Reply to Anonymous Referee #2

This paper presents a (very brief) description of an improved HCOOH from IASI retrieval method, evaluates the new method performance against the prior (Razavi et al., 2011) method, and carries out validations against both in situ FTIR HCOOH measurements and CTM (Imagesv2) output. There are numerous interesting results, however some topics need to be addressed more fully. Furthermore, while the paper is well organized over all, the paragraphs are often not well connected and it does not read very smoothly. I believe the paper should be published in ACP after some moderate revisions.

The authors would like to thank the Referee for his careful reading of the manuscript and for his constructive comments. This paper certainly benefited from these corrections. A detailed point by point reply (in blue) is provided hereafter.

1.Comments on content

a.Section 2.2:

i.This section could use some more details on which forward model and which retrieval code were used for the OEM retrievals. Are these the same as in Razavi et al., 2011 (hereafter just Razavi)? This is not stated. Even if they are, providing this information adds clarity to the paper.

We thank the referee for this comment. We added the following sentence (in bold):

"In the current work, the main difference with the previous IASI HCOOH determination in Razavi et al. (2011) is the use of retrieved total columns over selected regions to determine conversion factors, instead of the use of forward simulations. **The OEM implemented in the line-by-line radiative transfer model Atmosphit (Coheur et al., 2005) has been used as in Razavi et al. (2011).**"

With the following reference:

Coheur, P.-F., Barret, B., Turquety, S., Hurtmans, D., Hadji-Lazaro, J., and Clerbaux, C.: Retrieval and characterization of ozone vertical profiles from a thermal infrared nadir sounder, J. Geophys. Res., 110, D24303, doi:10.1029/2005JD005845, 2005.

ii. The authors state that they used a large variability in the retrieval (350%) based on the retrieval settings of Razavi. I think a better term would be a loose constraint. I also do not understand what is meant by "based on the settings".

We agree with the referee that the term "based on the settings" is confusing, so it was deleted. We also changed the "large variability" by "loose constraint".

b.Section 2.3.1

i. Is the thermal contrast really defined as the difference between the surface temperature and the air temperature right above the surface? A more appropriate variable for satellite IR sensors is the difference between the surface temperature and the temperature at the peak of the instrument sensitivity.

If the ΔBT from forward model runs is plotted as a function of thermal contrast (from the definition used in this paper), it will not be zero when the thermal contrast is zero. I suspect that the correction for ΔBT developed in this section would not be necessary if the definition I suggested were adopted.

I would like to see a plot of ΔBT vs thermal contrast for various profiles using both definitions. If the authors feel this does not belong in the paper (though I believe it is an

important point) they can submit the plots in their response. However, if the plots confirm my hypothesis I leave it to the editor to decide if this section should be omitted and the rest of the analysis redone.

The definition of TC is similar to that which was used for several previous papers ((e.g. Clerbaux et al. (2009) – cited in the manuscript; or Bauduin et al. (2014)). The correction on ΔT_b is also similar to the correction performed in previous studies (e.g. Razavi et al. (2011)).

S. Bauduin, L. Clarisse, C. Clerbaux, D. Hurtmans, and P.-F. Coheur: IASI observations of sulfur dioxide (SO_2) in the boundary layer of Norilsk , J. Geophys. Res., 119, 4253–4263, doi:10.1002/2013JD021405, 2014.

As can be seen on the Jacobians plotted in Fig.2, the maximum of vertical sensitivity of IASI is located at around 3 km. As suggested, the same test as performed in Fig.4a was done by perturbing the temperature profile at 3 km, by ±5K.

A "thermal contrast (TC)" defined here as the difference between the surface temperature and the temperature at 3km is calculated.

A similar but slightly lower correlation (see figure below) is found with the new calculation than the result found in the paper (r=0.58 hereafter, and r=0.61 in the paper). The equation found is also similar.

