**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.

1) 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.

2) 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.

3) 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.

4) 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.

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.

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.

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).

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.

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.

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.

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

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

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.

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.

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

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

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

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?

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.

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).

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.

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 $t_{End}$ and $t_{start}$.

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.

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).

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.

L309-L311: Delete?

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

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

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

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

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?.

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

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.

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?

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.

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.

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

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

L510-L512: Please clarify more.

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.

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.

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

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.

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?

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.

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?

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.

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.

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

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.

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

**Technical corrections**

L52: Insert “atmospheric” before “observations”.

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

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

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

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.

L499: 4 Data availability.

L507: 5 Conclusions

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

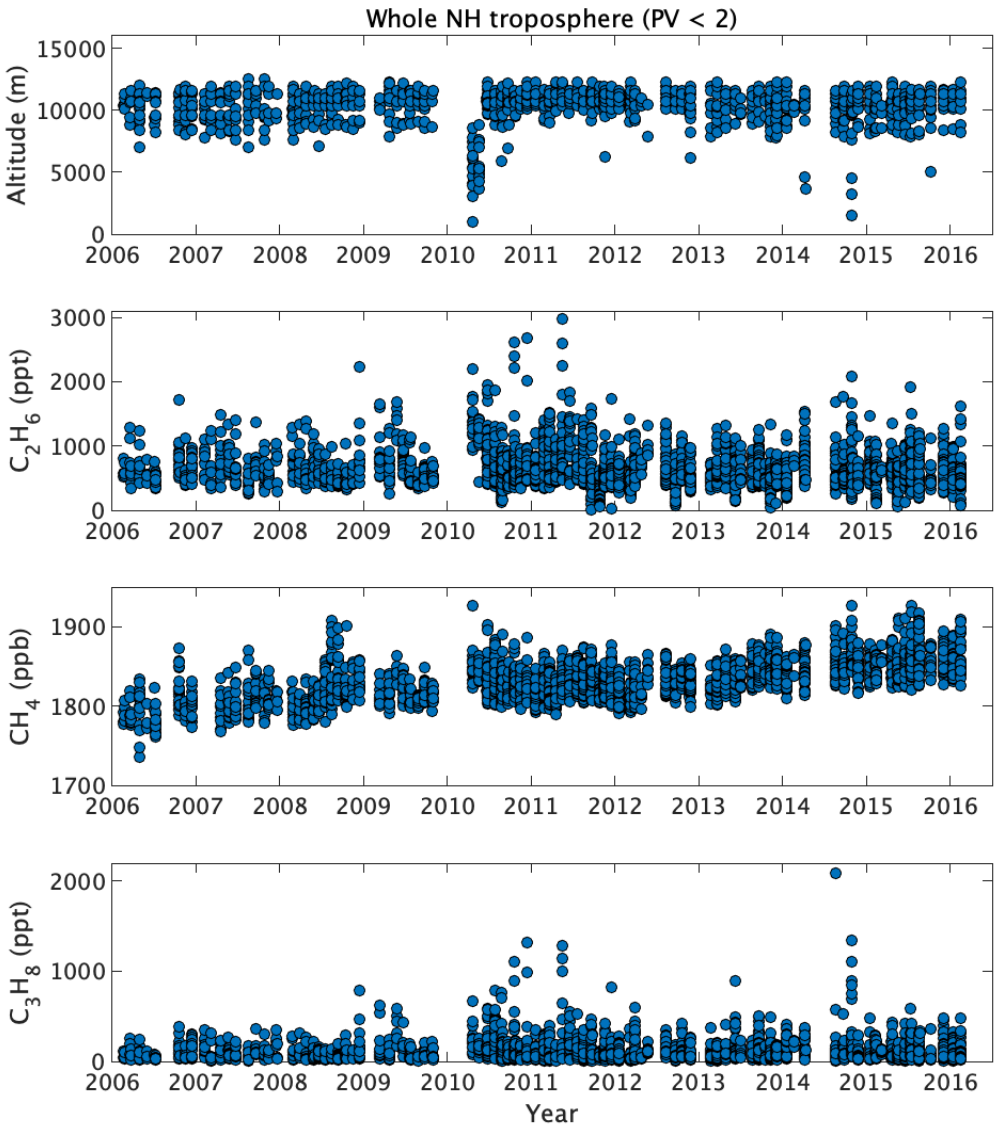


Figure A: The observed ethane, methane, propane mole fractions for the whole NH upper troposphere (same as Figure 5 in the authors' manuscript), together with sampling altitudes.

