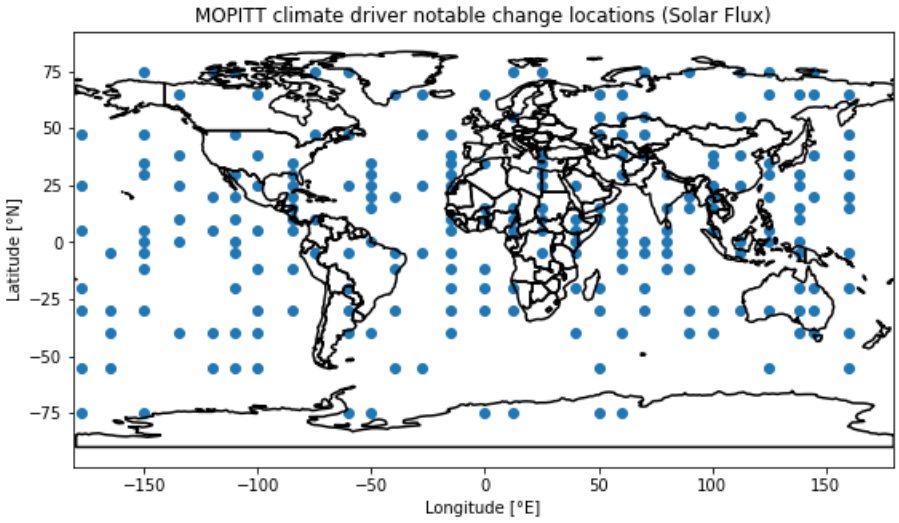

Figure 6: Same as Fig. 2, but for the Solar Flux.

2. A summary table of events including the threshold value for the key regions (with the event classifications) in the study could have been useful.

The non-uniformity of the events between grid cells in a region, alongside the loosely-defined boundaries of the regions discussed, obfuscate our ability to make any sort of table summarizing all events and their thresholds for a given region in a concise fashion. However, the event flag files that we produce contain the desired information for constructing such a table, and for this application we have also made sure to include the threshold values for each anomaly event location in the flag files. Using the event flag files produced here, a user can construct such a table for a focused study.

3. It would be good to add some information about the episodic total column CO intensities also in addition to the frequency and distribution.

We have produced a new version of the flag files that includes an additional field for the episodic total column CO intensities, alongside the anomaly threshold, location, and date of the event. Due to the variable nature of each enhancement event, the authors feel that delving into the episodic total column CO intensities in detail in the paper itself does not contribute meaningfully to the results. Changed L184:

From "Values outside of the threshold range are flagged as those potentially affected by episodic emission events, and the results are used to produce a set of enhancement flag files, which contain the location and time information for these flagged daily-mean observations."

To "Values outside of the threshold range are flagged as those potentially affected by episodic emission events, and the results are used to produce a set of enhancement flag files, which contain the location and time information for these flagged daily-mean observations, along with the threshold value for the grid cell in which the enhancement arises and the daily-mean CO total column and measurement error for the anomaly event."

4. It would be relevant to include the method used to re-grid MOPITT data to 0.5 deg.

Clarified by changing L131:

From "Prior to the application of the detection algorithm, the MOPITT L2 total column CO observations are gridded to 0.5° latitude by 0.5° longitude and used to generate daily-means at this spatial resolution."

To "Prior to the application of the detection algorithm, the MOPITT L2 total column CO observations are gridded to 0.5° latitude by 0.5° longitude and used to generate daily-means at this spatial resolution. Specifically, the MOPITT data is partitioned into discrete bins at this spatial resolution using the measurement latitude and longitude, and the daily-mean for each bin is calculated as the weighted average of the data, with weights assigned as the inverse square of the retrieval error of each measurement."

Reviewer 2

The study presents statistical analysis of outliers in satellite-measured CO using the MOPITT instrument. The long term measurements from MOPITT (2000–2022) make it an ideal instrument for evaluating outliers over a long time period. The authors use level 2, thermal infrared retrievals, regridded to 0.5 degree resolution, and evaluate temporal properties for each 0.5 x 0.5 degree grid box. They use advanced statistical techniques to agnostically quantify outliers. Outliers are evaluated for persistence and strength and linked to known emission events, which are often, but not always, biomass burning. Overall the manuscript is very well written and the analysis rigorous. I have several comments to be addressed below.