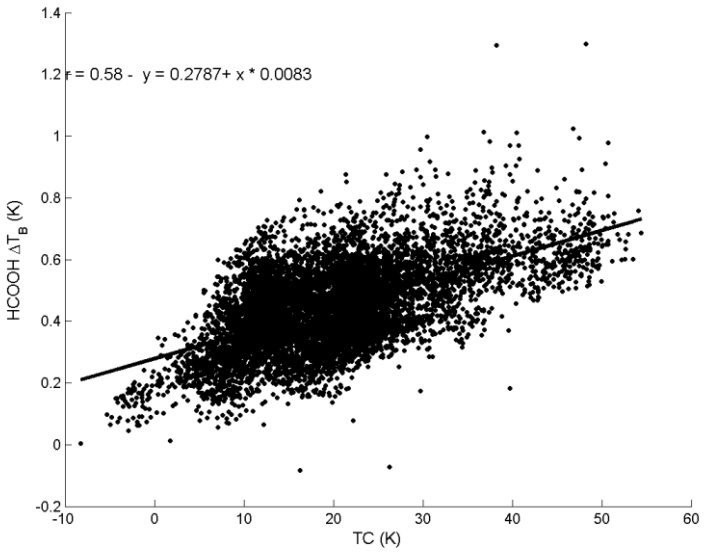

Fig. Scatter plot between the simulated ΔT_b and the TC=Tsurf-T3km, for one fixed HCOOH total column (0.6×10^{16} molec/cm^2).

ii. The last two sentences in this section appear contradictory. If rejecting negative values introduces a bias, then why are you rejecting negative values?

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 new sentence (in bold) reads:

"Note that this conversion could lead to negative total columns. If we eliminated all the negative values and kept only all the positive values, we would introduce an artificial bias in the average. **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**."

c.Section 2.3.2

i.The final paragraph in this section is very interesting. My experience with retrievals is that it is the lack of sensitivity to small changes in background amounts that leads to the very large errors on these values. It would be very useful to users of this data if the authors could provide an estimate of the algorithm sensitivity, i.e., what is the threshold detection value and how it varies with thermal contrast.

It is an interesting comment. We added this following information:

"Considering the detection threshold defined as 2σ on the ΔT_b (=0.30K), an indicative total column detection threshold was calculated using our conversion factors. To do so, forward simulations were performed for different total columns and TC. The result is illustrated in Fig. 6b and this shows that for the less favorable TC condition (TC=0K) the detection limit of HCOOH is close to 0.6×10^{16} molec/cm^2 (0.4×10^{16} molec/cm^2 for σ). This detection limit is improved with higher TC."

With this corresponding figure:

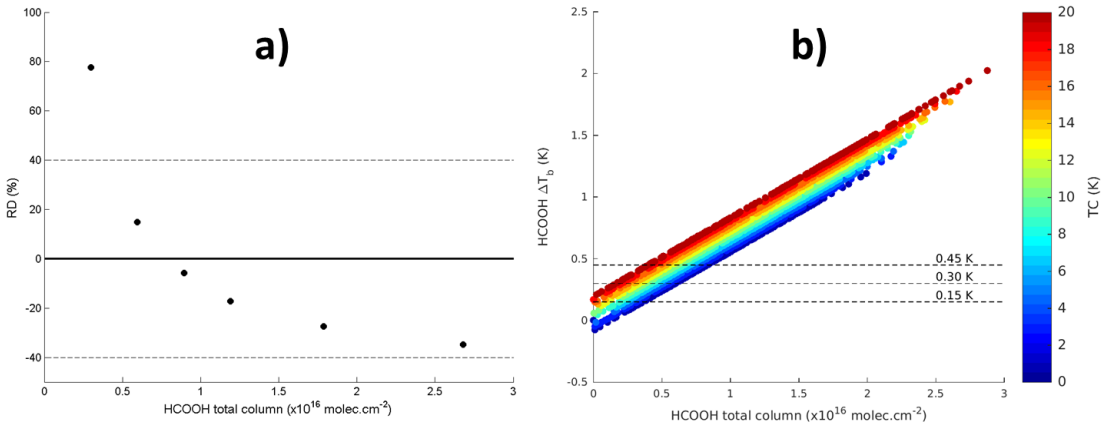

Figure 6. a: Variation of the mean relative difference between the total columns derived from the ΔT_b conversion and the a priori total columns (used as input in the forward simulations) according to the a priori used. The black solid line corresponds to a relative difference equal to 0 and the dashed black lines to ±40%. b: Variation of the simulated ΔT_b for different HCOOH total columns and TC. The dashed black lines correspond to a ΔT_b equal to 0.15K, 0.30K and 0.45K. 0.15 K corresponds to the IASI radiometric noise in the HCOOH spectral range (see Section 2.1).

