Reviewer #1 (Referee-report #2)

GENERAL COMMENTS

After carefully revising all the responses to the comments made to reviewers, I see that the authors have done a thorough revision of the previously submitted manuscript. Information regarding the model setup and important descriptions in the methodology were added. Also, numerous changes to the discussion section were made. The revised manuscript was greatly improved by the new information, clarifications, and changes to the title.

That being said, I have significant concerns regarding the model setup as described in the revised manuscript and concluded that their modeling approach has serious flaws, which are then reflected in the results. The authors used to run their model an ethane emission inventory (Table 1) that does not reflect the current scientific knowledge of 1) the sectoral distribution (percent-wise), 2) the magnitude of the emissions by sector and region, and 3) the geographical distribution of emissions for the represented period (2006-2016). I will further explain these:

1 and 2) Some emission sectors and their associated emission estimates are unrealistic (e.i. SWD-solid waste and wastewater, TNR, TRO). Just as the author responded to line 269 in response to reviewer comments, ”No independent estimation of SWD contribution was found in the literature". And even so, authors included that sector without justification with emissions equal to other well-known sources such as biomass burning-BIB.

2 and 3) The geographical distribution of anthropogenic emissions presented in Table 1 is not based on any study, and no justification is offered for such values. Asian emissions are too high compared to North America and Europe (Xiao et al., 2008, Tzompa-Sosa et al., 2017).

Furthermore, emissions from the ROW-Rest-of-the-World (excluding Asia, Europe, and North America) are estimated to contribute about ~17% to global emissions, whereas Mengze et al. report a 30% contribution, again without providing any justification.

The lack of attention to these characteristics in the model input ethane emissions ultimately diminishes the quality of the results and conclusions therein. Currently, the model simulations constitute a fundamental part of the manuscript since they were used to determine the upper- tropospheric and low stratospheric trends. It is no surprise that

the model did a poor job determining regional trends (Lines 327-328 from the red-line manuscript essd-2021-246- ATC2.pdf). It is important to note that the global ethane emission in the optimized model simulation (19.3 Tg/yr) is similar to other studies (Franco et al. 2016), and this could explain why the model did a better job at determining the NH upper tropospheric trend (lines 326-327).

Overall, I can see a significant scientific value in the measurement dataset they present and the linear trends derived from the observations only. I suggest either removing all the sections and conclusions related to the model results or re-run the model with science-based emissions of ethane, methane, and propane. Additionally, it would be interesting to run the model using the trend derived from observations and see how similar or different the model is compared to observations. Major changes need to be made to this manuscript before acceptance.

We thank the reviewer for the second-round comments on our manuscript, we find the comments and suggestions very helpful.

Both reviewers #1 and #3 have suggested the removal of all model simulations. After careful consideration, we have decided to do this so as to focus on the observations only. We have therefore made major changes to the manuscript to describe and explore the observations in greater detail.

SPECIFIC COMMENTS (LINE NUMBERS CORRESPOND TO THE RED-LINE DOCUMENT essd-2021-246-ATC2.pdf)

Line 324. “The model incorporates all known emissions via emission inventories so any...”

Suggest changing the wording to “The model incorporates emissions from various emission inventories described in table XX”. In that table, provide a list of all the names (with references) of the emission inventories used for each relevant compound (i.e., ethane, methane, propane). Also, include emission totals by region and sector for methane and propane since both compounds are mentioned in the paper (similar to Table 1).

This is now being removed from the manuscript because it is related to model setup and simulation.

Authors need to avoid subjective adjectives such as “well” (line 327), “good” (line 409) and instead provide values for every comparison made.

Done. We have carefully checked the manuscript to avoid subjective description.

Suggest adding "in the upper-troposphere and lower-stratosphere" in the title.

Done. Now the title reads: “Northern hemispheric atmospheric ethane trends in the upper troposphere and lower stratosphere (2006-2016) with reference to methane and propane”

Line 160 & line 490. Does the model include chemistry for other halogens, such as Br? Also, suggest adding a reference to Sherwen et al. [2016] that concludes that Cl may be an important C2H6 sink that can decrease the simulated global burden of C2H6 by about ~20%. Also, add other studies that estimated the impact of the lack of other halogens like Br.

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 175. It seems that there is confusion between the mole fraction term and emissions. The emissions were optimized to match the observed mole fractions. Thus, “We further optimized ethane mole fractions for each emission section” is technically incorrect.

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 179. What is the value of this minimum RMSE?

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 209-210. If the Prophet algorithm is designed for non-continuous datasets, why do results match with the NOAA algorithm even though it is designed for continuous datasets? Is this an expected result, or what does having similar values mean?

Because the Prophet algorithm is the only product that we know which works for non-continuous datasets, we cannot test and compare it directly with other algorithms using a non-continuous dataset. The NOAA algorithm has been examined many times in the literature and it has advantages in deriving trends for continuous datasets. We therefore used a continuous dataset to test the performance of Prophet versus NOAA algorithms. The agreement of results from both algorithms (as we expected) indicates that Prophet algorithm works for continuous datasets, and may indirectly prove its performance for non-continuous dataset. In another words, if Prophet algorithm didn't match with NOAA algorithm using a continuous dataset, it means that Prophet algorithm may not even be able to work for our non-continuous dataset. That's why we did this comparison.

The entire Section 3.1 should be moved to the introduction. Also, the first paragraph of section 3.2 should be moved to the introduction. The time of collection and location of the sampling should not be in the discussion but in the introduction and methodology section.