Thank you for the feedback. Please see below for our responses to your comments.

1. Temporal evaluation of outliers: The authors may have missed an opportunity to temporally evaluate outliers. Specifically, it would be interesting to test whether a temporal trend exists for data in Figure 4 (the number of days flagged as affected by outliers) or alternatively compare 2000-2011 values with 2012-2022 to see if there is a change in the number of outlier days between periods.

Included the following figure and caption, as well as the following paragraph comparing the number of enhancement events between two periods (2000–2010 and 2011–2021).

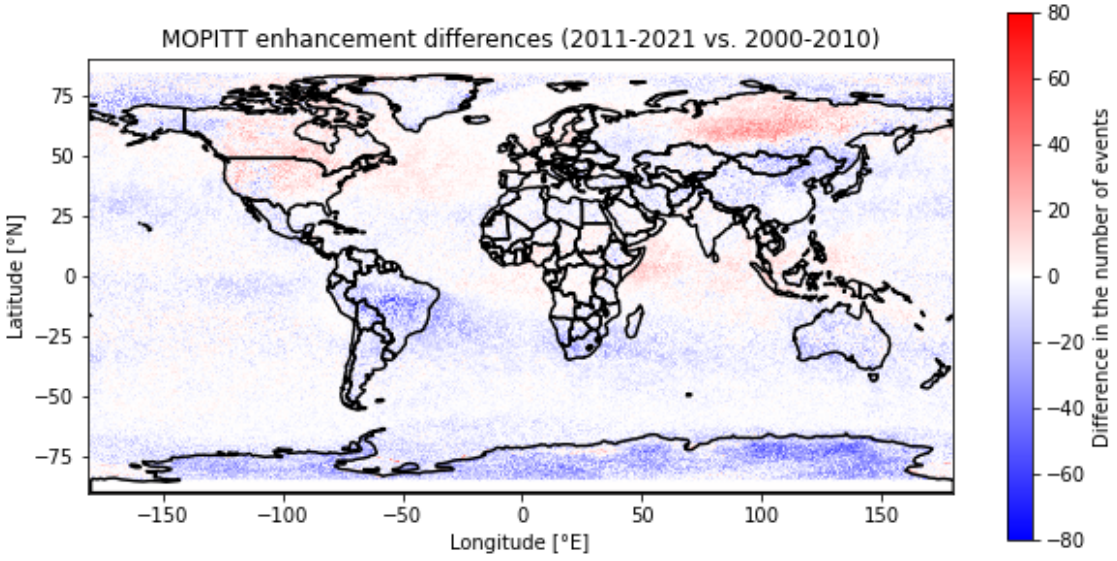

Figure 7: Differences in the number of enhancement events identified in the MOPITT L2 daily-mean measurements between two periods corresponding to the measurements made between 3 March 2000 and 31 December 2010 and those made between 1 January 2011 and 31 December 2021.

Paragrpah added after L236:

"The temporal distribution of the flagged observations in each grid cell can also be analyzed in order to identify changes in the frequency of enhancement events over time. While the sporadic nature of these enhancement events obfuscates the detection of trends when considering the evolution in the number of these events between individual years, examining multi-year periods allows for an overview of their change with time. To this end, the episodic emission events identified, shown in Fig. 4, have been separated into two periods of roughly equal length, corresponding to the MOPITT measurements made between 3 March 2000 and 31 December 2010 and those made between 1 January 2011 and 31 December 2021. The difference in the number of events in each of these two periods can then be readily calculated, as the number of enhancement events in latter period minus those in the former. The results of this are shown in Fig. 5, and from this plot several key features arise. Immediately evident is a small overall trend toward fewer enhancement events over most of the globe in the latter portion of the MOPITT measurement dataset, with the largest decreases happening over Antarctica. However, North America, Siberia, and the