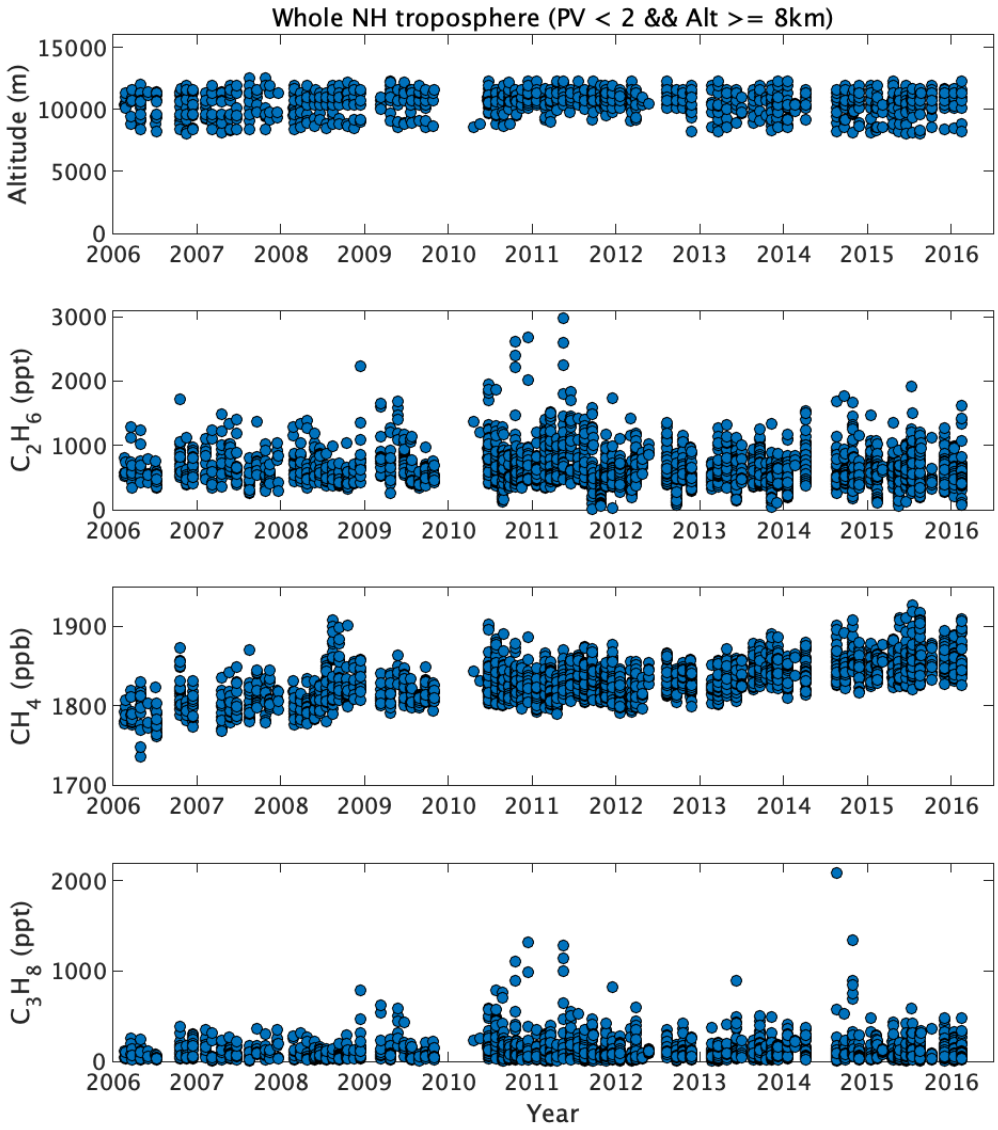


Figure B: Same as Figure B, but for sampling altitudes over 8000m.