d. Section 3.1

i.The authors use the result from section 2.3.2 (high errors on low amounts, lower errors on high amounts), to justify their lower results compare to Razavi. However, Razavi found a similar pattern in their data analysis, so I do not believe this is the correct explanation, or at least not the entire source of the lower values, which actually occur nearly everywhere. The authors should comment on why their results are in general significantly lower over most regions/periods with enhanced HCOOH.

It is a very good comment. That is why we added an additional explanation with this following sentence:

"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."

ii. The long list of features evident in Figure 8 should be written less like a list. The numbering of each discussion point is useful, but I think the points could be expanded on and better connected.
This part (lines 192-231 of the ACPD version) was rewritten and the list is now removed (see the last comment).

iii. The discussion on the Asian outflow is weak and unclear, as the IASI total columns are not provided for the PEM campaign periods/regions.
To clarify this point, we added the following details (in bold in the sentences):
"The Asian HCOOH outflow is well captured over the western Pacific (see label (10)). **The range of values** of the IASI total columns, from 2008 to 2014, broadly **agrees with our estimation of total columns using** the measurements from the aircraft PEM-West-B campaign conducted in February-March 1994 (Talbot et al., 1997a; 1997b) over a large region covering the latitudes 0°-60°N and the longitudes 110°-180°E. Indeed, measured HCOOH mixing ratio profiles during the campaign mostly ranged around 100-150 pptv from the boundary layer to about 12 km altitude, with peak values of up to 4 ppbv in fresh (< 2 days) plumes originating in China. Using these profiles, we estimated that this corresponded approximately to columns ranging from 0.2 to 0.9×10^{16} molec/cm^2 **while the IASI mean column is around 0.55×10^{16} molec/cm^2**. Over the remote Pacific, the IASI total columns, **for the studied period**, are larger than measured during the aircraft PEM-Tropics-A campaign in August-December 1996 (e.g. Talbot et al., 1999). They measured mixing ratios of the order of 20-40 ppbv in the boundary layer and 50-100 pptv in the free troposphere, corresponding to estimated total columns of $0.1\text{-}0.2\times10^{16}$ molec/cm^2. This overestimation is in agreement with the error budget from Fig. 6."

iv. A possible reason for the high values over India should be provided.
We added this information in bold:
"Over India, the largest total columns are observed from March to June **probably due to biomass burning (see label (7)). Indeed, those emissions present a marked seasonal variation with a maximum in March-May according to the GFED3 inventory (van der Werf et al., 2010), with 50%, 22% and 11% of annual emissions occurring in March, April and May, respectively**."

We also added this sentence (in bold) in Section 3.3.2:
"This phase difference is coherent with the cycle shown on the global distribution (Fig. 12). **The model underestimation over India suggests that either the emission factors of HCOOH (or its precursors) are underestimated in the model, or the biomass burnt estimated by GFEDv3 is too low for this region. The underestimation for forest fire emissions is a severe issue as shown by Chaliyakunnel et al. (2016) over tropical forests.**"

With the corresponding reference:
Chaliyakunnel, S., Millet, D.B., Wells, K.C., Cady-Pereira, K.E., and Shephard M.W.: A Large Underestimate of Formic Acid from Tropical Fires: Constraints from Space-Borne Measurements, Environ. Sci. Technol., DOI: 10.1021/acs.est.5b06385, 2016.

e. Section 3.2
i.This section requires at least some description of the FTIR: spectral resolution, noise level, sensitivity.
We added these lines (in bold) and now it reads:

"The IASI HCOOH retrieved columns in this work have been compared with ground-based FTIR measurements. This comparison was done without smoothing the data since the averaging kernels (AKs) were not provided by our retrieval method. This comparison is presented at four sites: Jungfraujoch (46.55°N 7.98°E) in Switzerland, Wollongong in Australia (34.41°S 150.88°E), Saint-Denis (20.88°S 55.45°E) and Maido (21.07°S 55.39°E) at La Réunion (Fig. 9). The current retrieved columns have also been evaluated with those using the methodology from Razavi et al (2011) over the same sites (Fig. 10).
"A complete description of the FTIR instruments and the retrieved HCOOH data can be found in Zander et al. (2010), Paton-Walsh et al. (2005), and Vigouroux et al. (2012), for the Jungfraujoch, Wollongong and Saint-Denis stations, respectively. For the Jungfraujoch, spectra were typically recorded at spectral resolutions of 0.004 and 0.006 cm^{-1}. For the present subset, a mean signal-to-noise ratio (SNR) of 895 was computed, with 10-th and 90-th percentiles of 525 and 1525, respectively. A uniform scaling of the HCOOH a priori was performed, and no information was available regarding the sensitivity of the retrieval with altitude. For Wollongong, the spectral resolution was 0.004 cm^{-1} and the SNR was around 1000-2000. Over La Réunion, the HCOOH retrieval parameters were the same for Saint-Denis and for the more recent data at the Maïdo station. The spectral resolution of La Réunion spectra was 0.007 or 0.011 cm^{-1}, depending on the time of the measurement. The SNR was about 1000-2000 depending on the spectra."

ii. The paragraph starting at line 240 is especially confusing, as the OEM results are not shown. A plot or table would be helpful.
These details (below) are now available in a supplement.

Table S1 Mean bias in 10^{16} molec/cm^2, normalized mean bias in percent (in parentheses), between the daily FTIR measurements and the daily IASI co-located mean data, retrieved by OEM and between the daily FTIR measurements and the columns using the conversion factors. The IASI data were collocated within 0.5° of the site location. The number of coincidence days is given for each site. For this test, only the five first days of each month in 2009 were retrieved as done in the Section 2.3.2 retrieval approach.

	Saint-Denis (12 days)	Wollongong (13 days)	Jungfraujoch (10 days)
IASI column, OEM	0.31 (110.65%)	-0.04 (-13.11%)	-0.86 (-73.96%)
IASI column, converted	0.26 (91.72%)	-0.05 (-15.14%)	-0.8 (-69.24%)

