Reply to Anonymous Referee #1

The authors thank the Referee for his careful reading of the manuscript and for his thorough review. A detailed point by point reply (in blue) is provided hereafter.

This paper presents results from an improved retrieval method for obtaining total columns of formic acid globally during seven years from IASI radiance spectra. This simple and computational inexpensive method has been used before for other species and presents some improvements to results obtained before for HCOOH by the same group. It is based on the calculation of conversion factors starting from a representative set of formal retrievals with the optimal estimation method, and using these to convert brightness temperature differences to column amounts. The results presented show improvements to previous attempts to derive reliable global distributions of HCOOH. This work has the potential to be published in ACP after a significant improvement of the manuscript is carried out by the authors.
In particular, I would like to point out that it is very poorly written in terms of wording and sentence structures, which makes the text at times very difficult to read and follow. Some (but clearly not all) minor corrections are listed below. It is important that the text is revised and improved by someone with good English skills. Also, some sections could be shortened and the key points could be better explained in a more concise manner without leaving important information out. Additionally to this, please consider the following points for improving the content and structure of the manuscript:

The text was revised in order to improve the English.

l.61 why would this be necessarily from a direct flux of HCOOH, couldn't it be also secondary formation from other unknown VOCs?
This comes from the conclusions given by Millet et al. (2015), -see reference in the manuscript:
In this paper it is said that "This indicates one or more large missing HCOOH sources, and suggests either a key gap in current understanding of hydrocarbon oxidation or a large, unidentified, direct flux of HCOOH."
Moreover, the secondary formation from other unknown VOCs is included in "hydrocarbon oxidation".

l.94. Please comment on how (and why) the Averaging Kernels of the ground-based FTIR retrievals can be compared to the normalized Jacobians from the IASI retrievals. I could't find any information on this also in section 3.2
We now provide this information in the section 3.2:
"The AKs indicate the vertical sensitivity of the retrieval. The Jacobians express the sensitivity of the radiative transfer model and the IASI instrument (through its instrumental function) to the variation of HCOOH in the atmosphere. Both functions then give a good indication of the vertical sensitivity for each data set."

Fig 2. This figure would be more appropriate later on in the section describing the comparison with FTIR. Also, separate into two adjacent plots with common y-axis and individual x-axis for both the normalized Jacobians (left) and FTIR AK (right). Avoid the inset and use larger labels if it is to fit into a one column of the text.

As suggested by the reviewer, we enlarged the labels for the Jacobians plot and the inset is removed. We also decided to separate both plots (Jacobians and averaging kernels) as below. Now there are 2 figures:

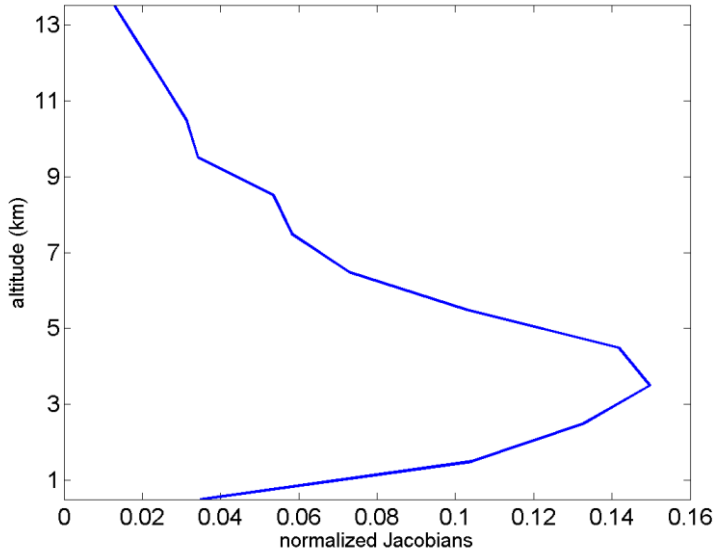


Figure 2. Mean normalized Jacobians of all retrieved spectra (over the 7 selected regions) as a function of altitude.