Done. We have moved the section 3.1 and the first paragraph of section 3.2 to the introduction. We have also removed the paragraphs in the introduction that are related to model and emission inventories.

Lines 277-278. Are there any studies that talk about a weaker mixing of stratospheric air over EUR? If so, add references.

We did a literature search but not find any studies reporting a weaker stratospheric mixing over EUR.

Lines 304-307. Can authors provide values for trends due to sampling location vs. trends derived from observations to support this conclusion?

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 338-371. As suggested in the general comments, model sectors and model geographical sectors (section 3.3.2) contribution analyses should be deleted because

sources, totals, geographic distributions, and trends do not reflect current scientific knowledge.

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 390-392. Having peaks in all regions does suggest regional and global increases in emissions, but not an increase in fossil fuel emissions. Thus, remove "fossil fuel" from this sentence.

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 412-414. This result should show the estimated emission in 2006 and the subsequent increase due to the increasing trend derived from observations. Otherwise, state the exact year at which those estimated emissions by sector correspond.

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 541. Remove the word "budget". Also, see the previous comment on how to present this trend value. Possibly write it as "averaged" emission for the same period.

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 545. The underestimation is only ~30% based on the optimized model emissions (13.3 Tg vs. 19.3 Tg). When re-writing this sentence, make it clear which emissions this comparison refers to.

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 551-552. "…was caused by fossil fuel-related emissions, likely from oil associated and natural gas sources". The conclusion of the 2010-2011 peak coming from oil and gas emissions is stated in the text (line 390) only as a possible explanation and is not based on any of the main results (see Fig 3) and discussions made in the text (lines 330, 390, 430, 478). Even figure 3 (lower panel with model optimized sectoral emissions) does not reflect this statement, as other sectors such as SWD, RES, and TRO show the same increase in 2010-2011. Therefore, this conclusion should be removed from here and from the abstract since the authors have no strong results to reach such a conclusion in the manuscript.

This is now being removed from the manuscript because it is related to model setup and simulation.

Lines 552-553. "The global ethane trends cannot be well simulated by advanced atmospheric chemistry modeling ". This is a strong claim that is not supported in the text. If there are clear examples of studies that have had struggles modeling ethane trends, they should be cited here and in the text so that this conclusion is consistent with the main document.

This is now being removed from the manuscript because it is related to model setup and simulation.

TECHNICAL CORRECTIONS

Line 174- Move “(13.2 Tg/yr)” to the end of line 172.

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 187- Add upper case in "Real..."

This is now being removed from the manuscript because it is related to model setup and simulation.

Line 321. Model estimation or optimized model results?

This is now being removed from the manuscript because it is related to model setup and simulation.

Add the word "modeled" every time the text refers to regional trends. For example, in line 475, it should say: "The modeled regional trends..."

This is now being removed from the manuscript because it is related to model setup and simulation.

Reviewer #3 (Referee-report #1)

Review comments on “Northern hemispheric atmospheric ethane trends (2006- 2016) with reference to methane and propane”

Author(s): Mengze Li et al. MS No.: essd-2021-246

MS type: Data description paper

General comments

Li et al. present long-term ethane mole fraction data obtained from airborne observation in the UTLS region in the Northern Hemisphere through IAGOC-CARIBIC project for the period 2006–2016. The authors also present simultaneously measured propane and methane mole fraction data as well as output of simulated atmospheric ethane using the EMAC model to understand the contributions of different source regions/sectors to the observed ethane variations.

After the revision, the manuscript and dataset has been generally improved, but at the same time it has also revealed significant issues for publication. Before going into the details, I would like to emphasize again that the observational ethane, propane, and methane data obtained from this project is very unique and useful for the scientific community, because there exist only limited airborne observation platforms that can collect the air samples in the UTLS region and can measure ethane, propane, and methane with high quality over the 10 years. I think the data can used by many researchers in this field after the issues as below have been resolved.

The current manuscript is mainly focusing on the comparison of the trend components derived by Prophet algorithm between observations and model outputs, which even precedes the presentation of original observational data. I think this has hidden the overall features of dataset as well as data quality and is not well suitable for the data description paper. I would like to ask the authors to pay more attention about the observation data itself (i.e., rigorous data classification in terms of altitudes, latitudes, longitudes, PV, and etc, also see my general comment 2). Now, the authors have newly added the timeseries of raw measured data together with their estimated trend by the Prophet (Fig 5 and Fig 8

and other figures in the supplement). The figures are very useful and helpful to understand the overall structure of dataset – for example, how the sampling frequency is for 2006-2016 in this project, how the variance of the raw data is, and how the trend is properly estimated. For the future revision, I would like to suggest that the figures are presented before presenting the comparison with model nor respective components by Prophet, as long as the authors purpose is the publication of their observational dataset. I think to guarantee the quality of the dataset it depends on whether the QC of the observation data is done properly and the presentation of observation data itself is reasonable, rather than the comparison with the model output.

We appreciate a lot that the reviewer provides very helpful and constructive comments for second-round revision. We are also glad to see that the reviewer finds our data very unique and useful for the scientific community. In this second iteration, we have removed all the model setup and simulations and now focus entirely on the measurement data, as suggested by both reviewers. We have made major changes to the manuscript, added resampling methods to estimate trend uncertainty, and discussion on data classification, as suggested by the reviewer.

