Letter to the Editor

Tropical tropospheric ozone and carbon monoxide distributions: characteristics, origins and control factors, as seen by IAGOS and IASI

Maria Tsivlidou et al.

In the second round of peer-review, I received strongly contrasting reviewer report and thus asked a third reviewer for their input. The third reviewer points out that, while they do not share the general concerns of reviewer 1 regarding the methodology, they conclude that major revisions are needed before the manuscript can be published with ACP. Thus, I encourage the authors to perform another major revision of the manuscript, considering all three new reviewer reports and focusing on presentation quality of the article. The paper will be accepted if the scientific results and conclusions are presented in a clear, concise, and well-structured way.

Dear Editor,

we have carefully addressed in the revised version the comments from Referee#1 and the new Referee#3.

We have also responded to the 3 points raised in Referee#2's second assessment (Report#1), although we strongly disagree with his/her comments which are not at all unbiased. Referee first recognizes that we have considered his/her first comments but he/she objects to new ones to reject the publication. Furthermore, Referee focuses his/her new comments on the less concluding details of the paper disregarding most of the results which are concluding. Finally, his/her judgment is based on an incorrect assessment of the Figures with no explanations.

- First, Referee keeps on repeating that the model only reproduces 10% of the anomalies when this is absolutely not the case. The paper shows that the model reproduces the vast majority of anomalies. As explained in our answer, we do not understand why the Referee argues on supposed large underestimation by SOFT-IO. For the 20 clusters and in the 3 vertical layers and the 4 months, SOFT-IO reproduces on average 93,5% (relative difference) of the observed CO. The underestimation is slightly larger on average in the LT, but still SOFT-IO reproduces 87% of the anomalies. The largest underestimation is observed over the Sahel cluster in the LT in January, with 52,8% of the anomaly reproduced by the model, still far from the 10% claimed by Referee. The only occurrence of a 90% bias is over Mumbai in January in the 40hPa surface layer (see the answer to Referee#2 for details about biases in the high resolution profiles). Nevertheless, first our conclusions and main results are all based on the broad LT, MT and UT layers for which the model results are robust as explained above and second, there are only 2 profiles for Mumbai in January which makes this bias statistically irrelevant. This fact was already acknowledged with a warning in the first version of the paper. Therefore, the repeated mention to this accidental 10% by Referee#2 is far from fair. Then, Referee#2 suggests that we deliberately hide the worst results. This is not true, the figures presented in the paper and in the Appendix are rather exhaustive, even too much according to Referee#3. The underestimations of the model were clearly explained in Section 2.2 lines 164-166 and lines 172-175 of the 1st round revised version. We give more precise numbers of SOFT-IO limitations in this 2nd round revised version. Imperfections and biases are a common feature of models such as SOFT-IO and we did not try to hide it on the contrary. We would like to emphasize that such a model evaluation is useful for a large community using emission inventories. Referee#2 then argues that we state that our model "does a poor job" but that we still want to use it. This appears contradictory to his/her allegation that we aimed at hiding our model limitations. Moreover, we never said that our model "does a poor job", we give an objective and fair scientific statement that all modelers know, models are imperfect by definition, because they use inputs (emission inventories, meteorological fields) and schemes (chemical schemes, convective parameterization, turbulence) that are all perfectible. Our model does a good job, contrary to what is claimed by Referee#2 without any scientific argumentation, and Figures and numbers provided by SOFT-IO speak by themselves. The version of SOFT-IO we use is based on a state of the art Lagrangian model (FLEXPART last version); a state of the art biomass burning inventory (GFAS last version) which is the only that provides biomass burning injection height and that is operational; a state of the art anthropogenic emissions (CEDS2); and state of the art meteorological analysis and forecast (ECMWF). All of them are separately or together largely used in hundreds of peer-reviewed publications of high level scientific journals. If we follow his/her sarcastic statement, no publications in atmospheric science based on a model should be published, as all models and input or subparts are impacted by uncertainties.

- Finally, Referee#2 argues that our conclusions are biased by the emission inventories we use based on a preprint (Wiedinmyer et al. in discussion, since Feb 2023) not yet accepted and which, as such, could absolutely not be known and used by the authors at the time of submission. This is not a fair and correct way of reviewing a paper.
  Our study is based on the analysis of two datasets (IASI and IAGOS) and using the best SOFT-IO version at that time, and we never claimed that our conclusions were set in stone. The large majority of model studies in atmospheric science use one state of the art biomass burning inventory, and one state of the art anthropogenic emission inventory, and conclusions are obviously related to this particular choice of inventories, but also of meteorological fields, transport schemes, chemical schemes, version of the model etc. This is the same for all studies using models in a given configuration, so his argument is inadmissible. We have revised the conclusions when fire and anthropogenic contributions are of the same order, as the other inventories could change the results, but not when fire or anthropogenic very largely dominates.

In addition and most importantly, we would like to point out that this second review of Referee#2 does not comply at all with the reviewers' code of conduct in many respects. In particular, "The reviewer of a manuscript must judge objectively the quality of the manuscript and respect the intellectual independence of the authors. Under no circumstances is personal criticism appropriate.” (see https://www.atmospheric-chemistry-and-physics.net/policies/obligations_for_referees.html)

Many comments of Referee#2 are indeed ironic, insulting and inappropriate in a scientific discussion such as “it borders on intellectual dishonesty” or “I believe these statements speak for themselves” or making fun of the comparison we did between CEDS2 and MACCITY (the last one was largely used by the community at the time of the study) without any scientific argumentation.