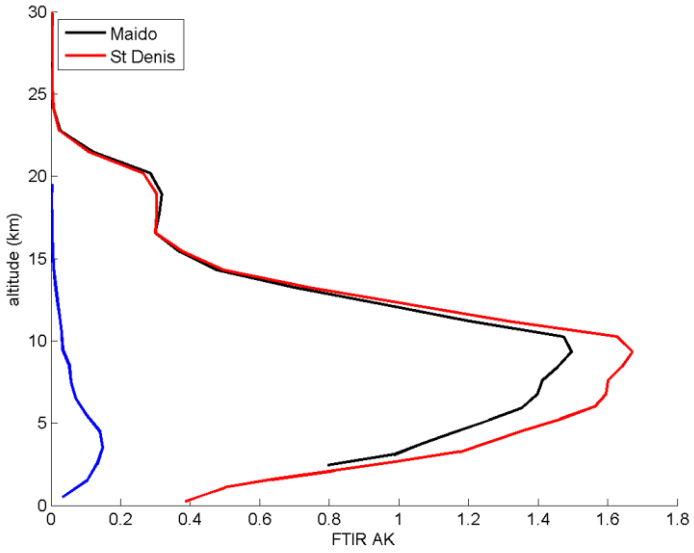


Figure 11. Mean total column AK for the FTIR ground-based measurements over Maido (black) and Saint-Denis (red) at La Réunion. Both stations are shown by green stars in Fig 1. Both FTIR stations have a degree of freedom for signal (DFS) close to 1. As reminder, the mean normalized Jacobians from Fig. 2 is plotted in blue.

This Fig 11 is located in the Section 3.2 "Comparison with ground-based FTIR measurements"

Fig 2 caption. What is a "degree of freedom of signal"?? Ground-based FTIR retrievals often report the degrees of freedom (DOF) with respect to the independent layers sensed. A value of 1 would mean that no information in the vertical distribution is accessible. Does DFS refer to the same thing?

Thanks for noting this: we change the 'of' into 'for' so that DFS = degree of freedom for signal. Yes the concept is similar for ground-based instrument and satellite retrieved data and both acronyms DOFs and DFS are used in the literature, e.g. for DFS:
Deeter, M. N., H. M. Worden, D. P. Edwards, J. C. Gille, and A. E. Andrews (2012), Evaluation of MOPITT retrievals of lower-tropospheric carbon monoxide over the United States, J. Geophys. Res., 117, D13306, doi:10.1029/2012JD017553.
Worden, H. M., Deeter, M. N., Frankenberg, C., George, M., Nichitiu, F., Worden, J., Aben, I., Bowman, K. W., Clerbaux, C., Coheur, P. F., de Laat, A. T. J., Detweiler, R., Drummond, J. R., Edwards, D. P., Gille, J. C., Hurtmans, D., Luo, M., Martínez-Alonso, S., Massie, S., Pfister, G., and Warner, J. X.: Decadal record of satellite carbon monoxide observations, Atmos. Chem. Phys., 13, 837-850, doi:10.5194/acp-13-837-2013, 2013.

A DFS close to 1 means that the signal corresponds to a column, with no profile information available.

l.101 use the more conventional expression with B as subscript ΔT_B here and throughout the manuscript.
As suggested, we changed it to ΔT_b in the text, in the captions and on the axes for the Figs. 3 & 4.

l.102 I think the use of spectral microwindows is here more appropriate than "spectral channels".
"Spectral channels" is the appropriate term for this study as the calculation of ΔT_b is based on specific channels and not on a full microwindow.

l.110 what are these "mean RMS" differences? Do you refer to residuals? Please be clearer.
The RMS is the square-root of the differences between the observed and the fitted spectra. The RMS mentioned in the manuscript is the mean of all RMS, thus the mean of all square-root of the residuals.
We agree, the word "difference" was forgotten in the text and it was confusing. We have deleted it.