1) The current data classification in terms vertical axis is based on whether it is lower or higher than 2 PVU only. I found the data regarded as upper troposphere contains even much lower altitudes than expected (946 m at minimum). The change of such spatiotemporal sampling density has significantly affected the estimated trends and conclusion. For example, one of the main conclusions of this paper (and dataset) is that a sharp increase of atmospheric ethane was observed in the upper troposphere in 2010-2011, the cause of which is estimated to be fossil fuel related sources. However, according to the Fig 5 and Fig 8, it is found there is no observation data around half month before and after 2010, and then two flight data just after the sampling program restarted in 2010 showed overall higher mole fractions with relatively small variance, which looks affecting the rapid change of trend estimated by Prophet (Fig 5). I found most of these flight data were mainly sampled at altitudes lower than 8000 m (please see the attached Figure A). I would say such an increase in 2010 is not trend in the upper troposphere nor caused by the fossil emission increase, but is just the reflection of the vertical profile in the trace gases (see Figure B). Similar points can be said to the stratospheric data because it contains

different strength of stratospheric data (from 2 PVU to ~12 PVU, see Figure C-D). Please reconsider how the data classification is reasonable to be presented as the data description paper.

We thank the reviewer for pointing this out. We have now (1) excluded samples collected below 8,000m from analysis, (2) added more figures and discussion on the data classification for the lower stratospheric trends, see section 3.3.4. As suggested by the reviewer, we now identify the tropospheric samples as altitude >8,000m & PV<2, and the lower stratospheric samples as PV>2. We also presented the trends of the lower part (2≤ PV<6 PVU) and upper part (PV≥6 PVU) of the lower stratosphere, to investigate the changes of trends within the lower stratosphere and the difference when using different data classifications.

2) Under the temporal sparseness of data in several years and in several regions, the analysis using Prophet would have underestimated the uncertainty in estimated trend. I don't think the uncertainty can be evaluated just by changing the fitting parameters "changepoint_prior_scale" (as authors did), because the estimated trend cannot usually avoid the overfitting when there is a large temporal data gap. If the authors would still keen to estimate the trend from Prophet, I would like to request to do some bootstrap tests, for example, jackknife method, to evaluate the robustness of estimated trend. Such tests should be especially important when estimating the trends for each geographical region, due to the sparseness of their data compared with whole NH dataset (i.e., ASI in the upper troposphere and NAM, ASI, and EUR in the lower stratosphere). Any discussions and interpretations on the seasonality and trend (and comparison with model) should be made after such cautious statistic test is done.

This is a good point. We now use resampling methods to evaluate the trend uncertainty (section 2.2). We have 20 sub-regions (4 categories: upper troposphere, lower stratosphere, lower part of the lower stratosphere, upper part of the lower stratosphere; 5 regions of each category: whole NH, EUR, NAM, ASI, ROW), dataset of each region is resampled 20 times with each subset consisting 90% of the sample number. For each subset, we run 10 times "Prophet" algorithms using 10 levels of fitting scale, and use the average of 10 fittings as the non-linear trend for that subset. For example, there are 3,288 upper tropospheric samples, we run 10 fittings with Prophet to estimate the trend, then we

resample the dataset of 3,288 samples 20 times randomly, each subset from resampling consists 3,288*90%= 2,959 samples, we run 10 fittings with Prophet for each subset, and plot the trends of 20 subsets as the trend uncertainty for the original dataset which has 3,288 samples. That means, for each sub-region, we run 10 times level of fittings with Prophet for each of the 21 datasets (1 original + 20 subsets from resampling), that is 210 times run of Prophet. We have 20 sub-regions, so in total 4,200 run of Prophet was done and the results are shown in the figures and manuscript to evaluate the robustness of trend uncertainty.

3) I think the procedure of model optimization is not so unique nor well organized, which makes me wonder if the model output is useful for understanding observation data and is worth publication, contrary to the significance of observation data. If the authors main purpose is to publish the observational data, I would suggest they focus more on the presentation of observational data itself and statistical analysis after rigorous data classification and may remove the model output from the publication of data. Otherwise, if their focus is both the observation data and model output, I would suggest they develop and sophisticate their model optimization procedure for the future use of model data.

We have now removed all the model outputs and focus on the presentation of observational data.

Considering the above, I would recommend this paper for publication after major revisions.

Specific comments

Overall: There are many confusions between “mole fraction” and “trend”. After reading through the manuscript several times, I can imagine the authors somewhere refer to the trend component derived from Prophet as trend (e.g., Fig 3-8 and related sentences), but it is not always easy for readers to follow because we generally expect the ethane trend as in Table 2 or its related sentences (the unit, %/year or ppt/year). Also there still exists multiple places where it is hard to distinguish if the authors are referring to modeled results or observations.

We have removed all the model outputs. We have also added clarification on different trends in section 2.2 Trend analyses. We now refer the trend from linear fitting on the entire time period as “linear trend”, and the trends derived from Prophet algorithm as “non-linear trend” because Prophet outputs trends for each day when the day has observations, and such trends are non-linear.

L25-L26: Now I wonder if the statement of emission estimate can be removed because the aim of this study using model is not to validate the emission inventories and also because the emission optimization methodology is not well sophisticated.

We agree. This is now being removed from the manuscript because it is related to model setup and simulation.

L28 "An ethane plume": I would recommend replacing it by "An sharp increase of ethane". "An ethane plume" sounds like more short-term intrusion of polluted air masses but here it seems saying about inter-annual variations. However, as I commented, I think it is also reflected by changes in sampling altitudes (please see general comment 2).