There seems to be an obvious conflict of interest with a strong will to prevent the publication of our work. Such behaviour should not be allowed in a publication of the rank of ACP.

Once again, we express our gratitude for the thorough review process and the comments of Referee#1, Referee#2 and Referee #3, which has helped strengthen our manuscript. We are confident that with the suggested modifications, our paper will be well--positioned for publication in the ACP journal.

Thank you for your time and consideration. We look forward to hearing from you regarding the final decision on our manuscript.

Sincerely,

Maria Tsivlidou

Answer to Referee#2 (blue in the text)

2nd Review of Tsivlidou et al. “Tropical tropospheric ozone and carbon monoxide distributions: characteristics, origins and control factors, as seen by IAGOS and IASI”.

I would like to commend the authors for incorporating most of the comments that were provided after the first round of review. I believe the manuscript in the present form is easier to read and provide some needed justifications for some of the methodological choices.

We thank Referee for his/her comments, indeed we provided substantial answers to all his concerns expressed in the first review.

Unfortunately, I remain largely unconvinced by some of the responses:

i) SOFTIO clearly struggles at reproducing the magnitude of the CO anomalies in all clusters of the tropical band. On some occasions, as little as 10% of the CO mixing ratios are accounted for by the model. If so little of the CO anomalies is explained by SOFTIO, how can the authors conclude on the sources of those anomalies? Let me try to be clearer here: if SOFTIO can only represent 10% of the

CO anomaly at a given cluster, then the source attribution is only valid for those 10%. The remaining 90% are unaccounted for, and this should be clearly highlighted in the paper. This is the reason why in my original review, I had asked the authors to show on their figures how much of the CO anomalies are NOT accounted for by SOFTIO. Otherwise, and it may not be the intention of the authors, it borders on intellectual dishonesty

We do not agree with this comment. Referee focuses on “some occasions” where SOFT-IO would reproduce 10% of the anomalies but this does not happen.

We provide below quantitative arguments that demonstrate that Referee repeated statement about 10% is wrong and unfair:

- SOFT-IO reproduces remarkably well the shape of the CO vertical profiles and also their seasonal variability as shown in the profile Figs. 3, 5, 6 and 7.
- SOFT-IO performances on anomaly frequencies:

When focusing on the strongest CO anomalies (higher than the 75th percentile of IAGOS measurements), which are supposed to be less simulated by the model, we clearly see that SOFT-IO detects the strongest anomalies 85-90% (in frequency) of the time (Table 1 below). Lowest values are observed over Addis Ababa and Bogota in the LT because these sites are located above 2 km altitude, with few measurements in the LT layer (below 750 hPa).

- SOFT-IO performances on anomaly amplitudes:

Table 2 displays the absolute (left) and relative (right) differences between IAGOS and SOFT-IO anomalies. To be comparable, background has to be added to SOFT-IO (or subtracted to IAGOS measurements). The relative difference is then calculated as (IAGOS-(SOFT-IO+background)) / IAGOS, and allows to compare observed and simulated anomalies in the 3 different layers, for all the clusters, averaged for the 4 months.

On average for the 3 layers, SOFT-IO reproduces more than 93% of the amplitude of observed anomalies, and the largest underestimation occurs in the LT with 87%.

When looking at the LT layer where observed anomalies are the strongest and where model performances are supposed to be the lowest, January and April are the months with the lowest performances, with respectively 82,4% and 78,5% of the anomalies reproduced in average for the 20 clusters, and a standard deviation of respectively 20% and 13%.

Looking at individual clusters, months and layers, the 2 lowest performances occur in the LT over Sahel and Mumbai in January, with respectively 52.8% and 53,1% of the amplitude of the anomalies reproduced by SOFT-IO, which is far from 10% .

Of course, the profiles display larger differences especially close to the surface in polluted areas where large scale models have more difficulties to represent anomalies resulting from local contributions. That is why our conclusions are based upon the results concerning the three broad layers representing the LT, MT and UT. Furthermore, profile Figs. 3, 5, 6 and 7 with the x-scale extended to the largest CO mixing ratio for all months and all airports from a region probably induced the wrong idea that the model does a “poor job” reproducing the anomalies. Nevertheless, if we look at the same relative differences for the high resolution profiles than for the 3 broader layers (discussed above), above the surface level the model represents in majority the CO anomalies at 40 hPa vertical resolution with a bias lower than 40% (see Figure 1 below for Bangkok). Looking at the surface 40 hPa level, the bias reaches 80% only on 4 occasions: at Sahel in January, Central Africa in April or Lagos in April and October (see Figure 2 below for Lagos).
The ONLY case with 90% bias is at the SURFACE level over MUMBAI in JANUARY. But it is based on a low statistics (2 profiles in January) and a warning about this problem is already given in the paper: “It has to be noted that the number of profiles over Mumbai (6) and Hyderabad (19) are lower than the threshold established for representativeness (see Sect. 2.1.1).”

