"Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations" by I. De Smedt et al.

Response to reviewer #1

This is an interesting and relevant publication, which I recommend to be published in ACP. The manuscript describes in great detail the retrieval algorithm used to infer tropospheric formaldehyde (HCHO) columns from the OMI and GOME-2 satellite instruments. The derived products are then evaluated and analyzed for temporal and spatial variability and trends.

The authors provide a clear description of the products and tools used and the analysis applied. The paper is very well written and easy to read. Accompanying their analysis the authors give a good demonstrations of the limitations of the products and their interpretation while at the same time highlighting the value and possible applications of the data set. This is very important information for users of satellite derived HCHO products.

Their analysis also yields some interesting scientific results that reflect the changes in human activity on tropospheric chemistry such as emission reductions in Western Europe and the Eastern US or changes due to deforestation in the Amazon.

We thank the reviewer for their supportive comments. We have provided detailed responses to comments raised and have adjusted the manuscript accordingly where appropriate.

I have a few comments and questions I appreciate if the authors could address:

(1) I am confused whether the SCIAMACHY time series was also processed with the new algorithm. In Section 6 GOME-2 and SCIAMACHY are used combined for analysis, so I assume both rely on the same retrieval algorithms? A good agreement between the two products is mentioned on Page 12265, line 5 with reference to Fig 14-16, but to me the graphs do not seem to make this clear. Have there been any more comprehensive comparisons to ensure consistency between the different products?

The SCIAMACHY and GOME-2 time series were not retrieved using exactly the same algorithm, because it is not possible. More precisely, the pre-fit of O4 and BrO is not possible with SCIAMACHY because of a spectral polarisation structure around 360 nm. However, H_2CO SCD are finally retrieved in the same interval (328.5-346 nm), and using the same external datasets (as mentioned on page 12250, lines 10-13). The air mass factor calculation is done using the same algorithm. The consistency between the SCIAMACHY and GOME-2 retrievals has been presented in details in De Smedt et al.; 2012 (Figure 12).

In figures 14-16, the morning time series combine SCIAMACHY (2003-2011), GOME-2A (2007-2013) and B (from 2013) measurements. Therefore, the morning observations from 2004 to 2006 are only composed by SCIAMACHY. We used the mean of SCIA and GOME2 columns between 2007 and 2011. From 2012 onwards, only GOME-2 data can be used, as the SCIAMACHY time series stops in Mar.2012. These details have been added in the description of the figures.

(2) Also related to Fig 14-16 I wonder if the authors have an explanation for why the trend series over California is not significant for OMI and only significant for SCIA/GOME-2. For all other regions

either both or only OMI show significance which as they mention is reasonable giving the higher HCHO columns at the OMI overpass.

We agree with the reviewer, and unfortunately, we don't have a good explanation. In the morning observations, we constantly observe an increase of H_2CO columns over California. With GOME and SCIAMACHY, we found for the period 1997-2009 (De Smedt et al. 2010):

- Los Angeles: 0.7 ± 0.3 x10^{14} molec.cm^{-2} (2.2% ± 0.9)
- San Francisco: 1.0 ± 0.3 x10^{14} molec.cm^{-2} (4.3% ± 1.6)

Similar values are now found between 2004 and 2014, adding GOME-2 to the morning time series (1.1 ± 0.3 x10^{14} molec.cm^{-2}, 3.3% ± 1.0).

The afternoon OMI time series does not show this positive trend, which is more in line with what is expected from anthropogenic emissions. Simulations with a 3D CTM are needed in order to assess the impact of, for example, fire events or temperature changes on biogenic NMVOC emissions (Stavrakou et al., 2014). We cannot go further in the interpretation of the observations at this stage.

(3) Fig 14-16 as referenced in the text are not in line with the Figure order.

Thanks. Corrected.

(4) Section 3.2: I could not quite follow how this correction is applied. It is stated that the median column over the Pacific is subtracted from the slant columns together with a polynomial latitudinal fit and then replaced by the latitudinal dependence of modelled HCHO columns (the same model as used for a priori I assume). Would one expect that over the Pacific the HCHO columns are then near-zero? And why are the corrected columns larger than the uncorrected columns. It would help if this part is rewritten to describe the individual steps in a very clear way.

