"Development and characterisation of a state-of-the-art GOME-2 formaldehyde air-mass factor algorithm" by Hewson et al. [2014]

Response to Reviewer 1

General Comments:

The manuscript by Hewson et al. is interesting and potentially holds useful messages for the community. I am enthusiastic about the calculation of per-pixel AMF errors, but why not push through and provide a complete error analysis for tropospheric HCHO columns? The authors also make an important point at the end of section 5.2: data users focusing on regional studies, and there will only be more of them, should aim to recalculate AMF using profile information which can resolve the spatial characteristics of their target domain. This is a good point, also with an eye on the future 7x7 km2 TROPOMI instrument. This was done by e.g. Vinken et al. [2014] and Lin et al. [2014], who improved on coarse TM4 profile by using GEOS-Chem 0.5 x 0.67 degrees to better resolve shipping lane/Chinese emissions effects in the AMF calculation for OMI NO2.

We thank the reviewer for their supportive and insightful comments. We have addressed all issues and have adjusted the manuscript accordingly where necessary. We are slowly working towards another paper that will present (1) a full error analysis for our GOME-2 product, and (2) include its validation against in-situ measurements (e.g., MAX-DOAS, aircraft). We hope the reviewer will understand our desire to present those results a subsequent manuscript.

Specific Issues

1. First of all, I find it difficult to believe that the AMF errors are "dominated by uncertainties in the HCHO profile shape". The method to compute the profile uncertainty contribution to the AMF error is not described clearly. Yes, HCHO below and above certain model levels are manipulated, but based on which hypothesis? How realistic are the perturbed profiles? I agree with reviewer#3 that a comparison with aircraft profile variability as done by Millet et al. [2006] makes much more sense.

To avoid repetition, please see our detailed response to Reviewer 3, which clarifies this issue.

2. Then on the albedo-related AMF errors; in FRESCO+ cloud retrievals, the MERIS albedo climatology is used, but for the HCHO AMF a completely different climatology is used based on TOMS (360 nm). Using wavelength-corrected (412 -> 340 nm) MERIS values would improve consistency in the retrieval approach and in the error analysis. I don't see any benefit in using the TOMS albedo or the 'improved' OMI climatology: it holds for a different time period, 1979-1993, or a different time-of-day (13:40 hrs), and both have been retrieved from a different sensor (i.e. different viewing geometries), and the TOMS dataset is spectrally not representative for 340 nm. The authors must have weighty arguments why they prefer the TOMS or OMI albedo climatology over the MERIS 412 albedo set, which could easily be spectrally scaled to 340 nm using the GOME Koelemeijer albedo climatologies. I recommend to either replace the TOMS/OMI UV albedo's with MERIS 340 nm equivalent albedo's, or the authors should convince the readers why the OMI albedo may still be useful for GOME-2 retrievals at 340 nm.

At the time of our initial submission, a GOME-2 reflectance product was not available. Hence we adopted the approach of other published leading retrievals, e.g., De Smedt et al. [2012], to use the OMI reflectance at 342 nm in the computation of GOME-2 AMFs. Whilst it was recognized that the use OMI product was not ideal in this case, it was generally accepted to being preferable to scaling the MERIS black-sky albedo (BSA), for which data at UV wavelengths does not exist. It is 'possible' to scale the MERIS BSA to 335 nm using the Koelemeijer et al. [2001] data, using a similar approach to Boersma et al. [2004]. However, this is process is subject to its own uncertainties since the wavelengths of the Koelemeijer et al. [2001] data are 335 nm and 416 nm, and MERIS BSA is at 412 nm, i.e. not an exact match anyway.

Since our submission a GOME-2 reflectance product ***has*** *become available at 340 nm, additionally with a spatially and temporally resolved error estimate. We have now integrated this dataset into our AMF algorithm to ensure consistency, i.e. GOME-2 HCHO AMFs are computed with corresponding GOME-2* *surface reflectances.* ***Given this new GOME-2 albedo is now used in the AMF computation, we do not believe it is appropriate to use scaled MERIS reflectances, as suggested by the referee.***

3. The lack of consistency between the clear-sky albedo and the albedo used for deriving the cloud fraction introduces additional errors in the HCHO AMF. The cloud fraction retrieved in FRESCO+ holds, given the surface albedo used in the FRESCO+ retrieval. Any error in the MERIS surface albedo would normally be compensated by the retrieved effective cloud fraction (if albedo is biased low, a high-biased cloud fraction still explains the TOA reflectance), but only as long as the MERIS database is used for the clear-sky AMF. Since the TOMS or OMI albedo climatologies are not consistent with the MERIS climatology, these compensating effects collapse, leaving the authors with an unknown contribution from albedo inconsistencies in their AMF values. The best would be if the authors resolve this issue by using MERIS albedo's at 340 nm for the clear-sky AMFs, but if they think that OMI-MERIS inconsistency poses no problem, they should explain why that is.

We accept that the reviewer makes a valid point here. However, as shown in Figure 8 of De Smedt et al. [2012], the impact on switching from the Koelemeijer et al. [2001] data to the MERIS albedo of Popp et al. [2011] within the FRESCO+ algorithm had only a relatively small effect on GOME-2 HCHO global seasonal means. But it is unclear whether this would also be the case if the GOME-2 surface reflectances were used in FRESCO+.