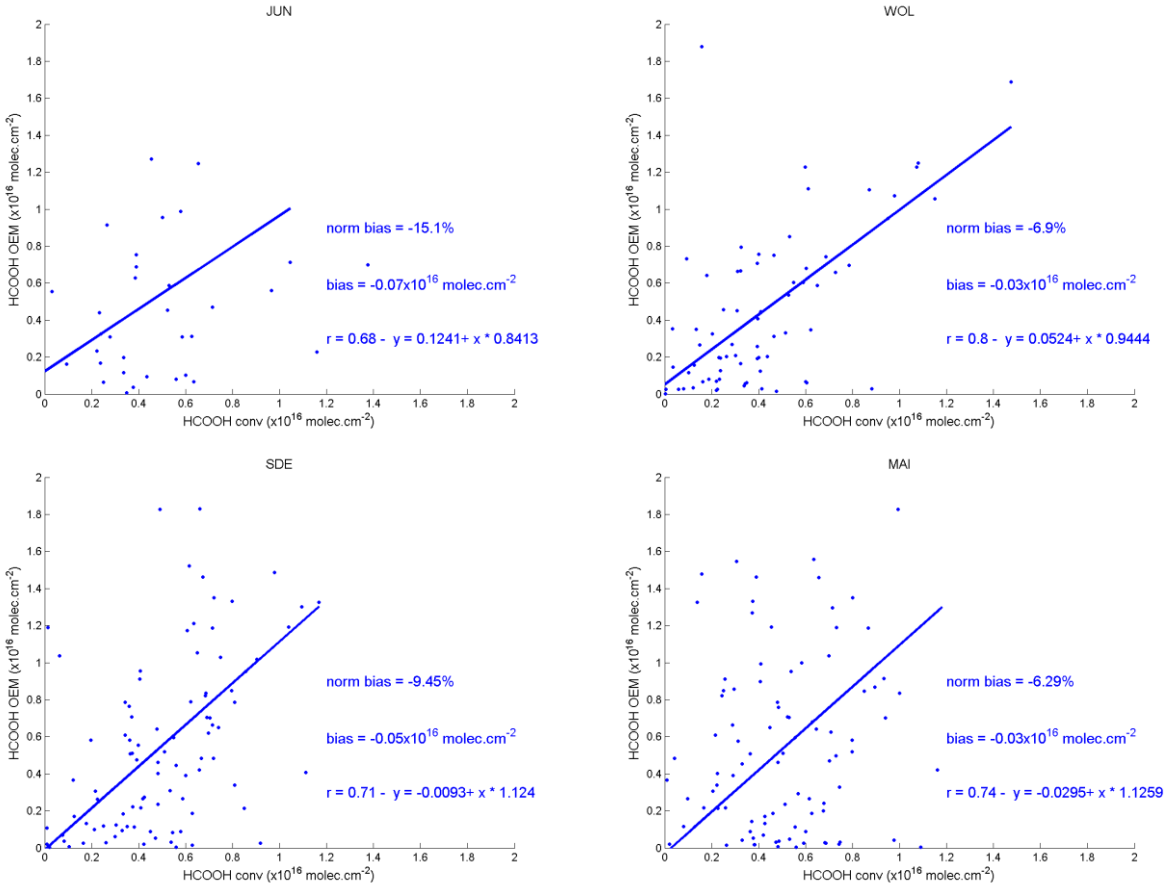

Figure S2. Scatterplot between the HCOOH columns retrieved by OEM (HCOOH OEM) and the corresponding columns calculated by the conversion factors (HCOOH conv). The IASI spectra were selected at ±0.5° around each site location (JUN = Jungfraujoch, WOL = Wollongong, SDE = Saint-Denis, MAI = Maido). The mean bias, the normalized mean bias, and the correlation coefficient are reported. The blue line corresponds to the linear regression and the corresponding equation is also provided. For this test, only the five first days of each month in 2009 were retrieved.

And we now refer to these results as (in bold):
"The current IASI retrieved columns were also compared with a set of columns retrieved by OEM around the sites. For each OEM-based retrieved column, the corresponding column using the conversion factors was calculated, showing that the current dataset and the OEM-based retrieval are in agreement (correlation ranging from 0.7 to 0.8, with an underestimation of the columns calculated with the conversion factors between 6 and 15%) **(Fig. S2).** It is also worth noting that similar biases were found between the columns retrieved by OEM around the ground-based locations and the FTIR columns as between the columns retrieved in this work and the FTIR ones **(Tab. S1).**"

iii. Why does the FTIR AK peak higher and have a broader peak?
It is normal to not have the same peak between a ground-based FTIR and a satellite FTIR. Both instruments are measuring in the IR but the spectral resolution is different and the geometry is different.
The DFS for the FTIR is close to 1, meaning that the signal corresponds to a column. Since 98% of the HCOOH is located below 20 km of altitude, it means that the FTIR mainly measured the column between the surface and 20 km.

iv. Where does equation 3 come from?
The columns are scaled with a height factor of 7.4. This 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 the sentence to the paper "This simplified formula is a variation of the hypsometric equation (Wallace and Hobbs, 1977)" with the corresponding reference.

2. Minor changes
a. Introduction
Line 24: ...dependence on thermal contrast is taken into account...
Corrected

Line 31: ...highlights the difficulty of retrieving total columns from IASI measurements over mountainous regions...
Changed

Line 48: ..., and to a lesser extent through oxidation by the OH radical.
Done

Line 61: ... the existence of unknown direct fluxes of HCOOH.
Changed

Line 63: Nadir looking atmospheric sensors can derive global distributions of trace gases, but with less vertical sensitivity than airborne or some ground-based measurements, such as FTIR instruments. Their extended spatial coverage allows for observations over remote regions ....
We prefer to keep the previous version: "Nadir looking atmospheric sensors allow to derive…".

Line 69: Are the ACE data also monthly?
No, per season. We then 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)."

Line 70: ... a low radiometric and high spatial coverage. HCOOH is a weak absorber, so it is challenging to ...
We changed this by "HCOOH is a weak absorber, so it is a challenge to …"
This change was requested by Referee 1.