The reference sector correction has been described more in details in our previous papers (De Smedt et al., 2008 and also in De Smedt et al., 2014).

As daily radiance spectra (selected in the Equatorial Pacific) are used as background spectra in the DOAS fit, the retrieved slant columns actually represent the difference in slant column with respect to the slant column contained in the reference spectra. The SCD (N_s) are therefore expected to be equal to 0 in the Equatorial Pacific region. However, this is not always the case, because of spectral artefacts, that can result in positive or negative biases in the SCD, and that depend on the latitude and season.

In the equation: $N_v = \frac{\Delta N_s}{M} + N_{v,0,CTM}$

ΔN_s are indeed forced to 0 in the Pacific Ocean, by the reference sector correction. Then, the vertical columns of the IMAGES model in the reference sector ($N_{v,0,CTM}$) are added. This background value of H_2CO, due to methane oxidation, ranges from 1 to 4x10^{15} molec.cm^{-2}, depending on the latitude and on the season. It presents a smooth variation with the latitudes.

We added more details in section 3.2.

(5) Figure 9: It would help to adapt the colorscale of this Figure to show the same scales for both GOME-2 and OMI yet resolve hotspots.

We would prefer to keep different color scales for this figure, because the aim is to show in particular the effect of the improved spatial resolution of OMI. Both instruments are able to detect hotspots, but with a difference in columns of about $2x10^{15}$ molec.cm^{-2} (effect of horizontal resolution, but also of diurnal variation and difference in sensitivity). If we change the color scale, hotspots will disappear in the GOME-2 map.

References

De Smedt, I., Van Roozendael, M., Danckaert, T., Van Gent, J., Theys, N., Lerot, C.: TROPOMI/S5P ATBD of Formaldehyde data product, S5P- BIRA-L2-400F- ATBD, 2014.

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., Valks, P., Hao, N., and van der A, R.: Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos. Meas. Tech., 5:2933-2949, Special Issue: GOME-2: calibration, algorithms, data products and validation, 2012.

De Smedt, I., Stavrakou, T., Müller, J. F., van Der A, R. J. and Van Roozendael, M.: Trend detection in satellite observations of formaldehyde tropospheric columns, Geophys. Res. Lett., 37(18), L18808, doi:10.1029/2010GL044245, 2010.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H. and Van Roozendael, M.: twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8(16), 4947-4963, 2008.

Response to reviewer #2

This paper provides a consistent long time record of satellite HCHO data from OMI and GOME-2 using an updated version of retrieval algorithm. The authors then validate the new HCHO retrievals using global ground-based DOAS HCHO observations regarding total HCHO columns, HCHO vertical profiles as well as seasonal and diurnal variations in HCHO columns. Finally, interesting global and regional trends in satellite HCHO columns are examined and identified based on the validated retrievals. The paper is generally well written and describes space-based observations and trend clearly and thoroughly.
I think it is certainly appropriate for publication subject to the following additions/modifications.

We thank the reviewer for their supportive comments. We have provided detailed responses to comments raised and have adjusted the manuscript accordingly where appropriate.

1. It needs to clarify whether aerosol scattering has been taken into consideration in computing scattering weights. If not, why? Along the same line of scattering weights, could the authors comment on uncertainties due to cloud? And to what extent could (changing) aerosol load impact the retrievals and eventually affect the HCHO trend seen from space in China and India?

Aerosol scattering is not taken into account in the scattering weight LUT, because aerosol scattering effects are very difficult to separate from cloud effects, and because the needed information on aerosol properties is not available at the global scale. A full treatment of aerosols in radiative transfer