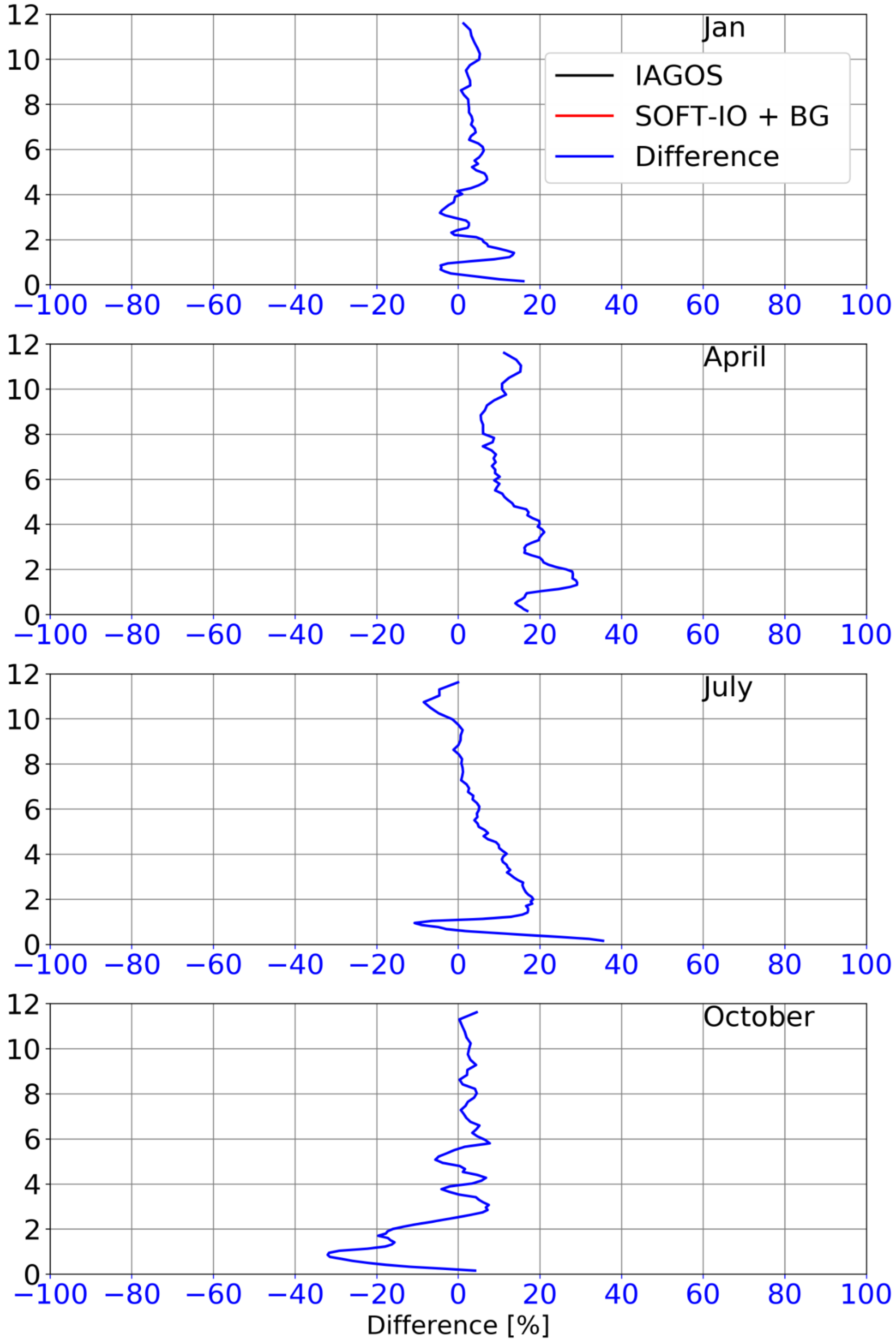

Figure 1: Vertical profile of relative difference between IAGOS and SOFT-IO anomalies (IAGOS- (SOFT-IO + BG) / IAGOS) averaged for January, April, July and October over **Bangkok**.