We have removed this statement, and added figures and text in manuscript to reflect the changes in sampling altitudes.

L29-31 “and higher temporal-spatial resolution data of ethane are needed.”: I would suggest deleting it here since there is no discussion about the issues of spatiotemporal resolutions of emission inventories in the main text.

Done, this has been deleted.

L52-L76: As long as I understand, the author is reviewing the study of emissions estimates from top-down approach in this paragraph, and discussing the issue of discrepancy between top-down and bottom-up estimates in the next paragraph. I would suggest clarifying it since there are several confusions as below.

This paragraph has been deleted as it is related to model and emission inventories.

L54-L56: The sentence can be deleted since this paragraph is the summary of top-down approach (as long as I understand); thus the discussion on inventory is just confusing.

Done.

L66: Insert "based on top-down approach" after "global emission estimates".

This paragraph has been deleted.

L76-L80: This sentence is not clear. Please rephrase.

Done. Now the revised sentence reads:

"Previous investigations of the distribution, emissions, lifetime, and atmospheric trends of ethane have been mostly based on surface-based measurements. These have been either from a regionally focused intensive field measurement campaign (e.g. Kort et al. (2016)) or from networks of remote sampling stations (e.g. Franco et al. (2015), Helmig et al. (2016))."

L108-L109: By inspecting the dataset, I found 946 m sampling altitude data exists as the minimum altitude. I would strongly recommend the author presents the information on the frequency of sampling altitudes and reconsider the data classification.

Done. We have shown the sampling altitudes of all the data in Figure 1, and classify tropospheric samples as altitude>8,000 m and PV<2 for further analysis, as suggested by the reviewer.

L109-L114: I would suggest the author explains the measured trace gases by each three detectors separately or just explains the measurements of ethane, propane, and ethane, respectively. Under the current writing, the reader cannot still find the correspondence between each detector and measured trace gas.

Done. We have now explained the measurement by each detector separately.

L135-L137 "Three injections of calibration standards": What does it mean? Need clarification.

Done. We have revised the text and it reads now:

"Three additional calibration standards samples were measured in between samples of each flight sequence in order to monitor the quality of measurements and reduce uncertainty."

L164-L165 "to reduce the uncertainty": What is the "uncertainty" here? Why combining AIR, BIB, and BIO into one sector can reduce the "uncertainty"?

This is now being removed from the manuscript because it is related to model setup and simulation.

L166-L176: Actually, this part is still not clear and confusing to me.

This is now being removed from the manuscript because it is related to model setup and simulation.

L168-L172: What is "the total amount"? The global total anthropogenic ethane emission? Despite the optimization, it is said the value used in this study is ~11.8 Tg/yr whereas the value used in Pozzer et al. (2020) is 13.2 Tg/yr? Are these global total ethane emissions including natural sources; thus they are not consistent even after the adjustment?

This is now being removed from the manuscript because it is related to model setup and simulation.

L173-L174 "We further optimized …": Do you mean the 45% increase is done from CAMS-GLOB-ANT v4.2 inventory itself or from that after multiplied by 2.47 as mentioned above? In the current manuscript, it sounds like the authors attempted to optimize the global total ethane emissions in the latter (two-step) method. If the optimization method is the former, I would suggest deleting or much reduce the sentence in L168-L172 to avoid confusion and clarity.

This is now being removed from the manuscript because it is related to model setup and simulation.

L174-L176: What is “whole dataset”? Do you mean many iterations were performed by changing the scaling factors to determine the minimum of RMSE? Please clarify more. The authors didn’t present the raw output of model simulations, but I think at least it is desirable to present the correlation plot between observations and model outputs after the optimization (maybe in the supplement).

This is now being removed from the manuscript because it is related to model setup and simulation.

L205-L208: Please see my general comment 3. I wonder if this approach is reasonable when estimating the trend uncertainty in this case whose data is sparse in time.

We have added resampling methods to estimate trend uncertainty, as suggested by the reviewer, details please see our response to general comment 3.

L210-L213: What is “start” and “end” representing here? For example, in L366-L368, how the start and end of the data were determined? Is it also applied to the output derived from Prophet or applied to the raw observational data? Is it strongly dependent on the measurement data at the tEnd and tstart.

We have replaced this with linear fit, and explained more in detail in section 2.2.

L216-L251: The two paragraphs are just listing of information and hard to read. Please summarize and be more concise. The descriptions of emission estimates are also mixed in L232-L234 and L234-L235. Since the authors intended to summarize atmospheric ethane trend here, the sentence of emissions can be removed for the readability. The headline can also be changed as “Literature perspective of global atmospheric ethane trends” in this case.

We have now put these two paragraphs in the Introduction section as suggested by another reviewer. We have removed the parts regarding emission estimates as they are not relevant to the study. We have also summarized the main messages in the beginning of each paragraph so that it is easier for readers to follow.

L272-L273: Any possibility that the sampling altitudes in EUR are lower than the other regions? As I commented in L108-L109, the information of sampling altitudes for each region is also necessary to be presented (and please reconsider the data classification when needed).

We have added Figure 1 and Figure S3 to show the distribution of sampling altitudes, and added text in the manuscript to address this (section 3.2.2). Indeed, the sampling altitudes in EUR are lower than other regions.

L285-L359: Please see my general comment 1. I think the comparison of estimated ethane trends between observations and model (Figure 3-4) should be followed by the presentation of raw observational data itself (Figure 5). This is important for potential data user because they use the raw data not the estimated trend derived by Prophet.

Done. We have now removed all the model output.