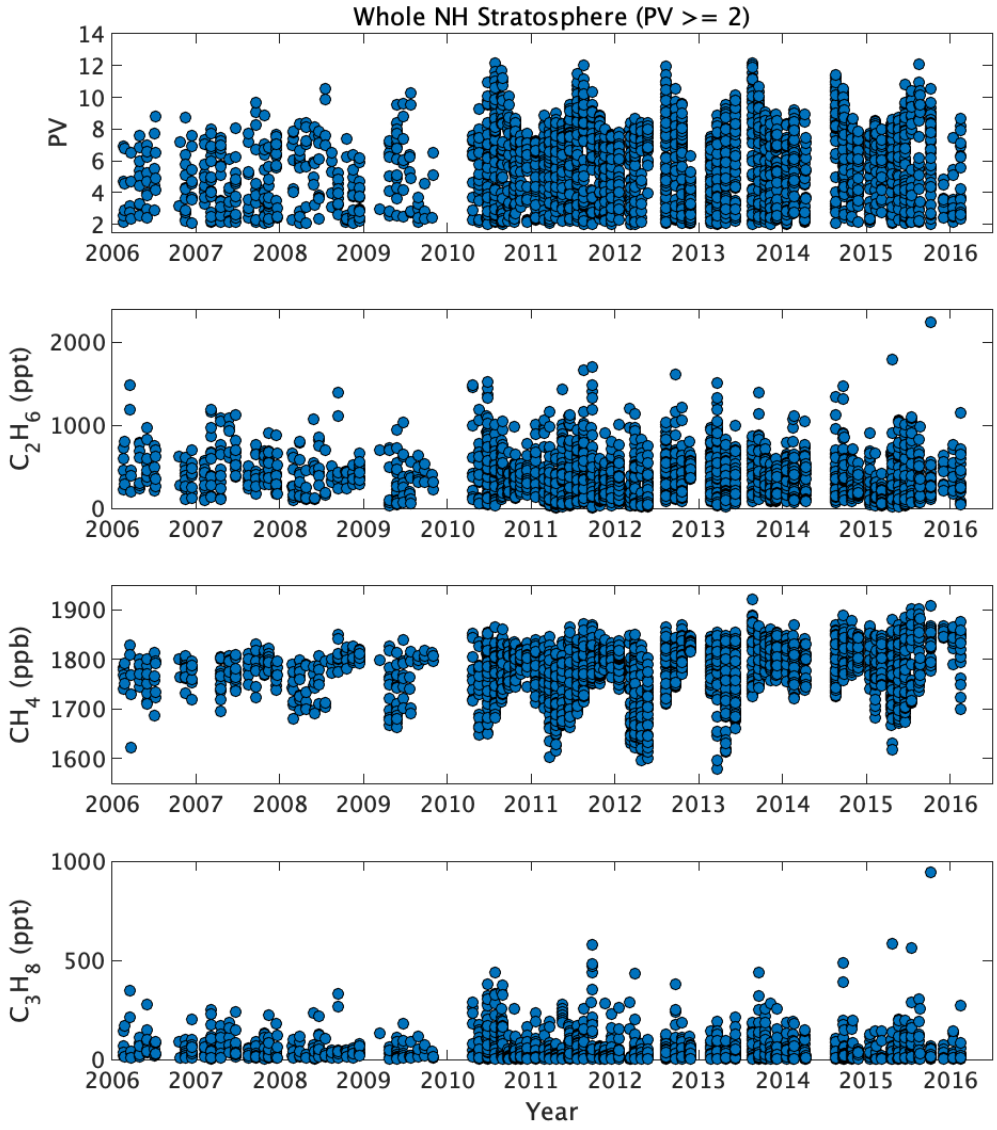


Figure C: The observed ethane, methane, propane mole fractions for the whole NH lower stratosphere (same as Figure 8 in the authors' manuscript), together with PV.