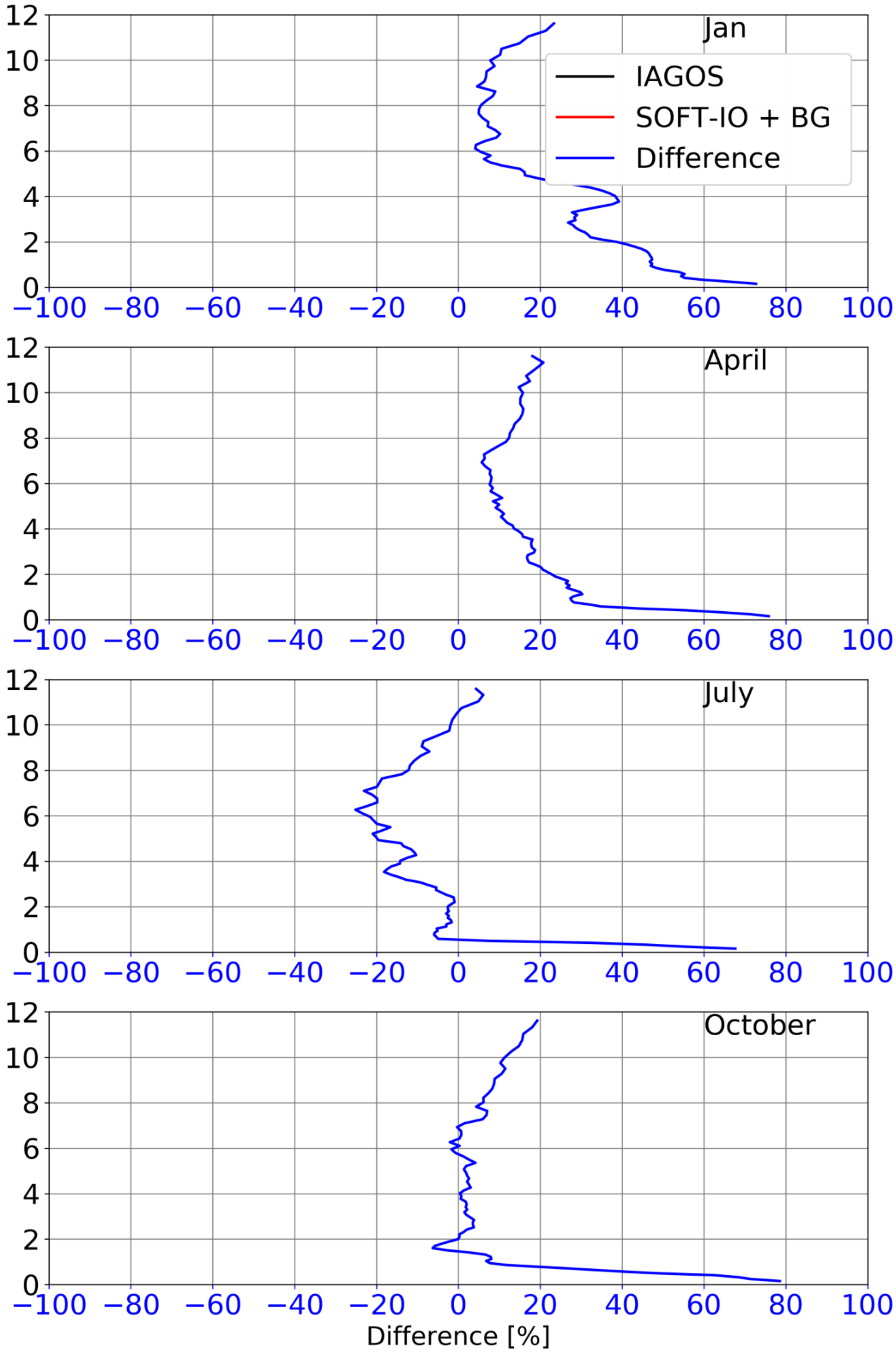

Figure 2: Same as Figure 1 for **Lagos**.

SOFT-IO's general underestimation (in ppb) is already discussed in the manuscript (Section 2.2, lines 161-163 and lines and 169-170), and the comparison between IAGOS and SOFT-IO (without the background that has to be added on SOFT-IO plots for correct comparisons, as explained in Section 2.2) is clear enough on Figs 3, 5, 6 and 7. The paper has already been commented as too dense with too many figures so we think it not appropriate to add some more.

We have added additional percentages of absolute and relative differences between IAGOS and SOFT-IO for the strongest anomalies in the text of Section 2.2, lines 170-174 of the 2nd round revised version.

	Accounted anomalies (%)		
	LT	MT	UT
Abu Dhabi	93.72	84.67	75
Thailand gulf	98.57	92.94	89.62
Central Africa	96.72	93.97	94.5
Sahel	95.38	93.29	92.94
South China	99.04	94.49	94.29
Gulf of Guinea	99.45	97.87	96.64
SBrazil	94.93	85.05	85.28
Addis Ababa	54.65	90.05	85.49
Bangkok	98.71	94.99	94.52
Bogota	1.23	89.84	88.34
Caracas	97.1	83.53	78.16
Ho Chi Minh City	98.01	88.94	88.32
Hyderabad	94.69	94.17	91.69
Jeddah	95.54	88.14	86.82
Khartoum	90.63	82.27	84.81
Lagos	99.6	96.57	96.31
Madras	98.4	89.93	90.06
Manila	96.93	85.47	84.41
Mumbai	97.72	95.52	92.35
Windhoek	90.52	86.12	88.27

Table 1: Detection frequency of the CO anomalies higher than the 75th percentile of the IAGOS observations.