L309-L311: Delete?

Done.

L323-L324: As commented in L28, the annual increase is not short-term.

This is now being removed from the manuscript because it is related to model setup and simulation.

L329-L331: This sentence is not clear to me. Please rephrase it.

This is now being removed from the manuscript because it is related to model setup and simulation.

L333-L334: FEF looks rather increasing over 2011-2013 in Figure 3a.

This is now being removed from the manuscript because it is related to model setup and simulation.

L335-L336 “We note…”: This sentence is just suddenly appearing here. Delete?

This is now being removed from the manuscript because it is related to model setup and simulation.

L337-L346: This may not be an important finding from this study and can be removed. The author didn’t optimize each emission sector but just roughly scaled up by 45% from CAMS-GLOB-ANT v4.2?.

This is now being removed from the manuscript because it is related to model setup and simulation.

L366-L374: Please see the comment in L210-L213. Also how to estimate the max and min?

Done. This has been removed and replaced by linear fit methods.

L377-L380: As commented, the high mole fractions in the first two flights in 2010 would have been due to the lower sampling altitudes. I think it is an overinterpretation to relate it to the increase of fossil emissions in 2010-2011 at this stage. After data classification with altitudes, it would be useful to present some scatter plots of ethane- methane or ethane-propane when the ethane mole fractions are high levels in order to relate with fossil sources.

We have removed the interpretation on fossil sources, and now the manuscript focuses on data itself.

L389-L391: I wonder why the authors write this here. As long as I understand, here the authors are attempting to present not the accuracy of their model, but the validity of their optimized global total ethane emission?

This is now being removed from the manuscript because it is related to model setup and simulation.

L407-L431: Please see my general comment 3. Please reconsider investigating the uncertainty before going into the detailed discussion even if the authors still wish to keep these parts.

Done, we have removed all the model simulation.

L463-L464: I think this is an important issue on the data classification, not limitations. Also, as I commented, the issue of temporal data gap should also be critical when estimating the trend.

We have added more discussion on data classification, and used resampling methods to examine the trend analysis uncertainty, please see our replies to general comment 2.

L472 “the WMO tropopause”: Replace it by “the WMO thermal tropopause”.

Done.

L474-L476 & L481-L483: I don’t think the authors need to add these sentences.

Done. We have removed these sentences.

L501-L504: Please add the information of the calibration scale, instrument, measurement precision, the number of measurements, and standard deviations into the dataset.

Will be done when I upload the dataset.

L510-L512: Please clarify more.

This part has been removed because it is related to model simulation.

L515-L516: I think the remarks on emission estimates need to be minimized because the current optimization method is very rough and the authors are also saying the aim of model simulations is to better understand the contributions from each emission sector, not improving the emission inventories.

Done. We have removed all the model simulation.

L518-L519 “a factor of three”: As long as I read through the paper, the author has increased the global ethane emissions by 45% from CAMS-GLOB-ANT v4.2 to match their atmospheric ethane observations, not by a factor of three.

This part has been removed because it is related to model simulation.

L526-L527: There is no presentation on the simulation results of methane in this study. It sounds very sudden.

This part has been removed.

Table 1: Please specify “inventory” as CAMS-GLOB-ANT v4.2. Also the row of “Biogenic emission” and “Biomass burning” also needs a headline. Please add the explanation that the sum of BIO + BB and (a) or (b) is equal to total sources.

This is now being removed from the manuscript because it is related to model setup and simulation.

Table 2 “~”: What does it means? The uncertainties presented here represent the same manners? Also please unify the usage of parentheses in the first column; otherwise add the new column to describe the locations of measurements?

We have added explanation in the table legend: “Table 1. Summary of studies reporting ethane trends in the (a) troposphere and (b) stratosphere. Parentheses in first column indicates the locations of measurements.”

Figure 5: The trends in Propane are difficult to distinguish (also in Figure 8, S6-S9). As commented, these time series plots are very important and preferred to come first rather than comparisons with model outputs. Also useful to show the sum of all components (i.e. best fit curve) derived from Prophets.

This figure has been now deleted because it is related to model simulation.

Figure S3: Not clear the meaning of "trend analysis uncertainty". What is the meaning of percentage here? The uncertainty range is generally smaller for the period 2006–2009 when the number of data is small, potentially due to the over-fitting to the sparse data?

This figure has been now deleted because it is related to model simulation.

Figure S4: I don't think this figure is necessary even in the supplement. The contributions of each emission sector to the atmospheric simulated ethane trends are already shown in Figure 3 and Figure 6. This figure looks rather confusing.

This figure has been now deleted because it is related to model simulation.

Table S2, Figure S5, Figure S11: Isn't it the sectoral contribution to "simulated ethane mole fractions" not "ethane trends"? If it is saying the values of trend component derived from Prophet, please clarify so.

This figure has been now deleted because it is related to model simulation.

Figure S6-S9, S12-S15: Same as the comments in Figure 5.

This figure has been now deleted because it is related to model simulation.

Figure S7 and S13: It looks the observational data number is very sparse and difficult to estimate the robust trend? Need further statistical test to investigate the effect of time data gap (e.g., bootstrap method) if the authors still wish to present their trends.

As the reviewer suggested in previous comments, we have added resampling methods to evaluate the robustness of trend analysis and added text in the manuscript to raise the point of data gap.

Figure S12-S14: There are vertical lines in the left sides. Need removed.

This figure has been now deleted.

Technical corrections

L52: Insert “atmospheric” before “observations”.