l.117 again, unclear of how you define channels. Are these mean brightness temperatures within a spectral range (microwindow) or rather just a value at a specific wavenumber?
We used the value at a specific wavenumber:
Tb HCOOH at 1105 cm^{-1}
Tb ref 1 at 1103 cm^{-1}
Tb ref 2 at 1109 cm^{-1}
Thus the sentence "The reference channels used for the calculation of ΔT_b were chosen on both sides of the HCOOH channel (1105 cm^{-1}), i.e. at 1103.0 and 1109.0 cm^{-1}." is correct.

l.123 Why not use tao for termal contrast as in previous studies?
This definition of thermal contrast was used in several previous studies from our group (e.g. Clerbaux et al. (2009) – cited in the manuscript). It is a standard definition.

Fig 3. Labels are missing in the plot to the right, if same as the one to the left, just include it to the x-axis. Also for the color palette. Use a) and b) to describe the plots as in Fig 4.
Done

l.125 The linear regression for obtaining the conversion factors from the correlation between the OEM method and the ΔT_B's was gathered from retrievals performed in different areas of the world, representing different conditions, which is good. The question is if these areas are treated separately would result in very unique conversions factors (something not shown in Fig. 3) and which could be used to improve the conversion from ΔT_B to total columns. Is it sufficient to consider a correction for the dependance to the thermal contrast? Please comment.

Let's start the answer by a clarification.

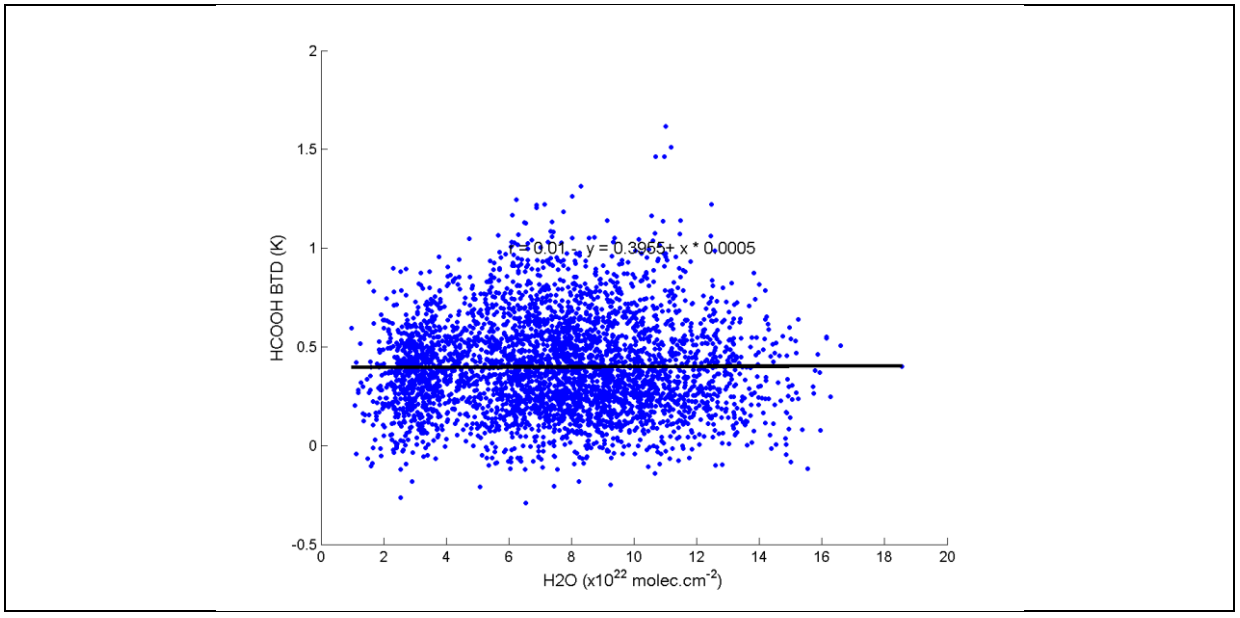
From line 125, we described the correlation between the columns retrieved by OEM and the ΔT_b. We highlighted the difficulty to convert these ΔT_b to columns using these coefficients since there is still an impact of the thermal contrast. Hence the conversion factors were found in section 2.3.1 and illustrated by Fig.4.