would be possible if clouds are aerosols were represented separately as scattering layers and if detailed information on aerosol optical properties was available at the global scale. However, the aerosol impact is partially accounted for by the cloud correction scheme. Indeed, to a large extent, the effect of the non-absorbing part of the aerosol extinction is implicitly included in the cloud correction, because the LER cloud algorithm is expected to overestimate the cloud fraction in the presence of aerosols (Boersma et al., 2004; 2011). Furthermore, observations with cloud fractions exceeding 40% are filtered out, excluding a large part of the high AOD scenes (Theys et al., 2015).
In Leitao et al. (2009), the aerosol effects on the retrieval of tropospheric trace gases using UV/visible nadir measurements were investigated. They showed that, when realistic vertical profiles are applied, the aerosol effect has a relatively small impact. Theys et al. (2015) have also performed simulations with LIDORT using anthropogenic SO_2 and aerosols profiles retrieved from ground-based MAX-DOAS measurements performed in Beijing and Xianghe (Clémer et al., 2010). The results confirmed the limited effect of the aerosols on the air mass factors (maximum 15% in average, for very high aerosol loading). The explanation is a compensation of enhancement and reduction of the trace gas signal due to aerosols, because the vertical profiles of tropospheric gases and aerosols are mixed.
Finally, a trend in aerosols should result in a trend in the cloud fractions because the LER cloud product is sensitive to scattering aerosols. We have done a sensitivity trend study on the cloud fractions and on the tropospheric AMFs. We find a positive trend in cloud fractions over Beijing and India, but also a positive trend in the cloud altitudes. These effects have respectively a positive and a negative impact on the tropospheric AMFs, and no trend is observed in AMFs. Furthermore, the positive trends are already present in the H_2CO slant columns.

In section 3.3.1, the following sentence has been added: “No explicit correction is applied for aerosols but the cloud correction scheme accounts for a large part of their scattering effect (Boersma et al., 2011). The uncertainty related to aerosol effects is estimated to be lower than 15% in average (Leitao et al., 2008; Theys et al., 2015).”

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., et al.: An improved tropospheric NO2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4(9), 2011.
Castellanos, P., Boersma, K. F., Torres, O. and de Haan, J. F.: OMI tropospheric NO2 air mass factors over South America: effects of biomass burning aerosols, Atmos. Meas. Tech., 8(9), 3831–3849, doi:10.5194/amt-8-3831-2015, 2015.
Clémer, K., M. Van Roozendael, C. Fayt, F. Hendrick, et al. (2010), Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, Atmos. Meas. Tech. Discuss., 3(1), 111-145.
Leitao, J., A. Richter, M. Vrekoussis, A. Kokhanovsky, et al. (2009), On the improvement of NO2 satellite retrievals - aerosol impact on the airmass factors, Atmos. Meas. Tech. Discuss., 2(6), 3221-3264.
Theys, N., De Smedt, I., van Gent, J., Danckaert, T., Wang, T., Hendrick, F., Stavrakou, T., Bauduin, S., Clarisse, L., Li, C., Krotkov, N., et al.: Sulfur dioxide vertical column DOAS retrievals from the Ozone Monitoring Instrument: Global observations and comparison to ground-based and satellite data, J. Geophys. Res. Atmos., 2014JD022657, doi:10.1002/2014JD022657, 2015.

2. In my opinion, it would be great to have some comparison between the updated HCHO retrievals from this paper and other available HCHO products, e.g., OMI HCHO from González Abad et al. [2015] and OMPS HCHO from Li et al. [2015]. Some numbers in hotspot regions and background from different products are enough.

Those 2 recent papers include comparisons between our OMI HCHO product and respectively the SAO OMI and OMPS HCHO products (Gonzalez Abad et al., 2015b) and the NASA OMPS HCHO product (Li et al., 2015). We provide here a summary of their comparison results. References to their

papers have been added in the manuscript. However, we would prefer to keep a more in depth comparison of the different satellites products for another work/publication.

Figure 6 in the recent paper of Gonzalez Abad et al. (2015b) show comparisons between BIRA OMI, and SAO OMI H_2CO columns (here in Table 1). Seasonal changes and spatial patterns are similar between the two OMI products. The quantitative agreement is within 25% during peak seasons, while relative differences are larger for mid to low H_2CO columns, from 50 to 80% in Indonesia.

Figure 3 and Table S1 in Li et al. (2015) compares the seasonal patterns of BIRA OMI and OMPS PCA H_2CO (here in Table 2). In general, the two H_2CO retrievals demonstrate similar seasonal changes and spatial patterns. Quantitatively, the OMPS PCA H_2CO retrievals are about 15–20% smaller than those from BIRA OMI in several source regions during the peak season. For example, the regional mean H_2CO for the southeast U.S. for August 2013 is ~20% smaller than BIRA OMI H_2CO. For months with lower H_2CO loading, the two retrievals generally agree better in the tropics than in the middle latitudes. This may reflect greater uncertainty in both retrievals under more challenging conditions during winter such as larger solar zenith angle and weaker signals. Despite these differences, the overall agreement between the OMPS and OMI retrievals is encouraging and should lend confidence to both products, given that they are independently produced from two different instruments using fundamentally different approaches (PCA versus DOAS). For example, OMPS has lower spatial and spectral resolution but higher signal-to-noise ratio than OMI.