	Absolute difference (ppbv) IAGOS - (SOFT-IO + BG)			**Relative difference (%)** (IAGOS - (SOFT-IO + BG))/ IAGOS		
	LT	MT	UT	LT	MT	UT
Abu Dhabi	5.75	-0.89	-4.18	3.18	-1.59	-4.04
Thailand gulf	-7.71	-20.74	7.43	0.97	-5.63	2.59
Central Africa	58.79	16.29	20.37	25.32	9.92	15.46
Sahel	39.88	3.56	5.20	10.11	-1.26	2.51
South China	-10.00	-1.73	0.77	-7.23	-1.97	-1.78
Guinea Gulf	43.40	10.89	9.50	12.64	4.90	6.11
SBrazil	0.42	1.33	20.92	-4.76	-0.71	12.62
Addis Ababa	92.53	-0.23	1.21	19.36	-3.54	-0.74
Bangkok	11.68	6.79	2.66	2.50	4.94	0.60
Bogota		43.38	16.08		27.14	13.60
Caracas	49.80	10.22	6.45	28.11	9.39	6.00
Ho Chi Minh City	94.07	-0.71	-2.64	29.75	-1.28	-2.42
Hyderabad	4.52	2.85	1.45	2.56	0.56	-0.14
Jeddah	20.40	6.02	2.26	14.63	4.91	0.04
Khartoum	29.25	2.88	2.89	19.21	1.71	1.50
Lagos	127.52	10.24	13.09	28.78	3.61	8.15
Madras	18.05	2.58	-1.67	9.95	0.61	-1.99
Manila	41.85	0.03	-0.08	22.13	-0.83	-2.19
Mumbai	43.95	4.08	2.65	24.23	3.60	1.56
Windhoek	10.43	10.67	14.17	6.63	8.10	11.38

Table 2: Absolute and relative difference between the observed (IAGOS) and simulated (SOFT-IO) anomalies averaged for 3 vertical layers over the 4 months (January, April, July, October).

ii) In addition, I don't understand the authors response stating that "For instance, models are persistently biased in the Southern hemisphere and in the tropics, particularly over polluted regions such as India and East Asia. As a result SOFT-IO has to be seen as a tool to perform source attribution and to quantify the relative part of a source influence to another, but not as a tool perfectly able to simulate the exact CO concentrations, but this is a problem of most of the models in

CO anomalies." To me, there are two serious problems with this response. If I was to rephrase it in a simpler way, it reads as "we know our model does a poor job at reproducing CO mixing ratios, but we are going to do it anyways because all models do equally poorly" and as "we can't reproduce CO mixing ratios, but we are still going to apportion sources contribution and disregard the remaining CO not reproduced by the model". I believe these statements speak for themselves.

Like any atmospheric model, SOFT-IO is not perfect as by definition, a model is a simplification of the true atmospheric processes (dynamic, chemistry), based on perfectible inputs (emission inventories, meteorological analysis) or subparts (chemical schemes, transport schemes). However our model does a very good job, as demonstrated in the reference paper of SOFT-IO (Sauvage et al., 2017), and also in this current study where more than 78% of the strongest anomalies are simulated (see Table 1). SOFT-IO uses a state of the art Lagrangian model (FLEXPART), state of the art meteorological ECMWF analysis and state of the art emissions inventories (GFAS and CEDS2 last versions), all largely used by the community and published in high standard peer reviewed journals. GFAS is for instance used in the Copernicus Atmosphere Monitoring Service (CAMS) model simulations (https://atmosphere.copernicus.eu/) (e.g. Flemming et al., 2017; Inness et al., 2019). It is operational and provides fire injection altitude. CEDS2 is used by some of the models included in the IPCC reports (e.g. Emmons et al., 2020; Horowitz et al., 2020; Griffiths et al. 2020).

Emmons, L. K., Schwantes, R. H.,Orlando, J. J., Tyndall, G., Kinnison, D.,Lamarque, J.-F., et al. (2020). The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2). Journal of Advances In Modeling Earth Systems,12,e2019MS001882. https://doi.org/10.1029/2019MS001882

Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015, Atmos. Chem. Phys., 17, 1945–1983, https://doi.org/10.5194/acp-17-1945-2017, 2017.

Horowitz, L. W., Naik, V., Paulot, F.,Ginoux, P. A., Dunne, J. P., Mao, J.,et al. (2020). The GFDL global atmospheric chemistry-climate model AM4.1: Model description and simulation characteristics.Journal ofAdvances in Modeling Earth Systems,12, e2019MS002032. https://doi.org/10.1029/2019MS002032

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition,

Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Griffiths, P. T., Murray, L. T., Zeng, G., Shin, Y. M., Abraham, N. L., Archibald, A. T., Deushi, M., Emmons, L. K., Galbally, I. E., Hassler, B., Horowitz, L. W., Keeble, J., Liu, J., Moeini, O., Naik, V., O'Connor, F. M., Oshima, N., Tarasick, D., Tilmes, S., Turnock, S. T., Wild, O., Young, P. J., and Zanis, P.: Tropospheric ozone in CMIP6 simulations, Atmos. Chem. Phys., 21, 4187–4218, https://doi.org/10.5194/acp-21-4187-2021, 2021.

iii) The fact that SOFTIO struggles to reproduce CO mixing ratios indicate that at least one or both the AN and BB emission inventories used in this study severely underestimate CO emissions in the tropics. The authors keep claiming that this is mostly due to the AN emission inventory, but with no scientific evidence for it. Their first argument is that "Furthermore, we comment the performance of SOFT-IO when the CO anomalies are attributed entirely to one source (AN) and to one source region. For instance, in the case of Africa (NH and SH) (line 261, page 14 and line 316, page 16 respectively in the original manuscript) and South America (line 445, page 21 of the original manuscript), we discuss the underestimation of the AN emissions during the transition periods, when the fires are suppressed." Looking at their Figure S1 showing GFAS estimations of BB, clearly there is fire activity in Africa that would affect the clusters in April. Their response is in direct contradiction with the data they show. Their second argument is that they performed a sensitivity analysis on emission inventories and found that AN emission inventories weighted more than BB emission inventories. Not surprising: the authors used GFED and GFAS as BB emission inventories, which are extremely similar to one another. On the other hand, they compared CEDS and MACCity for AN emission inventories, which comes down to comparing a Rolls Royce with a Toyota Corolla (and this is meant with no disrespect for Toyota). Of course, they would find a higher sensitivity of SOFTIO results to AN emission inventories. Now, what if the authors used the latest BB emission inventory, FINNv2.5? In their recent paper, Wiedinmyer et al. (2023) clearly show that both GFED and GFAS are lower by a factor of almost two compared to FINNv2.5 or FEER for CO emissions (see their Figure 4 pasted below).