eastern coast of Africa all show increasing numbers of enhancement events. Addressing first the general global decrease in the number of events, while the exact cause of this general decrease is uncertain, there is a high likelihood that this decrease is correlated with the decrease in global CO emissions over the MOPITT measurement period (Worden et al., 2013a; Hedelius et al., 2021). The global decrease in CO emissions would also likely reduce the number of observed enhancement events over Antarctica, which are transport-dependant in nature and thus strongly influenced by emissions elsewhere. In contrast to this, the regions displaying elevated numbers of enhancement events in the latter period are most likely affected by an increase in fire frequency and intensity associated with anthropogenic climate change (Dutta et al., 2016; Hart et al., 2019; Saito et al., 2022). Altogether, these findings indicate that the enhancement flags can also aid in understanding the changes in the behaviour of the CO total column over time on a global basis."

2. Comments regarding the EDF:
(i) Choice of threshold: The authors choose a threshold of 0.05 tolerance (95% confidence) to determine outliers. Were different thresholds explored? For example does a threshold of 90% confidence include more of the burning events. Does a threshold of 99% confidence reduce the number of outliers over Antarctica?

Three different threshold values were briefly tested, corresponding to 90 %, 95 %, and 99 % confidence levels, and, as expected, higher confidence levels, corresponding to lower tolerance values, led to fewer flagged enhancement events in the MOPITT dataset, and vice-versa. Both postulations from the Reviewer are correct in that a 90 % confidence level identifies more enhancement events than a higher confidence level, and a 99 % confidence level reduces the number of enhancement events over Antarctica. Desiring an intermediary between the stringency of the 99 % level, and the higher potential for CO column values near to the typical behaviour for a grid cell being identified as an enhancement event, the 95 % confidence level was chosen for this study.

(ii) Line 162: Is the EDF bimodal everywhere, or are some grid points explained by a unimodal gaussian and others by bimodal?

The empirically fit EDF is not bimodal everywhere, but there is a nearly equal amount of unimodal and bimodal EDFs, with approximately 49.7 % of the EDFs being unimodal and 50.3 % being bimodal. Both a unimodal and bimodal fit are tested for every grid cell, and the reduced χ^2 metric is used to identify the better fit, which is then used as the EDF for the grid cell.

(iii) Can anything be interpreted about the datapoints that are above the EDF, but below the threshold? For example at around residual values of -10 and between 40 and 50 in Figure 2. Are these used to determine how well the EDF fits the data?

The data points that are above the EDF, alongside all other data points, are used in the determination of the goodness-of-fit of the EDF, which is done through the calculation of a reduced χ^2 metric. This method is used to determine whether a bimodal or unimodal EDF is sufficient to describe the data without over-fitting the residuals (which might hinder the identification of enhancement events). No special focus is given to these data points beyond this.

(iv) Figure 2: Consider showing the equivalent unimodal gaussian fit as a supplement figure to demonstrate how it is inadequate.

Changed Figure 2 and its caption (shown below) to include a unimodal fit. Also changed L179:

From "This method requires an analytical estimate of the EDF for each grid cell, which is found by empirically fitting the histogram of the residual data using a unimodal and bimodal Gaussian distribution in order to account for asymmetry in the distribution. An example of such a fit is shown in Fig. 2. The reduced χ^2 metric is calculated for each fit and used to evaluate the goodness-of-fit for both fits of each grid cell. The fit with the better reduced χ^2 value is used as the estimate of the EDF."

To "This method requires an analytical estimate of the EDF for each grid cell, which is found by empirically fitting the histogram of the residual data using a unimodal and bimodal Gaussian distribution in order to account for any asymmetry or non-Gaussian features in the distribution. The reduced χ^2 metric is calculated for each fit and used to evaluate the goodness-of-fit for both fits of each grid cell. The fit with the better reduced χ^2 value is used as the estimate of the EDF. An example of a unimodal and bimodal fit are shown in Fig. 2, with the latter having been found to be the better fit, as per the reduced χ^2 metric, for the grid cell shown."