The input data are different between the three algorithms, including the a priori H_2CO profiles, surface reflectivity, cloud data, and also background H_2CO corrections. More detailed analyses will be necessary to reconcile the quantitative differences seen between the data sets, but the level of agreement is within the reported error bars of the products.

Table 1: Comparison between monthly averaged OMI H_2CO columns retrieved at BIRA and at SAO (from Gonzales Abad et al., 2015b)

Region	Period	BIRA OMI H_2CO	SAO OMI H_2CO	Relative Differences
South Eastern US	Jan. 2013	5.00 ± 2.50	2.70 ± 2.50	-0.46 ± 0.74
	Aug. 2013	10.00 ± 3.00	8.00 ± 3.00	-0.20 ± 0.43
Amazon	Apr. 2013	5.10 ± 3.00	6.00 ± 3.00	0.18 ± 0.84
	Aug. 2013	13.00 ± 3.00	10.00 ± 3.00	-0.23 ± 0.33
Equatorial Africa	Jan.2013	11.00 ± 2.00	10.00 ± 2.00	-0.09 ± 0.26
	Aug.2013	5.50 ± 3.00	7.00 ± 3.00	0.27 ± 0.79
Indonesia	Mar.2013	6.00 ± 3.00	9.00 ± 3.00	0.50 ± 0.75
	Aug.2013	4.60 ± 4.00	8.20 ± 4.00	0.78 ± 1.41
Northeast China	Jan. 2013	5.00 ± 2.50	6.00 ± 2.50	0.20 ± 0.71
	Aug.2013	15.00 ± 3.00	11.00 ± 3.00	-0.27 ± 0.29

Table 2: Comparison between monthly averaged OMI H_2CO columns retrieved at BIRA and OMPS H_2CO columns retrieved at NASA (from Li et al., 2015b)

Region	Period	BIRA OMI H_2CO	NASA OMPS PCA H_2CO	Relative Differences
South Eastern US	Jan. 2013	5.8 ± 2.1	0.6 ± 1.0	-0.90 ± 0.52
	Aug. 2013	9.3 ± 3.1	7.4 ± 2.3	-0.20 ± 0.42
Amazon	Apr. 2013	5.1 ± 2.9	5.4 ± 1.4	0.06 ± 0.63
	Oct. 2013	9.9 ± 5.3	8.4 ± 2.0	-0.15 ± 0.58
Tropical Africa	Apr.2013	6.2 ± 3.0	6.5 ± 1.4	0.05 ± 0.53
	Aug.2013	9.6 ± 2.7	9.2 ± 2.0	-0.04 ± 0.35
Southeast Asia	Apr.2013	12.3 ± 4.3	9.3 ± 3.8	-0.24 ± 0.47
	Aug.2013	4.4 ± 3.0	4.9 ± 1.6	0.11 ± 0.78

Eastern China	Jan. 2013	6.0 ± 2.8	0.7 ± 1.8	-0.88 ± 0.69
	Aug.2013	10.2 ± 2.8	8.5 ± 2.8	-0.17 ± 0.39

González Abad, G., Vasilkov, a., Seftor, C., Liu, X. and Chance, K.: Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval, Atmos. Meas. Tech. Discuss., 8(9), 9209–9240, doi:10.5194/amtd-8-9209-2015, 2015b.
Li, C., Joiner, J., Krotkov, N. A. and Dunlap, L.: A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophys. Res. Lett., 2515–2522, doi:10.1002/2015GL063204.Received, 2015.

3. I don't understand why GOME and SCIAMACHY HCHO retrievals are necessarily consistent with new OMI and GOME-2 retrievals. GOME and SCIAMACHY data were retrieved using old algorithm, if I understand correctly. Will difference in retrieval algorithms lead to uncertainties when combine all the data in looking at the long-term trend?