Indeed the factors will change with the used a priori columns. Since the idea was to have common coefficients for the full globe, all spectrum were gathered in a single set.

About Fig.3: If we separate different areas, we will obtain different correlations in each region, as it is characterized by different amount of HCOOH, temperature profile, etc.

About the dependence to the thermal contrast, this is a good remark. Razavi et al. (2011) took also into account the impact of H_2O in their conversion. It is not presented in the paper, but we also checked if the ΔT_b were correlated to the H_2O columns.

Hereafter you can see the scatterplot between the ΔT_b of the spectra used for the OEM-based retrieval over the seven regions and their H_2O column. No clear correlation was found (r=0.01).



We also performed a similar test than the test performed for the TC. We modified the H_2O profile by +10% and +20%. Thus 3 different profiles were used: H_2O ref, H_2O ref +10%, H_2O ref + 20%. We performed forward simulations and we checked the correlation between the simulated ΔT_b and the H_2O columns. We obtained the following scatterplot:

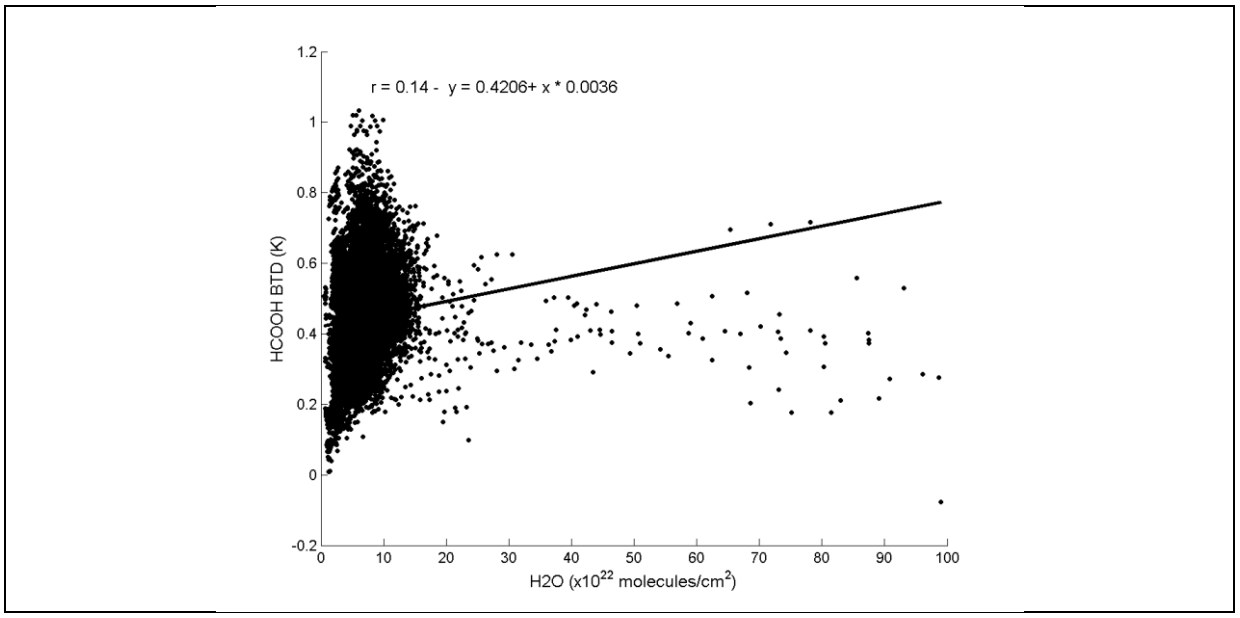


Fig 4b. Use same scales in the x and y-axis as in Fig 3a.
Done

l.152. What do you mean with only negative averages are being filtered out? You have just stated that negative values would produce a bias, so no filtering should be performed. Do you refer to columns used for the comparisons with FTIR and modeled data?
We wanted to say that the negative values were also kept to calculate the averages, but if these averages were found to be negative, then they were filtered out. We agree the sentences were confusing.
Now the sentence reads:
"For comparisons with zonal or temporal averages, the negative total columns were included in the average. But when the average was found to be negative, it was filtered out."