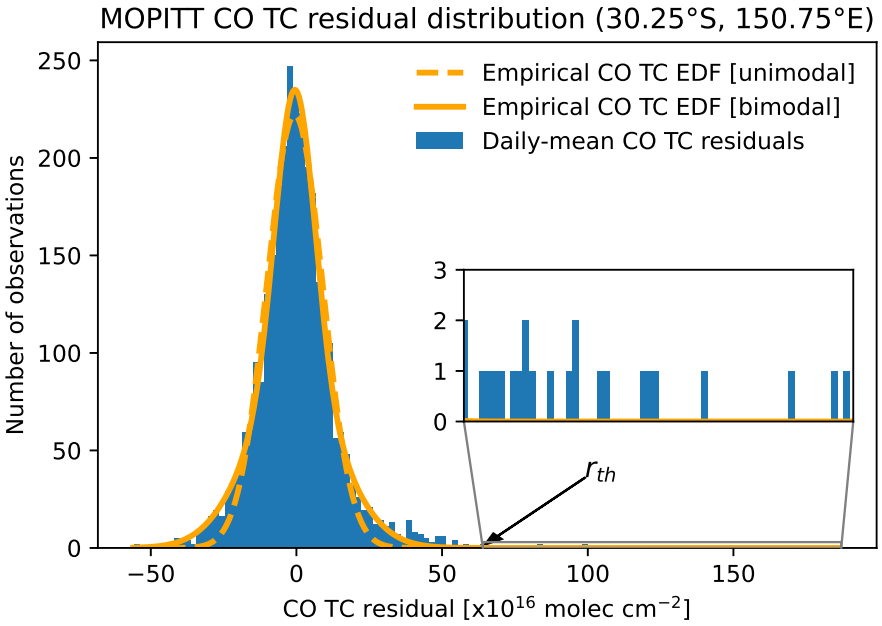

Figure 8: MOPITT daily-mean CO total column (TC) residuals, partitioned into discrete histogram bins using the Freedman–Diaconis method (blue bars), and the empirically-fitted expectation density functions (EDFs; orange lines) tested. For the grid cell shown, the interquartile range is 14.11x10^{16} molec cm^{-2} and there were 3015 observations, resulting in a histogram bin width of 1.95x10^{16} molec cm^{-2}. The residuals are fit with both a unimodal (orange dashed line) and bimodal (orange solid line) Gaussian distribution, with the latter having been found to yield the better estimate of the EDF for this grid cell as per the reduced χ^2 metric. Note the extended wings of the residual distribution that necessitate a bimodal Gaussian to properly capture the behaviour of the underlying EDF. The black arrow indicates the threshold value (r_{th}) for this grid cell, of 63x10^{16} molec cm^{-2}, and the inset shows a magnified view of the data found to be above this threshold value. Data are shown for a 0.5° by 0.5° grid cell centred over land in New South Wales, Australia (grid box centre at 30.25°S, 150.75°E), and cover the period 3 March 2000 to 31 July 2022.

3. Line 196-197: Is there evidence of atmospheric transport from the Northern Territory to NSW for the November 6, 2002 outlier? I would suggest fires in Victoria or within NSW would be much more likely to contribute to NSW outliers. Fires began in Victoria in September 2002.

Given the proximity, the 2002 Victoria wildfires seem to be a more likely contributing factor than those in the Northern Territories. Also noted and corrected a minor typographic error in the reported dates in the text here. Changed L196:

From "Outside of these days, five further observations are found to be coincident with other major Australian bushfires. These events include one on 22 January 2002, around the time of the Black Christmas bushfires, one on 6 November 2002, near the end of the 2002 Northern Territory bushfires, one on 23 January 2003, at the end of the Canberra bushfires, and a pair on 22 September 2006 and 12 January 2007 that align with the 2006–2007 Australian bushfire season."

To "Outside of these days, five further observations are found to be coincident with other major Australian bushfires. These events include one on 23 January 2002, around the time of the Black Christmas bushfires, one on 7 November 2002, near the end of the 2002 Victoria wildfires, one on 24 January 2003, at the end of the Canberra bushfires, and a pair on 23 September 2006 and 13 January 2007 that align with the 2006–2007 Australian bushfire season."

4. Paragraph starting 272: This could be due to the choice of colorbar, but I do not see evidence in Figure 5 for the patterns discussed about the Equatorial Africa region – specifically the westward transport is not evident to me in Figure 5.