- The algorithms are as consistent as possible between sensors. The main difference is that 3 intervals are used for OMI and GOME2, O_4 and BrO are pre-fitted, while HCHO is retrieved in 328.5-346 nm. For GOME and SCIAMACHY, the quality of the recorded spectra does not allow to use the pre-fit windows, but HCHO is retrieved is the same window as for OMI and GOME-2 (328.5-346). The other retrieval parameters are the same (cross-sections, albedo, a priori profiles, …). GOME and SCIAMACHY HCHO datasets have been reprocessed several times since the 2008 publication (De Smedt et al, 2008).
- Remaining differences between HCHO slant columns are corrected by the reference sector correction step. This removes global offsets, artificial zonal dependencies, and global long-term degradation effects.
- Consistency between GOME, SCIAMACHY and GOME-2 is addressed in De Smedt et al., 2012. It is not perfect everywhere but very satisfactory.
- To our point of view, the main reason for inconsistencies between the satellite columns is the large differences in spatial resolution. GOME pixels being much larger than the other sensors.
- In this paper, GOME is not included in our trend analysis. SCIAMACHY/GOME-2 and OMI datasets are treated separately. Furthermore, in the trend model, offset terms between SCIAMACHY and GOME-2 are fitted per region to account for possible effects of spatial resolution between the two time series (De Smedt et al., 2010; Hilboll et al., 2013). For these reasons, we think that if the trends are found consistent between morning and afternoon observations, they can be trusted.

4. GOME-2 has a larger footprint that OMI. Does this mean that GOME-2 is easier to be affected/contaminated by cloud? If so, it may not be fair to directly compare OMI and GOME-2 retrievals. This might also lead to bias in getting reliable diurnal HCHO profile, which is one of the main purposes of this paper, bracketed by GOME-2 and OMI. Can the authors comment on this?

We have performed several sensitivity tests to obtain the diurnal variations from GOME-2 and OMI columns. We have compared the cloud free observations to the cloud corrected columns. The differences observed between GOME-2 and OMI are almost identical, all the more so as the averaging period is several years. As stated in the paper, cloud effects are more random than systematic, and are not thought to influence systematically the averaged observed diurnal variations.

It is clearly acknowledged in the paper that uncertainties on the satellites H_2CO columns are large and therefore also the errors on their differences (as reflected in Table 4). Nevertheless, the sign of the differences between GOME-2 and OMI agrees well with both MAX-DOAS and FTIR

measurements. Early afternoon values are almost always equal to or larger than mid-morning values, except in Bujumbura where morning columns are larger.

5. I don't understand why 100 km is used in getting average HCHO columns. Some back-of-envelope calculations based on HCHO lifetime and local annual wind speed will be appreciated.

A radius of 100 km is used to compare satellite columns to ground-based measurements. For this kind of comparison, and unlike inversion exercises using models, the smallest possible radius should be used. Using 100 km is a pragmatic choice, because using a smaller radius does not include enough satellite pixels to reduce sufficiently the noise of the data, particularly in the case of GOME-2. It is a trade-off between the spatial resolution and the quality of the H_2CO time series.
We added the following sentence in section 5: "Although larger than the typical length of air masses sampled by a MAX-DOAS spectrometer, which is less than a few tens of kilometres (Gomez et al., 2014), this radius allows including enough satellite pixels to ensure significant analysis."

Gomez, L., Navarro-Comas, M., Puentedura, O., Gonzalez, Y., Cuevas, E., and Gil-Ojeda, M.: Long-path averaged mixing ratios of O3 and NO2 in the free troposphere from mountain MAX-DOAS, Atmos. Meas. Tech., 7, 3373-3386, doi:10.5194/amt-7-3373-2014, 2014.

6. I'm not totally convinced by some potential driving factors of HCHO trend proposed in the paper. Temperature should be the dominated driver of the interannual variability (IAV) in HCHO columns. The authors may want to acknowledge more that trends in temperature would have played a role in the changes/trends observed. The authors may also want to clarify why the instrument aging only plays a negligible role in the observed trend.

We agree, and have tried to make it clearer in the revised version of the paper.