l.178 The results part showing the description and interpretation of the obtained global distributions should go AFTER the comparison with ground-based FTIR measurements and the model results. Maybe as section 4: Results. It is important to know how reliable (or not) the data are before using them for interpretation.
This is an interesting remark but we presented this section as the first section in the analysis part since we compared the new dataset with the work done by Razavi et al. (2011). We decided to keep this structure.
The Fig. 7 highlights the difference on the global distribution between both retrievals and it is important to analyze the reasons for the differences before further interpretation of the data.

To clarify our analysis, we also added this information at the beginning of the section (in bold):
"Mean HCOOH global distributions (averaged on a 0.5°×0.5° grid) from IASI for the 2008-2014 period are presented in Fig. 7 and compared with columns obtained using the retrieval method of Razavi et al. (2011). Note that Razavi et al. (2011) retrieved only total columns over land. Except over Indonesia, lower values are observed over the source regions with the updated dataset. The previous section shows that large positive RDs are expected for very low true columns. Even if the columns from Razavi et al. (2011) are not the true columns, this could explain why the total columns for this study are higher over remote areas (e.g. deserts) than those obtained using the methodology described by Razavi et al. (2011). **It is also**

important to note that in Razavi et al. (2011), only averaged data in a 0.5°×0.5° grid with TC higher than 5K were considered. This implies that only data with a strong signal were used, probably overestimating the threshold of the ΔT_b and thus also the retrieved columns."

Moreover, the Figs. 8 and S1 help to interpret the peaks observed on the time-series in the following section, i.e. for the comparison with the FTIR measurements. Thus it is a good reason to present the global distributions before the comparisons with the FTIR and the CTM.

l.276 If the large biases found between the retrieved columns and this work does not come from using a simplified retrieval method as opposed to the OEM, then explain where the bias comes from. The explanation that IASI overestimates for background levels in La Reunion because of the larger errors in the conversion from brightness temperatures is not valid in the case of Wollongong, or is it?. Please provide with a more solid explanation.
The bias was not coming from our conversion method. The overestimation of the background levels over La Réunion and at Wollongong could be a result of a low detection limit due to a low local TC.

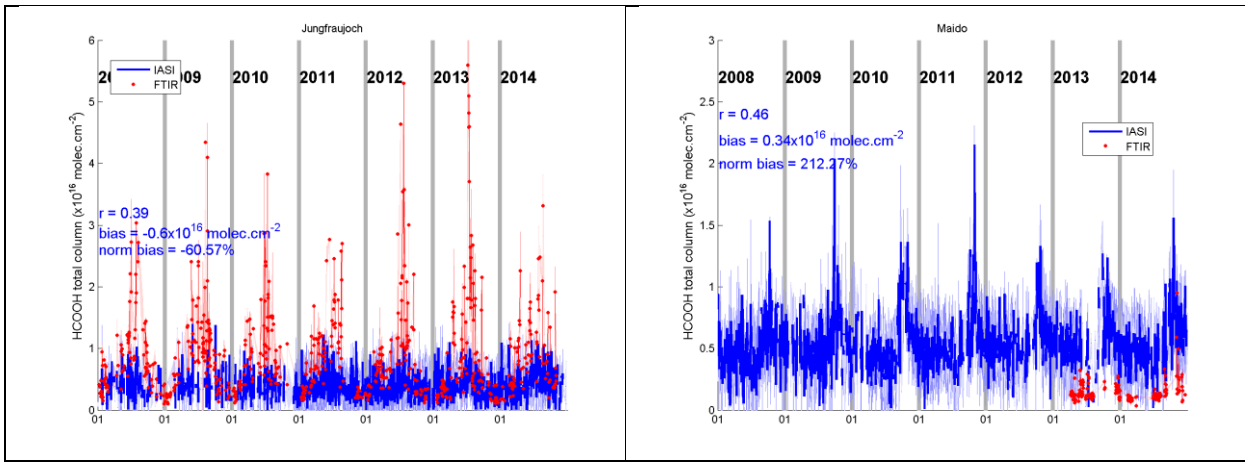
l.255 The altitude correction performed to both FTIR and IASI total columns is poorly argumented. Despite the fact that the correlation might improve, it may do so for the wrong reasons. I don't agree the authors should do this correction. It is quite feasible that a mountain site might not be sensing a plume further down while the broader pixel size of the IASI instrument covering lower altitudes might very well be detecting it.
You are right but if this correction is not applied, the comparison will be biased for the sites at high altitudes due to the absence of the lower levels. Indeed, over these high altitude stations the retrieved column are truncated since the lowest layer are not represented.