This is now being removed from the manuscript because it is related to model setup and simulation.

L253 “It is noted …above, any…”: Conjunction is required.

Done. We have revised the sentence, it now reads:

“It is noted that our aircraft samples have significantly different spatial distributions compared with the studies summarized in the Introduction section, therefore, any comparison should be made in a careful manner.”

L253-L257 “…of data quality, a process”: Also need conjunction.

Done. Now it reads:

“When comparing surface and airborne datasets from multiple locations to assess global atmospheric changes, as it will become increasingly important to ensure comparability of data quality. A process that has begun through the grounding of a World Calibration Center for VOCs, although this dataset predates this initiative.”

L273-L278: Very long sentence. Please separate it into two sentences.

Done.

L459-L497: I doubt the writing style in this section is following a proper manner. There are headlines starting from alphabet in parenthesis - some are complete sentences and some are not but all headlines start with lower cases.

We will consult with the journal editorial team to see if the writing style is proper and how to revise.

L499: 4 Data availability.

Done.

L507: 5 Conclusions

Done.

Figure 2: Insert “observed” before “over the whole …”.

This figure has been removed and replaced by other figures.

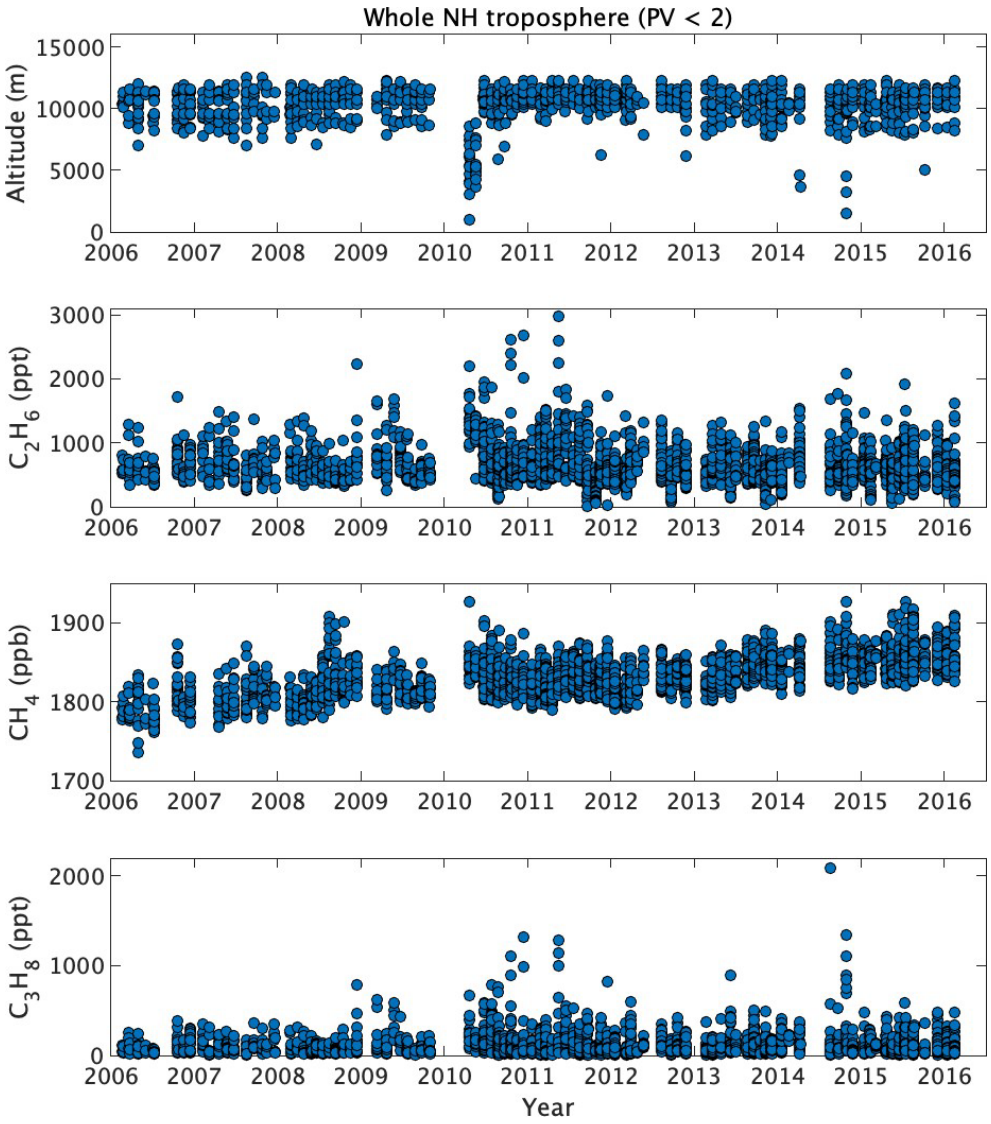

Figure A: The observed ethane, methane, propane mole fractions for the whole NH upper troposphere (same as Figure 5 in the authors' manuscript), together with sampling altitudes.