Line 75: Which discrepancies? Please elaborate.
An additional sentence was added:

"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})."

Line 77: ..., suitable for both enhanced and background ....
"large" was replaced by "enhanced" as requested.

b.Section 2
Line 87: ...October 2006 and has provided more than eight years ...
Changed

Line 88: ...September 2012. Owing to their wide swath each instrument ...
Changed. Now it reads: "Owing to its wide swath, each instrument delivers near global coverage twice per day at around 9:30 local time (AM and PM)"

Line 93: Suggested rewriting of this paragraph: Analysis of the mean of the normalized (by what??) Jacobians (Fig. 2) over the spectral range used by IASI for the HCOOH retrievals 1095-1114cm-1(Is this correct ? Later in the text the authors state they use the channels at 1103, 1105 and 11909 cm-1.) for a set of representative geographical regions (see Fig. 1 and next section) shows that IASI is sensitive to tropospheric HCOOH signal between 1 and 6 km.
Yes, the retrieval was done for the spectral range between 1095 and 1104 cm^{-1}, but the calculation of ΔT_b used the channels at 1103, 1105 and 1109 cm^{-1}.
Normalized Jacobians means Jacobians integrated vertically.
The sentence has been changed as below:
"Analysis of the mean of the normalized Jacobians (Fig. 2) over the spectral range used by IASI for the HCOOH retrievals (1095-1114 cm^{-1}), for a set of representative geographical regions (see Fig. 1 and next section), shows that IASI is sensitive to tropospheric HCOOH between 1 and 6 km."

Line 103: ...columns of HCOOH using a set of conversion factors derived from OEM retrievals.
Now, the sentence is: "In a second step, the ΔT_b were converted into total columns of HCOOH using conversion factors derived from a set of data retrieved by OEM."

Line 127: ...remote areas
Corrected

Line 156: ...of the method are ...
It was corrected and it is "…the drawbacks of the method are.."

Line 156: ...and the lack of an error budget.
It was added

Line 159: ... OEM retrievals. To provide an estimate of the algorithm error simulations were performed ...
The sentence has been modified.

Line 170: ..., Razavi et al. (2011), who find a mean RD ...
Changed. We wrote "Razavi et al. (2011), who found a mean RD…"

c. Section 3

Line 190: ...Equator, with the highest values between 0-10°N, but with large variability, as the maximum was $3.5x10^{16}$molec/cm^2, but the monthly mean in this region was only $0.5x10^{16}$molec/cm^2.

The sentence has been modified. Now it reads:

"They showed a gradient of columns from the Poles to the Equator, with the highest values between 0 and 10°N, but with large variability, as the maximum was $3.5x10^{16}$ molec/cm^2, but the monthly mean in this region was only $0.5x10^{16}$ molec/cm^2".

Line 192-231: As noted above, please make this section less list-like. Some specific changes:

Line 192: A number of features are evident and are discussed below:

(1) A particularly striking feature are the large hotspots over Russia ...

(7) These states are flagged as biogenic emission regions…

The paragraph now reads:

"Yearly global distributions between 2008 and 2014 with the updated dataset are also presented in Fig. 8 (on a 1°×1° grid).

These distributions highlight well the recurring source regions detected by IASI such as Equatorial Africa, the North of Australia, Amazonia and India, and also the long-range transport such as over the Atlantic Ocean from Africa. The long-range transport over oceans (Atlantic, Indian and Pacific) was not investigated in Razavi et al. (2011). The retrieved columns over the Atlantic Ocean are consistent with the FTIR data from ship cruises reported in the study of Paulot et al. (2011). They showed a gradient of columns from the Poles to the Equator, with the highest values between 0 and 10°N, but with large variability, as the maximum was $3.5x10^{16}$molec/cm^2, but the monthly mean in this region was only $0.5x10^{16}$molec/cm^2.

Several hotspots and distributions are detected and are numbered from (1) to (10) in Fig. 8.