The equation is a simplified formula which is a variation of the hypsometric equation (Wallace and Hobbs, 1977).
Wallace, J. M. and Hobbs, P. V.: Atmospheric Science: An Introductory Survey, 1977.

You can find an example of its use in:
De Mazière, M. et al., Comparisons between SCIAMACHY Scientific Products and Ground-Based FTIR Data for Total Columns of CO, CH_4 and N_2O, Proceedings of the Second Workshop on the Atmospheric Chemistry Validation of ENVISAT (ACVE-2), ESA-ESRIN, Frascati, Italy, 3-7 May 2004 (ESA SP-562, August 2004) ESC02MDM.

We added this sentence to the paper "This simplified formula is a variation of the hypsometric equation (Wallace and Hobbs, 1977)" with the corresponding reference.

We show below the same comparison between IASI and FTIR data as presented in the manuscript for Jungfraujoch and Maido, without the altitude correction:



l.260 If daily averages from FTIR are used instead of a more constrained time with respect to the IASI overpass time, then the authors should present the results showing that there is no improvement. I don't understand why they say the correlation does increase when using a +/- 2h criterion and still don't apply it. In my opinion for a compound with such a short lifetime, in the range of hours, a more constrained time criterion than daily averages should be used in this study.

We agree the sentence was confusing. We decided to add this sentence (in bold).

"A more stringent criterion of ±2h was tested but provided similar results, except over Maido where the correlation increased to 0.6 without improvement of the bias. **The advantage of this daily average is the possibility to derive the seasonal variation over each site**. Over all sites, the broad patterns of seasonal and inter-annual variations were similarly captured by IASI and the ground-based FTIR."

The reviewer can see an illustration of this sentence with this figure:

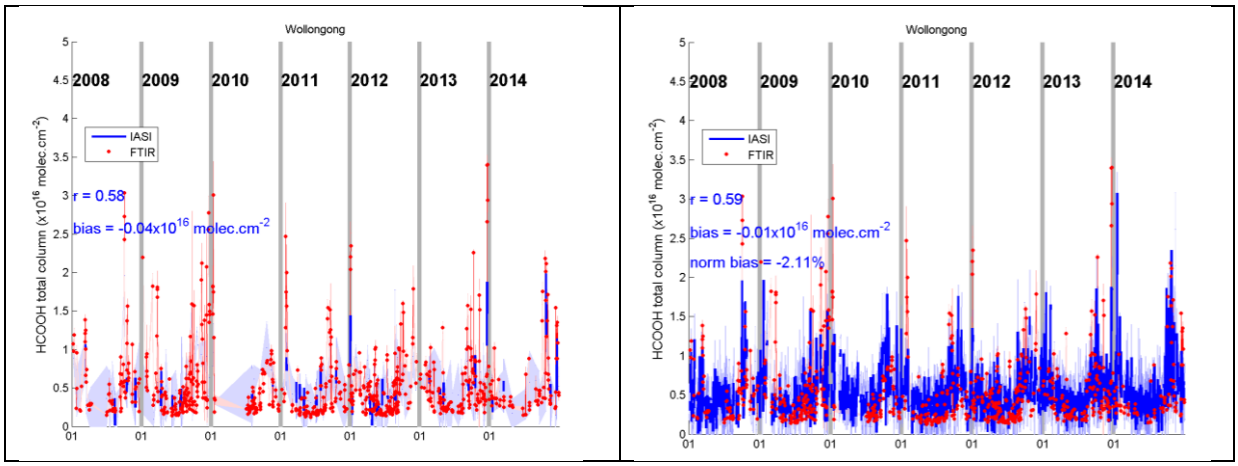


Figure. Left: Time series of HCOOH over Wollongong between 2008 and 2014 for IASI (blue) and the ground-based FTIR (red) measurements. The IASI data are collocated at ±0.5° and ±2h around each FTIR measurement. The correlation coefficient, the mean bias for all years is given in blue on each plot. The blue shade error bar corresponds to the standard deviation on the IASI daily means. **Right:** As left panel but the IASI curve corresponds to daily averages and the IASI data are collocated at ±0.5° around the site location. This plot is the time-series presented in the paper.

l.288 The authors decide to use a broad spatial coincidence criterion when comparing IASI with FTIR measurements, broader than in previous studies. Please provide with a more thorough explanation of why this decision was made. There seems to be enough IASI measurements to still have enough coincidences.

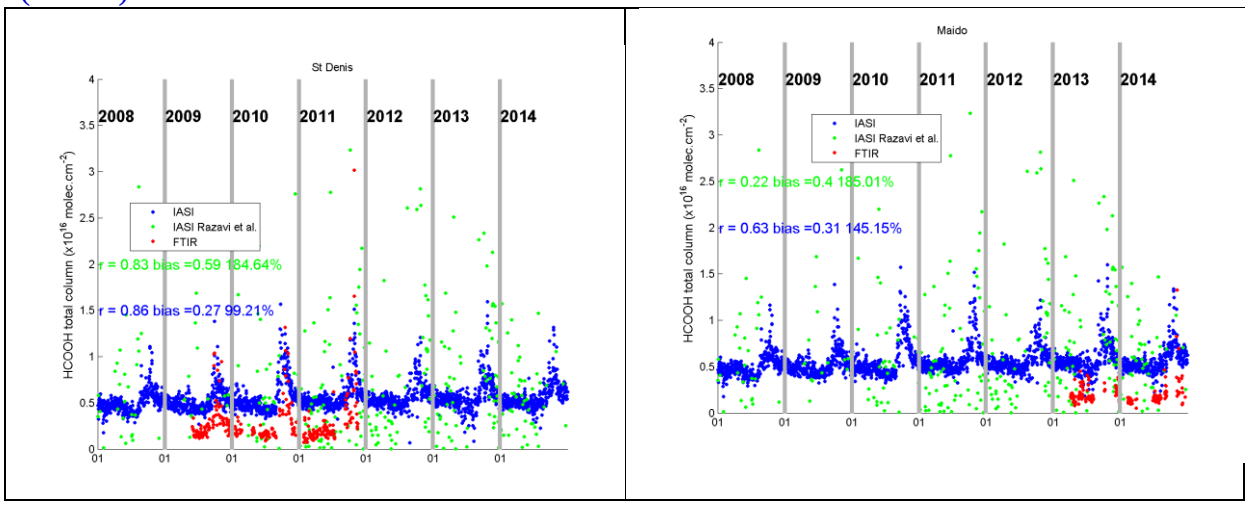
The idea was to compare this updated dataset with the work performed by Razavi et al. (2011).

Stavrakou et al. (2012) presented a comparison with the colocation criteria used in our work. The number of available data from Razavi et al. (2011) is also less important than the current study due to their stringent criteria in their conversion (averaged TC in a 0.5°x0.5° grid > 5K).