The patterns focused on in discussion arise in earlier Figures, which we realize were not clearly established in text. For clarity the region of interest was better specified and the figures containing the pertinent information were specified more clearly throughout the text. Changed L254:

From "Here we focus briefly on five regions in Fig. 6 that show a majority of their flagged observations are associated with major enhancement events under multiple classification criteria. These five regions roughly correspond to northern North America (Canada and Alaska), Siberia, the Amazon (Brazil and the western Atlantic Ocean), equatorial Africa, and the equatorial Indian Ocean including Indonesia. Antarctica is excluded from discussion here, due to the ease with which transport can dramatically enhance the CO column, as discussed above."

To "Here we focus briefly on five regions in Fig. 6 that show a majority of their flagged observations are associated with major enhancement events under multiple classification criteria. These five regions roughly correspond to northern North America (Canada and Alaska), Siberia, the Amazon (Brazil and the western Atlantic Ocean), the east coast of equatorial Africa, and the equatorial Indian Ocean including Indonesia. Antarctica is excluded from discussion here, due to the ease with which transport can dramatically enhance the CO column, as discussed above."

Also changed L272:

From "Equatorial Africa shows significantly different behaviour than the prior regions. High episodic event thresholds are observed over central Africa, extending westward out over the eastern Atlantic Ocean; however, there are very few enhancement events and very low major event fractions over this region. This implies a very regular annual cycle in CO emissions with few deviations. The exception to the low major event fractions occurs off the eastern coast of Africa, an area for which these fractions display high persistence and clustering."

To "The eastern coast of equatorial Africa shows significantly different behaviour than the prior regions. High episodic event thresholds, shown in Fig. 3, are observed over central Africa, extending westward out over the eastern Atlantic Ocean; however, there are very few enhancement events, as shown in Fig. 4, and very low major event fractions over this region. This implies a very regular annual cycle in CO emissions with few deviations. The exception to the low major event fractions occurs off the eastern coast of Africa, an area for which these fractions display high persistence and clustering."

L118 to L123: Consider adding in a discussion about how additional aerosol scenes, which have been previously misidentified as clouds, allow for evaluation of more data points in the MOPITT record, especially for polluted scenes.

Added a brief explanation emphasizing this impact. Changed L118:

From "As of V9, the criteria used to identify clear-sky conditions have been relaxed, which has led to significantly enhanced coverage of global CO. "

To "As of V9, the criteria used to identify clear-sky conditions have been relaxed, which has led to significantly enhanced coverage of global CO. This is particularly relevant in regions with heavy pollution, including those areas affected by large biomass burning events, as the aerosols in these scenes were frequently misidentified as clouds in previous versions and filtered from the MOPITT data record. "

L138: Please clarify the 15 day moving window multi-year average. Is it a 15-day moving-window climatology, 2000–2022?

Yes, it is a 15-day moving-window climatology. We reworded the text to better reflect this. Changed L137:

From "To this end, in each grid cell, the multi-year centred-moving-average was calculated for each day of the year using a 15-day window centred on each day in turn."

To "To this end, in each grid cell, a climatological multi-year centred-moving-average was calculated for each day of the year using a 15-day moving window centred on each day in turn."

Figure 1 caption: Specify the regression fit is to account for ENSO influence.

Changed the Figure 1 caption:

From "Time series of MOPITT daily-mean total column (TC) CO time series (in 10^{16} molec cm^{-2}) between 3 March 2000 to 31 July 2022 (top), as well as the deseasonalized time series and the regression fit of the deseasonalized time series (middle; blue dots for the time series, orange line for the fit). Fit coefficients were determined for this grid cell as a_0=228.7, a_t=0.11, and a_{MEI}=2.74. The residuals from this fit, used to identify episodic emission events, are shown in the lower panel (bottom). Data are shown for a 0.5° by 0.5° grid cell centred over land in New South Wales, Australia (grid box centre at 30.25°S, 150.75°E)."