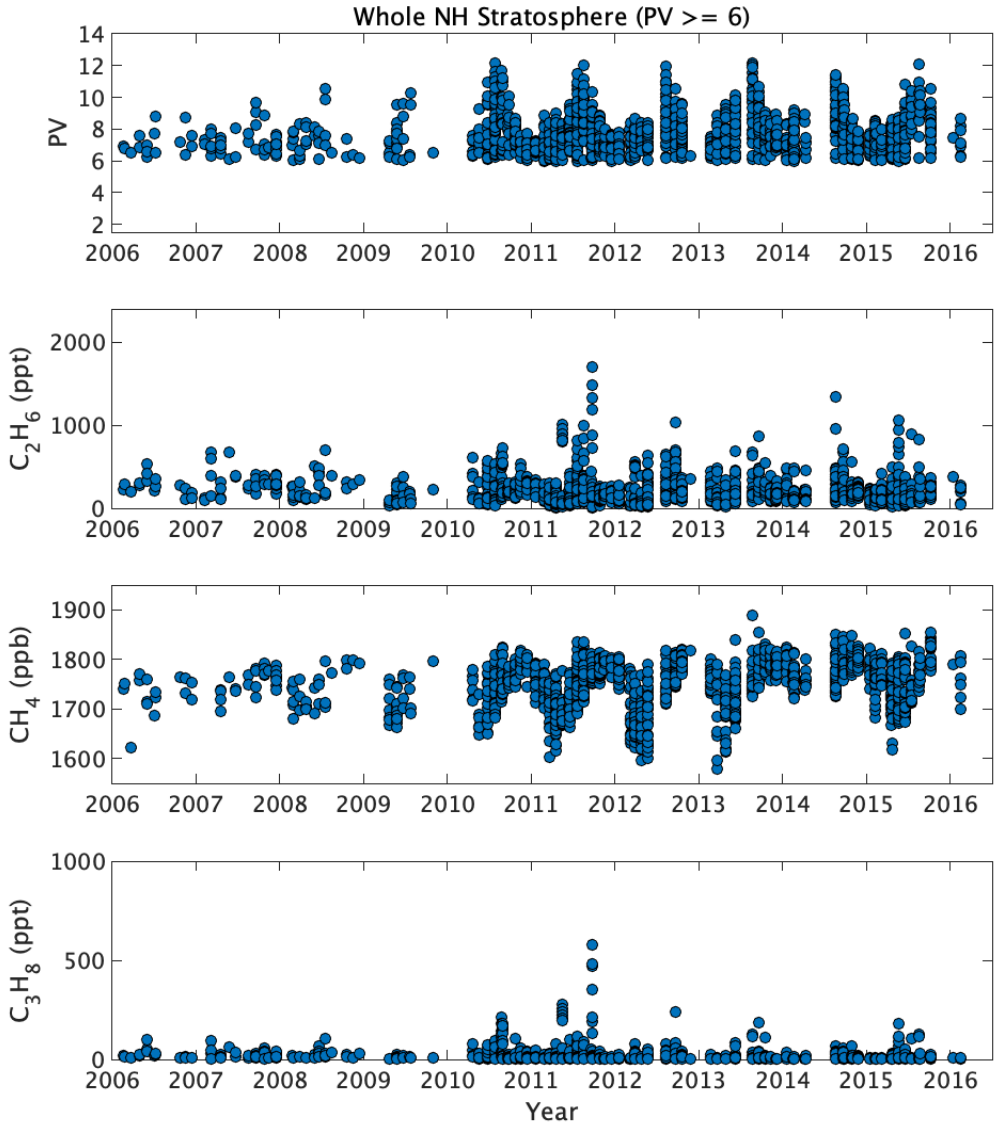


Figure D: Same as Figure C, but for PV over 6 PVU.