Formaldehyde columns are mainly formed by oxidation of VOCs, from biogenic, biomass burning and anthropogenic sources. Column inter annual variabilities are mainly driven by fire events and temperature changes (Millet et al., 2008; Barkley et al., 2009; Stavrakou et al., 2014). However, over industrialised regions, changes in anthropogenic emissions have also been identified as drivers of observed H_2CO column trends (De Smedt et al., 2010; Zhu et al., 2014; Khokhar et al., 2015; Mahajan et al., 2015; Stroud et al., 2015).

The instrumental degradation effects play a role in the long-term observations, which we try to take into account. The reference sector correction cancels any global variation of the H_2CO columns. In the case of OMI, we also take into account the change in sampling along the years, by using "sampling-corrected" OMI columns. Finally, the VCD errors are used in the trend analysis, and errors are provided with trend estimates. The results of our trend analysis are displayed only when they have been found statistically significant, and the fact that equivalent trends are observed independently with both datasets gives confidence in our H2CO column long-term variation estimates.

Some minor comments:

Page 12243, Line 17-18 "constrain NMVOC emissions in top-down inversion approaches" Not all of cited work involve inversion approaches, some of them instead simply assume linear relationship between HCHO columns and NMVOCs from a CTM to constraint NMVOC emissions.
Right. We then removed "in top-down inversion approaches".

Page 12245, Line 21, "Section 3" "Sect. 2", to be consistent.

Corrected

Page 12246, Line 4, “at 13:30 LT” “around 13:30 LT”, more accurate
Corrected

Page 12246, Line 13, “which enables daily global coverage” “which enables almost/ nearly daily global coverage”, more accurate
Corrected

Page 12247, Line 21, “Since 10 years . . .” Can the authors rewrite this sentence?
“Level 2 formaldehyde products developed at BIRA-IASB are provided via the TEMIS website. The algorithms used to generate these products are designed to be as consistent as possible, in order to optimise the overall coherency of the resulting time series.”

Page 12248, Line 11, “model a priori profiles is used at the” “model a priori profiles is sampled at the”, more accurate
Corrected.

Page 12255, Line 7, “The model time step is set to 4h” Can the authors clarify at which time period is GOME-2 and OMI a priori sampled?
The model is sampled at the satellite overpass time. This is realized by using correction factors applied to the photorates and the chemical rate constants which are derived from a full diurnal cycle model simulation with a time step of 20 minutes as described in Stavrakou et al. (2009a).

Page 12255, Line 26, “about 71 and 26 % of the global sink” HCHO loss due to OH oxidation and photolysis is about half to half. Can the authors say a little more about the two loss pathways? Maybe with some references?
At the end of section 3.3.2, we added the following:
“The photochemical production of H_2CO is estimated at ca. 1600 Tg annually. **The main formaldehyde sinks are the oxidation by OH (Sander et al., 2011) which leads to CO production and conversion of OH to HO_2, and two photolysis reactions which produce CO and HO_2 radicals. Based on IMAGESv2 model calculations, photolysis is by far the dominant removal process, estimated at 71% of the global sink, whereas the OH sink is less efficient (26%).** Dry and wet deposition account for the remainder (<3%). The global photochemical H2CO lifetime is estimated at 4.5 hours.”

Sander, S. P., J. Abbatt, J. R. Barker, J. B. Burkholder, R. R. Friedl, D.M. Golden, R. E. Huie, C. E. Kolb,M. J. Kurylo, G. K.Moortgat, V. L. Orkin and P. H. Wine : Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation No. 17, JPL Publication 10-6, Jet Propulsion Laboratory, Pasadena, 1320 2011 http://jpldataeval.jpl.nasa.gov.

Page 12260, Line 7-8, “has been developed to retrieve tropospheric trace gas total columns and profiles” Somewhere in the introduction part, the authors may want to clarify what they have retrieved are tropospheric HCHO columns.
Done: “This study focuses on **tropospheric** formaldehyde retrievals from OMI, using an algorithm historically developed within the TEMIS (Tropospheric Emission Monitoring Internet Service) framework and applied to morning observations from the GOME, SCIAMACHY and GOME-2 sensors (http://h2co.aeronomy.be).”

Page 12262, Line 27-28, “the slopes of the regression lines between” Is this RMA regression? It’s better be since you have errors in both X and Y. Also it seems to me that the intercept has been forced to zero, right? If so, why?