We clarified this point in the text:

"The FTIR measurements were also used to evaluate the current HCOOH columns with those using the conversion from Razavi et al. (2011) (Fig. 10). **The colocation criteria have been enlarged to ±4° as used in the evaluation shown in Stavrakou et al. (2012). The criterion was enlarged since the number of available data from Razavi et al. (2011) around the sites was less important than for the current dataset."**

The reviewer can find hereafter the time-series over La Réunion (Saint Denis and Maido) site using the colocation criteria used in the paper (daily averages, ±4°) and illustrating the lower number of data with the technique from Razavi et al. (2011) (in green) compared to our work (in blue).



Minor corrections

l.20 "There are, however, large uncertainties on the sources and sinks of HCOOH and is therefore misrepresented..."

Changed.

"There are, however, large uncertainties on the sources and sinks of HCOOH and therefore HCOOH is misrepresented by global chemistry-transport models".

l.24 "The dependence..." sentence unclear

The sentence has been modified as below:

"The dependence of the measured HCOOH signal to the thermal contrast is taken into account in the conversion method."

l.40 rewrite "is among the most ..."

done

l.49 rewrite to make a clear and correct sentence
The sentence is now:
"HCOOH is a short-lived species and its lifetime is mainly determined by the precipitation rate. The lifetime ranges between 2 days during the rainy season and 6 days in the dry season in the boundary layer (Sanhueza et al., 1996). The global lifetime in the troposphere is 3–4 days (Paulot et al., 2011; Stavrakou et al., 2012)."

l.54 rewrite "despite the..."
done

l.58 rewrite "of emissions..."
the "s" was added.

l.67 rewrite "provided..." since that instrument is no longer operational
corrected.

l.68 missing argument, ACE provides what in the upper troposphere and how often?
We added the information (in bold):
"…and the solar-occultation Atmospheric Chemistry Experiment (ACE) **provides seasonal global distribution** in the upper troposphere (e.g. González Abad, 2009)."

l.71 rewrite ", so it is a challenge to...radiances."
done

l.74 rewrite "during the summer..."
done

l.74-82 sentences poorly written, rewrite.
The sentences are now:
"These studies however highlighted discrepancies between the retrieved distributions and especially within enriched HCOOH air masses as, for instance, over large forest fires. Indeed, the total columns from R'Honi et al. (2013) were on average a factor of 2 lower than in Razavi et al. (2011) (around a factor of 1.5 for columns higher than 5×10^{16} molec.cm^{-2} and 2.3 for columns lower than 5×10^{16} molec.cm^{-2}). In this paper, we present an update of the method used in Razavi et al. (2011), in order to derive HCOOH distributions over both land and sea, suitable for both enhanced and background concentrations over the period 2008-2014."

l.76 which method?
The information in bold is added:
"…an update of the method **used in Razavi et al. (2011),…"**

l.77 rewrite "over the period 2008-2014"
done

l.86 rewrite "...Fourier transform infrared spectrometer."
Done. Now the upper-case letters are deleted except on the word Fourier.

l.87 rewrite "on board"

done

l.101. why would you use the word robust if the drawback of this method, apart of being computationally cheap, has large errors and no AKs (see l.156).
The term “robust” was deleted.

l.111 “This” refers to what? Use proper sentence structures.
Now, it is “This RMS value”.

l.112 redundant use of “conversion” in one sentence
It is changed. The sentence is now:
“The conversion factors allowing the calculation of total columns based on ΔT_b values…”

l.202-211 This paragraph should probably go a the end of the listing 1-10 as it refers to something general and not a specific region.
The initial order was chosen in order to separate the description of the annual distributions (Fig.8) and the monthly distributions (Fig.S1).
As requested by the other reviewer, this part was rewritten. We also moved the mentioned paragraph at the end of the section.

l.243. remove “way”
done

l.266 “strict of stricter”?
Thank you for finding this typing error. It is strict.