We disagree, as demonstrated before SOFT-IO does not struggle to reproduce CO anomalies at all! This is just the Referee biased vision. Also we only claim that AN are likely to be underestimated during the month when BB are at their minimum intensity in April. Over Northern Hemisphere Africa, BB emissions are much lower in April, with 1.02e-10 kg/m2/s on average according to GFAS, versus 3.47e-10 kg/m2/s during the fire season. This is also clearly visible on MODIS fire counts (for instance Fig.7 of Yamasoe et al., 2015 Tellus-b paper, cited in our study). During this specific month, AN contributes to more than 85% of the observed anomalies in the LT of West Africa, and as BB is minimum, AN are very likely to be underestimated. We have clarified this statement in the revised version (see more details at the end of this answer)

In addition, Referee cites a preprint paper (Wiedinmyer et al. in discussion, since Feb 2023) that was not even under discussion when we submitted our manuscript (28 September 2022). This is unfair and irrelevant.

Referee also makes fun of comparisons we did between some anthropogenic emissions inventories (CEDS and MACcity), but at the time of the reference Sauvage et al., 2017 paper and of the current work of our study (Tsivilidou et al.), these emission inventories were state of the art and largely used by the scientific community.

As a result we use state of the art emission inventories, GFAS and CEDS2, that are largely used in many published studies. The majority of the model studies that perform chemistry and transport modeling uses one state of the art BB inventory, same for anthropogenic, or for biogenics, or other sources, depending on the study, and does not provide sensitivity test between BB or anthropogenic emission inventories such as we did in Sauvage et al., 2017 or Tsivlidou et al. current study.

In this Wiedinmyer et al., 2023 paper (currently under revision), Referee just takes one figure stating that Finn v2.5 is larger than GFAS. This is true but on a global and yearly average (their Fig. 4). However there are important regional differences over the tropics in Wiedinmyer et al., 2023 (their Fig. 5) not mentioned by Referee: FINN is larger than GFAS over Africa (NHAF and SHAF), but GFAS is larger than FINN over Asia (EQAS), and there is no proof with this study that FINN or GFAS or another one may be a better emission inventory. There are also differences with other state of the art biomass burning emission inventories not mentioned by Referee (with QFED or FEER). Moreover, the only validation of FINN is through a chemical transport model (CAM chem) by comparing CO simulated tropospheric columns to the ones observed by MOPITT, just for the 2018 year or August 2018. There is no similar comparison in Wiedinmyer et al., 2023 paper using GFAS with the CAM chem model, or using different anthropogenic emission inventories (that also shows large differences) to see which emission inventory would best fit the total tropospheric CO columns.

In our study, we compare SOFT-IO using GFAS and CEDS2 using 17 years of in situ IAGOS high resolution vertical profile observations.In addition to the fact that GFAS is a reference in the community, as others emission inventories, we use GFAS because:

- It is an operational emission inventory that fits the near real time SOFT-IO calculations that we provide for the IAGOS users, which is not the case for other biomass burning emission inventories. GFAS is also used in the CAMS global model simulations.

- GFAS also provides biomass burning injection heights which are important for better simulations of biomass burning plumes (see sensitivity on fire injection in Sauvage et al., 2017). This is not the case of other biomass burning inventories.

All the conclusions we make in the paper concerning the origin of the anomalies are obviously dependent on the emission inventories, the Lagrangian model and the meteorological analysis, as it is for all conclusions done in atmospheric science using models that will depend on the model configuration, version and input parameters.

In the 2nd round revised version, we have revised the affirmations and conclusions (see below for details) when fire and anthropogenic contributions are of the same order, as the influence of other inventories (such as FINNv2.5) could change the results, but not when fire or anthropogenic very largely dominate the other one.
We have modified Section 2.2 lines 172-175 of first round revised version, by adding additional information of SOFT-IO performances:
"On average, SOFT-IO underestimates the observed CO anomalies by 10 ppb in the MT and UT, and by 45 ppb in the LT. A sensitivity test has shown absolute differences of 27% in the LT, 16% in the MT and 10% in the UT between SOFT-IO simulations using AN emissions from MACCity and from CEDS2. This clearly highlights the large uncertainty stemming from uncertainties in AN emissions."
by the following:
"On average, SOFT-IO underestimates the observed CO anomalies by 10 ppb in the MT and UT, and by 45 ppb in the LT. When looking at the differences between IAGOS and SOFT-IO anomalies (taking into account the background not simulated by the model), SOFT-IO reproduces on average more than 93% of the observed anomalies for the 3 layers and 87% in the LT. April is the month where the model gives the largest underestimation, with 78% of the anomalies simulated on average over the 20 clusters (13% standard deviation). The lowest performance is in the LT of Sahel in January, with 52% of the anomaly simulated by SOFT-IO." lines 169-174 of the 2nd round revised version.