In the discussion paper, it is not RMA regression but a simple least squares solution, indeed forced to zero. This has been changed in the revised version. We now use a cubic least square regression, where the line is fit by minimizing both x- and y-residuals simultaneously for weighted data points. This analysis works well in Beijing/Xianghe, where we have good correlations both for GOME-2 and OMI, but not in Bujumbura, where the correlations are too low. Therefore, we now only provide mean differences in Bujumbura. A Table has been added with more detailed numbers for both stations.

Table 5: Results of the comparisons between GOME2, OMI and MAX-DOAS columns, shown in **Error! Reference source not found.** for Beijing/Xianghe and in **Error! Reference source not found.** for Bujumbura. Three satellite VCs are used: IMAGES a.p. profile/no cloud correction, IMAGES a.p. profile/IPA cloud correction, and MAX-DOAS a.p. profile/IPA cloud correction. Mean differences (satellite-MAX-DOAS) are given in both stations. In Beijing/Xianghe, the slopes and offsets of a linear regression between MAX-DOAS and satellite columns are provided.

	GOME-2			**OMI**		
	IMAGES ap profiles No cloud correction	IMAGES ap profiles Cloud correction	MAXDOAS ap profiles Cloud correction	IMAGES ap profiles No cloud correction	IMAGES ap profiles Cloud correction	MAXDOAS ap profiles Cloud correction
Beijing/Xianghe						
# common days	711			807		
Mean Difference 10^{15} molec.cm^{-2} (%)	-4.4 (-33%)	-5.4 (-41%)	-1.4 (-11%)	-2.8 (-19%)	-3.4 (-24%)	2.1 (15%)
Correlation coefficient	0.82	0.85	0.84	0.88	0.87	0.88
Slope	0.80	0.64	0.91	0.61	0.62	0.97
Offset 10^{15} molec.cm^{-2}	-1.6	0.6	0.2	3.0	2.3	3.2
Bujumbura						
# common days	58			90		
Mean Difference 10^{15} molec.cm^{-2} (%)	-2.8 (-29%)	-2.6 (-27%)	1.1 (11%)	-3.3 (-39%)	-2.6 (-31%)	-0.9 (-9%)

Page 12263, Line 11-12, “bringing the satellites and ground-based observations to a satisfactory agreement within 15%.” I’m not convinced. It seems to me that you have used information from observations in getting vertical profile. Of course you can get a better agreement.

Using the max-doas profile shapes to re-calculate the satellite AMFs is equivalent to smoothing the max-doas profiles with the satellite total column averaging kernels. It allows removing from the comparison the error associated to the *a priori* profile shapes (Eskes and Boersma, 2003).
However, only the vertical distribution of the a priori profile impact the satellite AMFs, not their integrated columns (Palmer et al., 2001). It means that with the same max-doas profile shapes but integrated columns twice larger, the impact on the satellite AMFs would be the same, but the final column agreement would not be good. Therefore, if successful, this type of comparison validates the satellite slant columns, and the satellite AMF dependencies with albedo, clouds, … (all factors expect vertical profile shapes).

We added this precision in the manuscript: "It must be noted that the retrieved MAX-DOAS profiles also have their own uncertainties (Vlemmix et al., 2014), however using them to re-calculate the satellite AMFs allows to remove from the comparison the error associated to the a priori profile **shapes** (Eskes and Boersma, 2003). **Indeed, only the shape of the a priori profiles impact the satellite AMFs, not their total columns (Palmer et al., 2001).** The satellite averaging kernels (AKs) are much closer in shape to the FTIR AKs than to the MAX-DOAS retrievals, which may explain the better agreement of the columns (Vigouroux et al., 2009)."

Page 12263, Line 24-26, "The effect of the rather coarse resolution of the global CTM on the modelled profiles (here 2.5°× 2.5°) needs to be further investigated, as well as possible other effects of vertical transport and chemical processes." Just a comment on this. The coarse CTM may be doing OK for regional background from biogenic sources (e.g, HCHO from isoprene) but may never be able to model urban/point source right on top of biogenic background.
Agreed. Regional models would be more appropriate.
We added to the sentence: "The effect of the rather coarse resolution of the global CTM on the modelled profiles (here 2°x2.5°) needs to be further investigated, particularly for anthropogenic sources, as well as possible other effects of vertical transport and chemical processes."