A particularly striking feature is the large hotspot over Russia (close to Moscow) in 2010 as documented by R'Honi et al. (2013), due to intense forest fires during the summer and also in 2012 over Siberia (see label (1)). The current dataset presents a mean total column twice lower (2.0×10^{16} molec/cm^2) than the mean derived using the conversion from Razavi et al. (2011) (4.2×10^{16} molec/cm^2), within the emission area (50-55°N, 30°-70°E), on 27 July - 27 August 2010, in agreement with the conclusions from R'Honi et al. (2013). Over Russia, other large columns were also found over Sakha Republic and over Khabarovsk Krai, in 2008 and 2012 (see label (2)). It is also worth noting that the North American boreal emissions around the Hudson Bay were larger between 2008 and 2010 compared to other years (see label (3)). Over North America, and especially the US, we observe lower columns over Louisiana and Texas in 2008 and in 2014 compared to the other years (see label (4)); while larger total columns were measured over Northern Australia between 2012 and 2014, in comparison to the period from 2008 to 2010 (see label (5)).

The monthly means over the seven years are also presented with an animation (Fig. S1) in the Supplement. As already observed over Eastern Russia with label (2), in June 2010 and 2012, there were large concentrations, close to Khabarovsk Krai, compared to the other years in this region. Whereas intense fires were detected in June 2012 in this region, this was not the case in June 2010 (see maps on http://lance-modis.eosdis.nasa.gov/cgi-bin/imagery/firemaps.cgi). The absence of forest fires and the lack of hotpots in the biogenic emission inventory (Fig. 1) in this region points to the presence of an unidentified source, possibly of biogenic origin.

Large columns were similarly retrieved over a large region encompassing Laos, Thailand and Myanmar in April 2010, 2012, 2013 and 2014 (see label (6)). It matches well with the locations of fire hotspots detected by MODIS.

Over India, the largest total columns are observed from March to June probably due to biomass burning (see label (7)). Indeed, those emissions present a marked seasonal variation with a maximum in March-May according to the GFED3 inventory (van der Werf et al., 2010), with 50%, 22% and 11% of annual emissions occurring in March, April and May, respectively.
Larger total columns were retrieved in August 2010 along the Euphrates River compared to the other years (see label (8)).
The monthly distributions also highlight hotspots over the US, besides those shown in the annual distributions (see label (9)). In summer 2011, large signatures over the US were not confined to coastal regions; high total columns were also detected in the Mid-Western US such as over Kansas, Mississippi, Missouri or Oklahoma. These states are flagged as biogenic emission regions of VOCs by Millet et al. (2015), acting as secondary source of HCOOH. In July 2012, the emissions over the US were mostly confined to the Eastern part.
The Asian HCOOH outflow is well captured over the western Pacific (see label (10)). The range of values of the IASI total columns, from 2008 to 2014, broadly agrees with our estimation of total columns using the measurements from the aircraft PEM-West-B campaign conducted in February-March 1994 (Talbot et al., 1997a; 1997b) over a large region covering the latitudes 0°-60°N and the longitudes 110°-180°E. Indeed, measured HCOOH mixing ratio profiles during the campaign mostly ranged around 100-150 pptv from the boundary layer to about 12 km altitude, with peak values of up to 4 ppbv in fresh (< 2 days) plumes originating in China. Using these profiles, we estimated that this corresponded approximately to columns ranging from 0.2 to 0.9×10^{16} molec/cm^2 and the IASI mean column is around 0.55×10^{16} molec/cm^2. Over the remote Pacific, the IASI total columns, for the studied period, are larger than measured during the aircraft PEM-Tropics-A campaign in August-December 1996 (e.g. Talbot et al., 1999). They measured mixing ratios of the order of 20-40 ppbv in the boundary layer and 50-100 pptv in the free troposphere, corresponding to estimated total columns of $0.1\text{-}0.2\times10^{16}$ molec/cm^2. This overestimation is in agreement with the error budget from Fig. 6.
Overall, these monthly means highlight the seasonal variation of the HCOOH distribution around the world. The animation reveals clear variations in the HCOOH distribution due to the seasonality of biomass burning and vegetation growth. It is well shown with the large total columns observed during September and October 2008, 2012, 2014 in the Southern Hemisphere (over Amazonia, Africa and Australia). In 2010, the same features were noted except for Australia."