We also rephrase: "when the fires are suppressed" line 244 of the 1st round revised version by "when the fires are reduced" line 255 of the 2nd round revised version.

We have suppressed the sentences "The fact that SOFT-IO attributes approximately 80 ppbv of CO to local AN emissions (Figs. 3 panel 3b; A1 panels 1b and 2b), while the observed anomaly reaches 200–250 ppbv and no or few fires are detected by MODIS (Yamasoe et al. (2015); their Fig.7), indicates underestimation of the Northern Hemisphere African AN emissions" lines 245-248
and
"The measured CO maxima reaches 350 ppb, while SOFT-IO attributes 40 ppb to the aforementioned sources. This means that Southern Hemisphere African AN

emissions are likely underestimated” lines 281-283 of the 1st round revised version by the following (lines 258-261 of the 2nd round revised version):
“The annual CO surface maximum in Central Africa occurs also in April. LT CO is attributed by more than 85% to local AN emissions (Fig. 3 panel 4b and Fig. 4a) when few fires are detected by MODIS (Yamasoe et al. (2015); their Fig.7). The measured CO anomaly reaches +200 ppb (after removing the background from the observed CO), while SOFT-IO attributes 40 to 80 ppb to AN.This indicates that African AN emissions might be underestimated”.

Answer to Referee#3 (blue in the text)

General comments

Tsivlidou et al. present a valuable study adding to the understanding of CO/O3 distributions in the tropics. I very much appreciate the efforts the authors have gone through in their data analysis. I think the trajectory approach in the SOFT-IO model to disentangle the contribution of wildfire and anthropogenic emissions to the observed signals is sound and sufficiently well documented by Sauvage et al. (2017) to be used here.

However, I think that the manuscript is not yet suited for publication but needs to be significantly shortened and made more concise. A great part of the result section is dedicated to the lengthy and extremely detailed description of the observed and modelled profiles. Unfortunately the descriptions are difficult to follow, jumping back and forth between regions/locations, altitude regimes and the different figures, with the structure of the discussion often remaining unclear. Please comprise details and have a more structured discussion clearly presenting the similarities/differences between the regions and altitude regimes. The most relevant figures in the draft are Fig. 9 and Fig. 10. Given their importance and content they are not sufficiently discussed although they actually summarize the results of the preceding lengthy and detailed discussion.

We thank Referee for his/her objective and constructive comments and suggestions.
We have entirely modified the Result Section 3, which has been shortened by around 30% with regards to the previous version.
We have also modified the conclusions to be more concise and precise and to highlight the main results.

Specific comments

– There are too many abbreviations used. I understand that this is an attempt to keep it short, but to a degree that the text gets close to unreadable. Please use full

wording more often. In particular, the letter T is used in the abbreviations for tropics and troposphere which makes it even more difficult to keep things sorted.
The way abbreviations are embedded into the text sometimes seems strange with regard to grammar.
The abbreviations NT and ST have been replaced by Northern Tropics and Southern Tropics.

– I was confused by the term 'observational site' being used for a moving platform. 'Locations' would be more appropriate in my opinion. The term 'site' is usually used for a (temporarily) fixed installation of measurements equipment in one place.
Site has been replaced by location.

– I trust the ACP editorial team will eventually take care of this but the usage of italics in subscripts and units is inconsistent and wrong in many instances.
Corrected

– it is confusing to have both an appendix and a supplementary document
We present the figures from both the appendix and the supplementary material in the supplementary section in the 2nd round revised version of the manuscript.

– Line 522: Table 2 referenced here is not part of the draft.
We put Table 2 on page 22 in the 2nd round revised version of the manuscript.

Abstract
In my opinion the abstract is too long and not well organized. The main findings are unclear. I suggest to remove some details and make the abstract more concise. It should become clear what the main conclusions of the analysis are and why these are relevant.
We rewrote the main conclusions in the abstract.

– L 11: 'in above 6 km' does not make any sense
We meant that the O3 maxima over the Asian clusters is observed above 6 km in April, between 6.2 and 11.2 km depending on the location (Figure 8, page 23 of the 1st round revised version of the manuscript).

– L 13: What do you mean by 'The highest amount of transported CO'. Transport to Asia? Overall?
We mean the overall highest amount of CO (in ppb) exported from the source regions of Northern and Southern Africa.

Introduction
– L 32: Why is stratospheric influence as the least important process mentioned first?

Stratospheric influence has been mentioned last.

– L 37: I understand biomass burning throughout the manuscript refers to wildfires excluding usage of biogenic fuels which is attributed to anthropogenic emissions. This should be made clear here.
This has been specified.

– L 59 constraint → constrain
This has been corrected.

– L 68: 'offered' – the choice of word reads strange here
Offered has been replaced by allowed.

Data and Methods
– Line 121f: what do you mean by 'a distance criteria of 300-km'?
We have rephrased by "data are selected within a 300 km radius circle centered on the airport location".