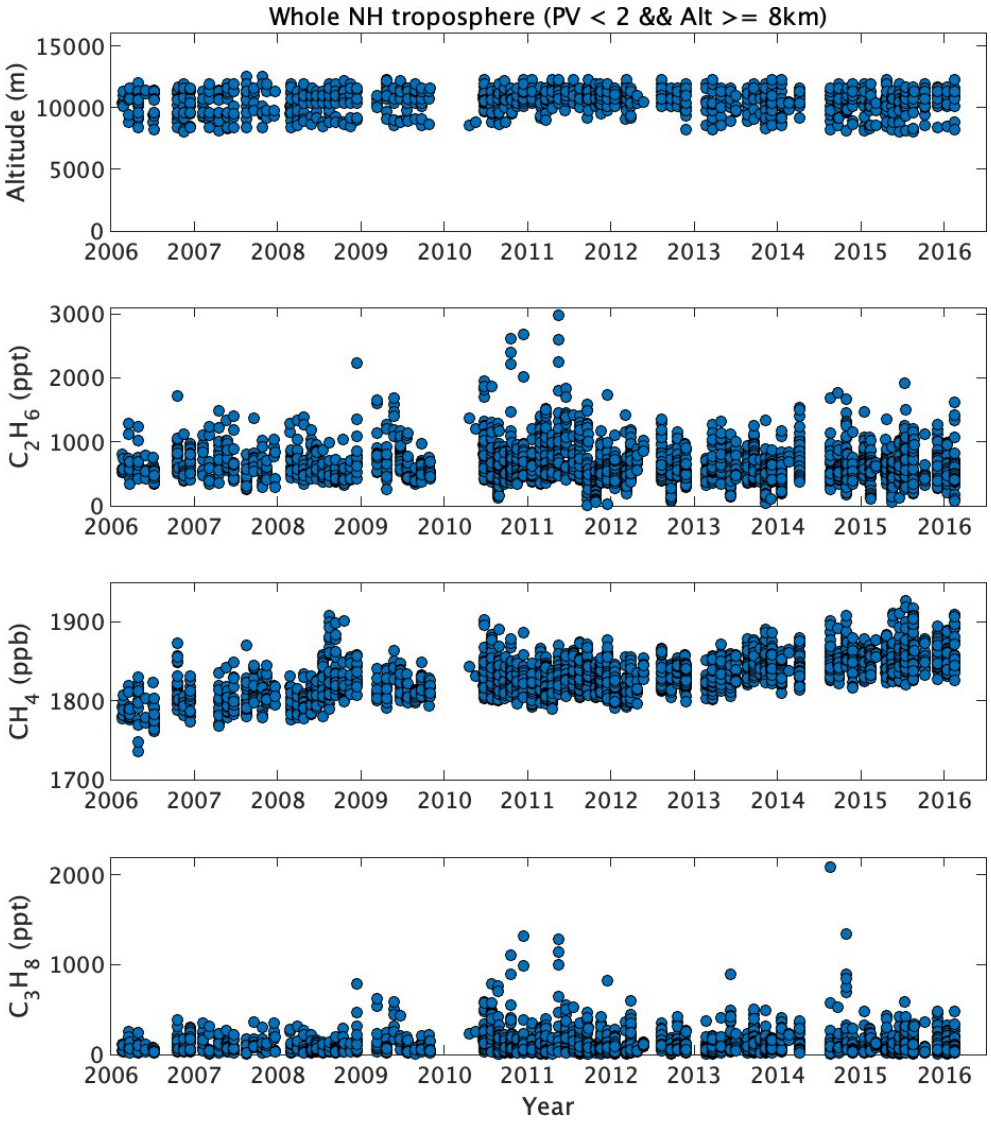

Figure B: Same as Figure B, but for sampling altitudes over 8000m.