The numbers on Fig.8 were also changed:

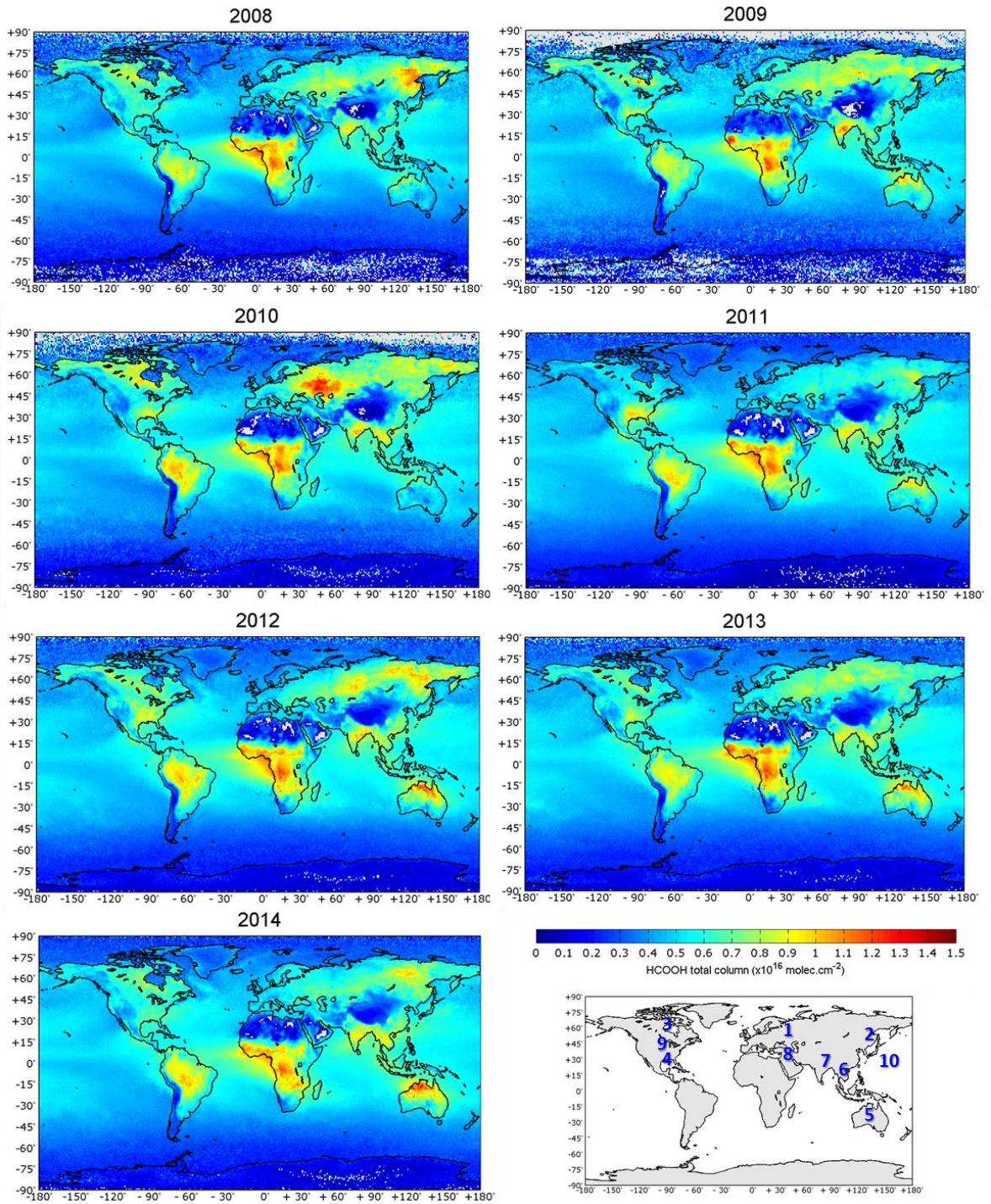

Figure 8. Annual HCOOH global distribution from 2008 to 2014, derived using the IASI radiance observations on a 1°×1° grid. Different sources or distributions described in the text are numbered in blue on the bottom-right map.