– Line 165: what do you mean by 'with bias lower than 10-15 ppb'? Please specify what bias exactly refers to here.
We have added the following in line 161-162, page 7 in the revised manuscript: "SOFT-IO captures the intensity of CO anomalies with a bias lower than 10-15 ppb **with respect to observed anomalies** for most of the regions and tropospheric layers".r

– Line 173f: percentages are not absolute differences. The statement does not make any sense to me.
The sentence containing this word have been deleted in the revised version.

– Line 177: Why two? Which two backgrounds are referred here? The two mentioned pressure surfaces?
As mentioned in line 175, page 7 in the original manuscript, the definition of the background CO (BG) is a source of uncertainty in SOFT-IO calculations. To quantify the impact of the BG definition on SOFT-IO contributions, we conducted a sensitivity test using 2 different definitions for the background CO:

- the monthly climatological median CO of a remote area away from polluted regions, between 300 and 185 hPa (UTcruise) for the period 2002 to 2020
- the monthly climatological median CO mixing ratio between 600 and 300 hPa for each location.

We have rephrased lines 175-177, page 7 in the revised manuscript as follows:
"~~In order~~ To assess this source of uncertainty, we used ~~the 600–300 hPa median CO mixing ratio as background for each location.~~ an alternative definition for the

background calculated as the median CO mixing ratio between 600 and 300 hPa for each location. The differences between ~~the two backgrounds~~ this background and the one used in the study are within 2.5-60 ppbv."

Results

All vertical profiles are discussed in terms of absolute altitude but in subsection 3.2. the different altitudes of the inbound/outbound airports are mentioned. In particular for the peaking altitudes presented in Fig. 8, I wonder what the results looked like if altitude differences relative to the local ground level were used.

The novelty of this study is that we analyse the tropical composition using IAGOS as a unified and consistent dataset for the entire tropics, as a globe. For this reason, we think that it is more appropriate to use the same methodology for the 20 locations taken into consideration for our analysis.

– Line 234: my reading from the figure would be 62+6=68

We have replaced the correct number.

– Line 240 and 248: Throughout the manuscript mixing ratios are discussed, not concentrations. Similar on several instances in the following.

We have replaced concentrations by mixing ratios.

– Line 246: the 'observed anomaly' to me is not evident in the figure.

We have rephrased as follows: "The measured CO anomaly reaches +200 ppb (after removing the background from the observed CO), while SOFT-IO attributes 40 to 80 ppb to AN".

– Line 255f: no need to cite Adon et al. twice within two lines. Skip first one.

This is done.

– Line 258: there is no obvious peak in the CO profile in the figures

We have rephrased by "Windhoek CO enhancement has the smallest magnitude among the African clusters".

– L265: space missing between brackets

Done

– L271: Reference to Fig 2l does not make sense, Fig. 2l shows CO.

Correct letter has been put.

– L445: If the vertical layers are defined on pressure as the vertical coordinate then why are km shown in the figures?

We show km in the figures in order to make it easier for the reader to understand the altitude of the CO and O3 maxima/minima discussed in the results and conclusions.

Figures

Overall, there are too many figures with too many panels and too small fonts.

We have increased the resolution and the fonts in Figure 3, 5, 6 and 7 in the revised manuscript. In addition, for clarity reasons we moved the O3 and CO verticals profiles over some clusters in the supplementary material (e.g. Sahel, Guinea Gulf (Fig. 3 panels 1-2 ab of the 1st round revised manuscript); Hyderabad and Thailand Gulf (Fig. 5 panels 1-2 be); and Addis Ababa (Fig. 7 panels 1c, 2c and 5).

The presentation of observations on an absolute mixing ratio scale and the modelled contributions as ∆CO is difficult to compare. Why are the vertical profiles shown not background corrected?

One of the main goals of the paper is to document the characteristics and seasonality of the tropical O3 and CO vertical distributions (lines 84-86, page 3 of the 1st round revised version). To do so, we present the absolute mixing ratios of O3 and CO based on IAGOS data which has not been analyzed before over the tropical band. As the paper has been commented as too lengthy with too many figures, we think it is not appropriate to add some more panels.

– Fig. 2: I suggest to have the panel labels in some lighter colour to make them visible.

Done

– Figures 3,5,6, 7 are poor resolution and cannot be zoomed which is essential given the small panels and fonts.

The size of the panels and fonts in Figures 3,5,6,7 are increased in the revised manuscript.

– Figure 8: I was wondering about the order in which locations are presented on the x-axis. It would be logical to have the locations by longitude which does not seem to be the case.

Done

Summary

The Summary largely rephrases the detailed discussion from above. Conclusions are presented alongside but are not worked out well. Please be more precise and separate the shortened descriptive discussion of the observations from the conclusions drawn.

Summary has been entirely rewritten. To separate the descriptive discussion of the observations from the conclusion we added the line 499, page 25 in the revised manuscript:

“**The results of the study indicate that** the highest O3 and CO mixing ratios are observed over Africa during the fire season in January, with anomalies located in the

LT (75 ppb at 2.5 km for O3 and 800 ppb at 0.3 km for CO over Lagos), and explained mainly by anthropogenic emissions, but with a strong contribution from fires."

– Line 611 should be 'NT' only

This is removed from the revised version of the manuscript.