Referee comment on "Seasonal, regional and vertical characteristics of high carbon monoxide plumes along with their associated ozone anomalies as seen by IAGOS between 2002 and 2019" by Lebourgeois et al.

This manuscript provides a statistical analysis of extreme CO values of IAGOS database for different regions, seasons and vertical layers (lower/middle/high troposphere), both in terms of origin (using SOFT-IO software) and in terms of impact on O3 production.

I found the manuscript to be well organised and clearly written. The methods used are scientifically sound and the figures chosen appropriately support the discussion and conclusions. The results provide an important overview of the CO plumes observed over 18 years of in situ measurements.

However, I find that there could be more links to the main processes involved, as well as a fuller discussion of how the results align with recent literature, including key publications using the IAGOS datasets that are cited in the article (but not only).

A discussion of how representative the in situ data used is of each major region should also be added, since I don't think that the data are randomly distributed within each region/layer.

General comments

1. Introduction.

I don't really see the relevance of the paragraph on satellite observations to this paper. However, a more precise description of previous results using CO and O3 observation datasets (in particular IAGOS) would be useful to better put the main results of this paper into perspective (in introduction and to discuss the results).

2. Methods

2.2: The description of the SOFT-IO software should include a paragraph on performance and uncertainties (in emission inventories and attribution using back trajectories). The important warning in § 122-127 could be written more clearly, and discussion of the performance in attribution in past studies could be helpful.

2.3.1: The data is divided into large regions for analysis. It would be important to discuss the location of the data analysed within each region and for each layer. If I understand the data correctly, the profiles correspond to specific airports and the flight paths also follow specific routes. I think it would be important to better discuss the representativeness of the data set for each region/layer.

2.3.2: For the type of source attributed to the anomalies detected (l.175-180) , why did you choose to use the main characteristic and why not add the fractional contributions? Is it because of too large uncertainties in attribution with SOFT-IO?

3. Results

For all regions and layers, I was wondering what fraction of the detected anomalies are successfully attributed to a main source using SOFT-IO?

The seasonal variations in the LT are mainly attributed to variations in the local sources. Does that mean that seasonal variations in the background levels has little impact? Even if anomalies are selected, the final concentration is an enhancement above background, both for CO and for O3.

l. 207: The authors mention “a cycle of O3 destruction in CO-rich air masses”: O3 is then lower than background in the corresponding area? It would be helpful to add some detail on the corresponding chemical processes (same comment for all regions). What could be the impact of other co-emitted compound such as aerosols? I understand that this is beyond of the scope of this paper but for each region, it would be helpful to have some reference to the literature on the subject.

As mentioned in the general comments, it would be important to better discuss the results obtained in light of the literature, in terms of source contributions but also in terms of O3 enhancement in CO enriched air masses.

Although carefully conducted, the analysis reads a bit like a list. Perhaps a summary scatterplot of O3 versus CO could be used to get a more general view of the data set? A colour code could be used, for example, to differentiate regions / layers, etc.

Specific comments

l. 51: Other O3 precursors have a long lifetime, CH4 for instance.

l. 101: The paragraph on SOFT-IO should be included in the next section which is dedicated to the software.

l. 213: need reference for the larger number of convective events during summer.

l. 218: ‘increased number of episodes…’ increased compared to what?

l. 244-245: again, increased compared to? Large increase in East Asia vs Siberia are attributed to quite different sources. What processes are at play? What transport pathways? Does that hold if not only the main features are kept, but the fractional contribution from different sources? Is the situation still that contrasted?

l. 259-260: “which is probably due to the higher emission height…” Could this statement be checked? I agree that injection heights may be important here. What height is considered in SOFT-IO for source attributions?

l. 315: Why is O3 particularly low in BB plumes for this region? Has this event (2015) been analysed in past publications (even using other methods)?

In fact I also have the same question for other regions, such as African/ME BL, etc.

l. 367: Can BB be called ‘wildfires’ in this region? I would think there is significant contribution from agricultural burning as well. The use of the term ‘wildfire’ should probably be reviewed throughout the manuscript.

l. 462: same phrases repeated.

All figures for the statistical analyses: mention the total number of points in the subsets.

Figures O3 (Fig 4, etc): What do colored dots represent? Average value? Why not show the full boxplot for each source? Not enough data?