To "Time series of MOPITT daily-mean total column (TC) CO time series (in 10^{16} molec cm^{-2}) between 3 March 2000 to 31 July 2022 (top), as well as the deseasonalized time series and the regression fit of the deseasonalized time series (middle; blue dots for the time series, orange line for the fit). The MEI term in the regression is included to account for the influence of the ENSO. Fit coefficients were determined for this grid cell as a_0=228.7, a_t=0.11, and a_{MEI}=2.74. The residuals from this fit, used to identify episodic emission events, are shown in the lower panel (bottom). Data are shown for a 0.5° by 0.5° grid cell centred over land in New South Wales, Australia (grid box centre at 30.25°S, 150.75°E)."

Figure 2 caption: add in (r_{th}) after "threshold value".

Changed the Figure 2 caption:

From "The black arrow indicates the threshold value for this grid cell, of 63x10^{16} molec cm^{-2}, and the inset shows a magnified view of the data found to be above this threshold value."

To "The black arrow indicates the threshold value (r_{th}) for this grid cell, of 63x10^{16} molec cm^{-2}, and the inset shows a magnified view of the data found to be above this threshold value."

L159: I suggest to provide an example of the interquartile range for the data in Figure 2.

Added a sentence addressing this in the caption for Figure 2. Along with other suggestions from above, changed the Figure 2 caption:

From "MOPITT daily-mean CO total column (TC) residuals, partitioned into discrete histogram bins using the Freedman–Diaconis method (blue bars), and the empirically-fitted expectation density function (EDF; orange line). The black arrow indicates the threshold value for this grid cell, of 63x10^{16} molec cm^{-2}, and the inset shows a magnified view of the data found to be above this threshold value. Data are shown for a 0.5° by 0.5° grid cell centred over land in New South Wales, Australia (grid box centre at 30.25°S, 150.75°E), and cover the period 3 March 2000 to 31 July 2022. Note the extended wings of the residual distribution that necessitate a bimodal Gaussian to properly capture the behaviour of the underlying EDF."

To "MOPITT daily-mean CO total column (TC) residuals, partitioned into discrete histogram bins using the Freedman–Diaconis method (blue bars), and the empirically-fitted expectation density functions (EDFs; orange lines) tested. For the grid cell shown, the interquartile range is 14.11x10^{16} molec cm^{-2} and there were 3015 observations, resulting in a histogram bin width of 1.95x10^{16} molec cm^{-2}. The residuals are fit with both a unimodal (orange dashed line) and bimodal (orange solid line) Gaussian distribution, with the latter having been found to yield the better estimate of the EDF for this grid cell as per the reduced χ^2 metric. Note the extended wings of the residual distribution that necessitate a bimodal Gaussian to properly capture the behaviour of the underlying EDF. The black arrow indicates the threshold value (r_{th}) for this grid cell, of 63x10^{16} molec cm^{-2}, and the inset shows a magnified view of the data found to be above this threshold value. Data are shown for a 0.5° by 0.5° grid cell centred over land in New South Wales, Australia (grid box centre at 30.25°S, 150.75°E), and cover the period 3 March 2000 to 31 July 2022. "

L160: Clarify that data for each 0.5 degree grid point will have a different bin width.

Added a sentence addressing this. Changed L159:

From "... where IQR(r) is the interquartile range of the residual data r and N is the number of observations (Freedman and Diaconis, 1981). This method is employed as it minimizes the difference between the generated histogram and the shape of the theoretical probability density function (PDF) that underlies the data (Freedman and Diaconis, 1981)."

To "... where IQR(r) is the interquartile range of the residual data r and N is the number of observations (Freedman and Diaconis, 1981). As a result of the differences in the interquartile range and the number of observations available in each 0.5° by 0.5° grid cell, the histogram bin width of each grid cell varies. This method is employed as it minimizes the difference between the generated histogram and the shape of the theoretical probability density function (PDF) that underlies the data (Freedman and Diaconis, 1981)."

L217 to L224: How were the sources that contributed to the observed threshold values determined or identified?

These four regions were singled out in the discussion of the threshold values due to their particularly high values. To the authors, who have worked extensively with the MOPITT CO data products since the launch of MOPITT, the regular large CO emissions from these regions are a well-known quantity, and the sources of these emissions listed (Amazon wildfires/deforestation, Indonesian biomass burning, sub-Saharan agricultural fires, and industrial activity in China) are considered by the authors to be the most prominent sources of regular CO emissions in each region.