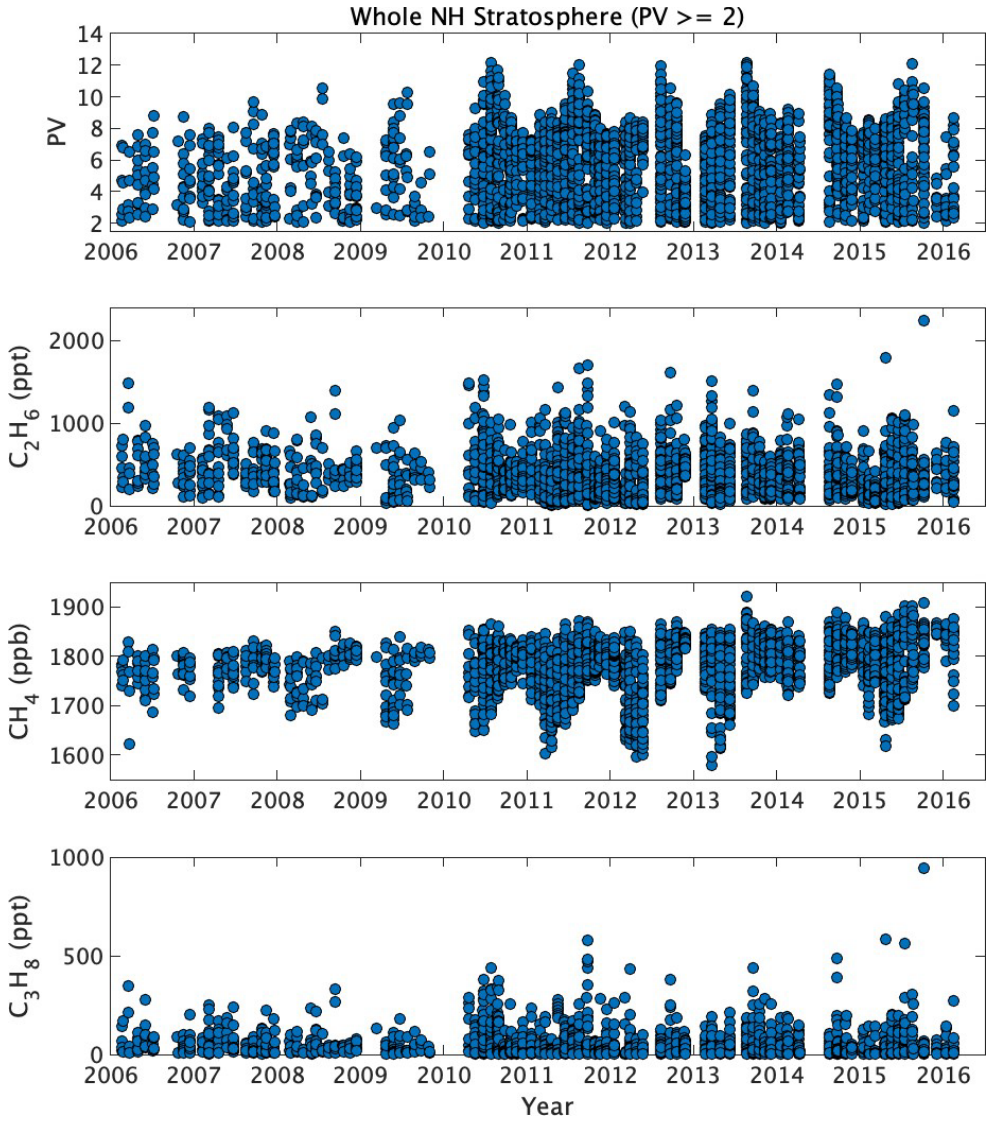

Figure C: The observed ethane, methane, propane mole fractions for the whole NH lower stratosphere (same as Figure 8 in the authors' manuscript), together with PV.

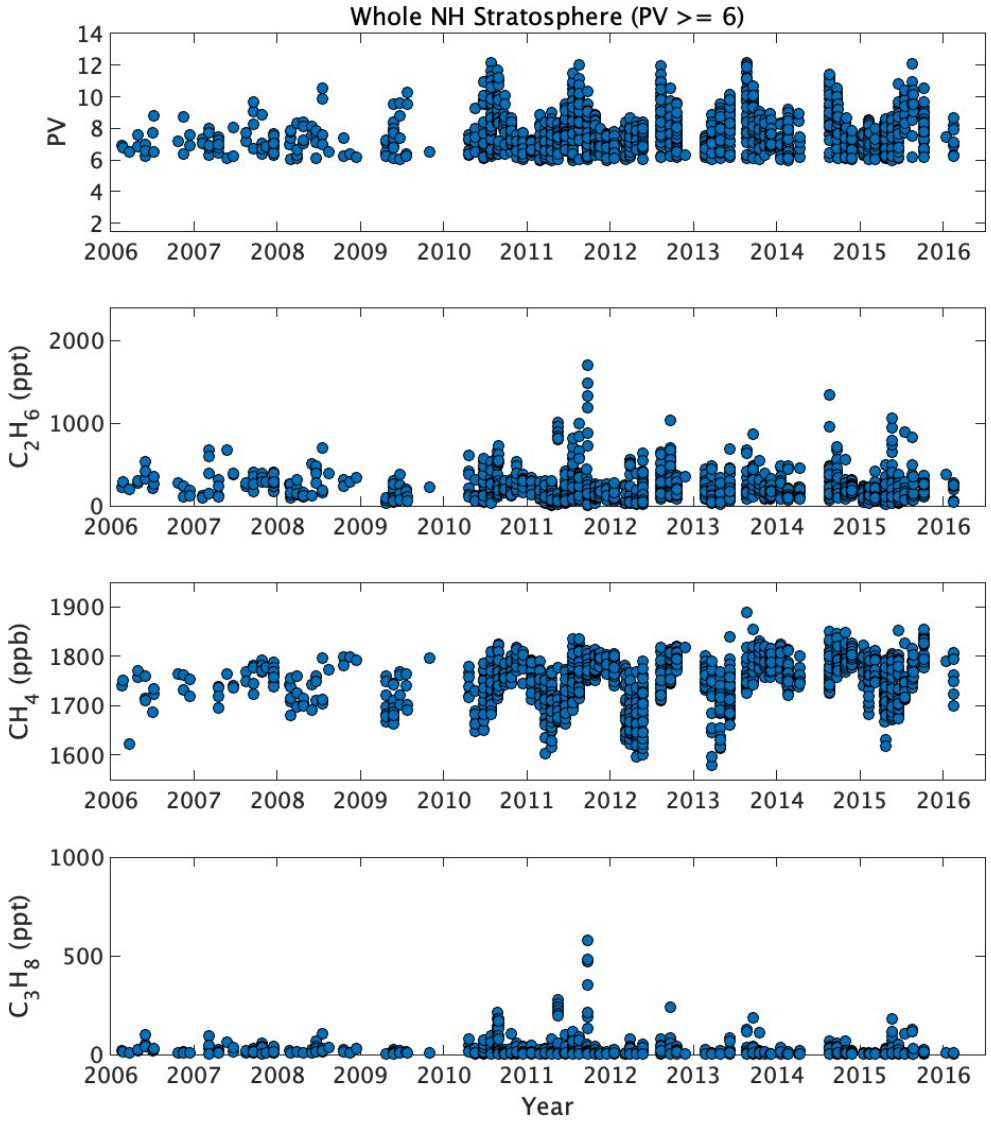

Figure D: Same as Figure C, but for PV over 6 PVU.