We do not have access to the FRESCO+ algorithm, and therefore cannot easily assess the impact that the new GOME-2 surface reflectances (which we have now implemented) would have on the retrieved effective cloud fractions and cloud top-pressures, and thus on our AMF computations. So some compensatory effects due to the use of the different albedo data sets may occur. We fully now acknowledge this issue within section 6 of the manuscript.

Since we now use GOME reflectances within our latest AMF algorithm (that are consistent with the observation geometry, time and AMF wavelength of the retrieval), it seems more sensible to communicate with the FRESCO+ team to implement the GOME-2 albedo data, to help overcome this problematic issue and avoid potential biases (rather than the other way around, i.e. by using the MERIS data in our AMF code). However, this is a long-term goal.

Given that the other leading GOME-2 HCHO retrieval of De Smedt et al. [2012] does not implement the MERIS albedo product, we hope the reviewer accepts our choice of surface reflectance implementation.

4. The above effects also apply on terrain height. A more sophisticated terrain height description for the HCHO AMF only makes sense if it is also applied to the cloud retrieval. From the manuscript, it is unclear if FRESCO+ accounts for terrain height in a manner consistent with what is proposed for the clear-sky AMF.

We have computed the AMF terrain height correction as done in similarly in other other published studies, but unfortunately we are not able to apply the correction in the FRESCO+ algorithm (see comment above). We ***now acknowledge this issue*** *within the new section 6 of the manuscript.*

5. The quoted uncertainty on the surface albedo is very large (0.05), and would imply that most frequently occurring albedo values over relevant areas are 100% uncertain. How did the authors arrive at this estimate for albedo uncertainty? More importantly, if they used this value, the contribution from the albedo error to the AMF error should be much larger than the _10% values over tropical forests displayed in the upper panel of Figure 8, as the sensitivity of the AMF to the local albedo is strong for low albedo values over tropical forests. I urge the authors to re-evaluate their methods, and especially the sensitivity of the AMF to the local albedo and they should explain clearly why they find so much lower albedo-related AMF errors than e.g. the 20-30% errors quoted for albedo-related AMF errors in the case of NO2 by Boersma et al. [2004], who used a much smaller albedo uncertainty of 0.02.

We use the error value given by Kelipool et al. [2008] for the albedo uncertainty [see Page1127, Line 4]. Kelipool et al. [2008] state the uncertainty of the OMI database is 0.01-0.02 for longer wavelengths but increases for shorter UV wavelengths. We therefore chose a high value of 0.05 to ensure we compute the maximum AMF uncertainty due to this source. The AMF errors associated with the surface albedo are *lower for HCHO than found for NO_2, since Rayleigh scattering is much stronger at UV wavelengths (resulting in a relative decrease of HCHO sensitivity to the surface. Again our computed errors are similar to those presented in De Smedt [2011] and De Smedt et al., [2008, 2012], i.e. of order 10%.* ***We therefore have high confidence that our estimates presented in the original manuscript are correct****.*

However, in the revised manuscript we now compute the AMFS using the GOME-2 surface reflectances and its associated error. Hence the AMF error is now calculated using per scene uncertainties rather than an assigned global value, as done previously.

6. The cloud error shown in Figure 8 is also very small over areas with a lot of HCHO; 5-10% at most. Especially for low cloud fractions, one expects a strong sensitivity of the tropospheric AMF to the cloud fraction (if this is calculated following the independent pixel approximation as stated by the authors), and hence much higher errors than quoted here. Is it possible that something is amiss with the calculation of AMF sensitivity to cloud fraction? The authors could include some typical dependency curves that illustrate the sensitivity of HCHO to albedo, cloud fraction, cloud pressure, and aerosols to convince readers that the error calculation is being done in a proper manner.

We calculate the AMF error due to cloud fraction uncertainty using the same method described in De Smedt [2011] and De Smedt et al., [2008, 2012]. As shown in Figure 9 of De Smedt et al. [2008] paper, contributions to the total AMF error from cloud fraction were estimated to be relatively small (approximately 0-10%) but contributions from errors in cloud height more important (approximately 10-40%), especially for low cloud situations. These findings agree with the error contributions presented in Figure 8 of our manuscript, since errors from cloud top-height are generally higher than those from cloud fraction. Furthermore, we have rechecked our code to reassure the reviewer that our error estimates have not been erroneously computed. Additionally, below we show below some sensitivity tests of the AMF to albedo, cloud fraction and cloud height. These plots are again ***consistent*** *with De Smedt et al. [2008; Fig 8], albeit with differences that reflect the choice and set-up of RTM and HCHO profile differences.* ***We therefore have high confidence that our error estimates are correct.***

However, to more strongly emphasize the impact of clouds we have performed two additional sensitivity tests, following the approach of Barkley et al., [2012], in which we assess errors associated with incorrect cloud fractions and cloud height, using a 'brute force' approach. In these simulations we assign an error of +0.1 in the cloud fraction, by increasing the cloud fraction used the independent pixel formula (after each observation is cloud filtered using its original value). We also test a systematic change in cloud-top pressure of -60 hPa. However, the impact of these cloud tests changes the AMFs by a small amount typically about 5%. Hence the overall effect of clouds on GOME-2 AMFs appear small. ***We have now include and discuss these results within Section 6, and acknowledge that validation using aircraft observations is necessary to clarify this error source.***

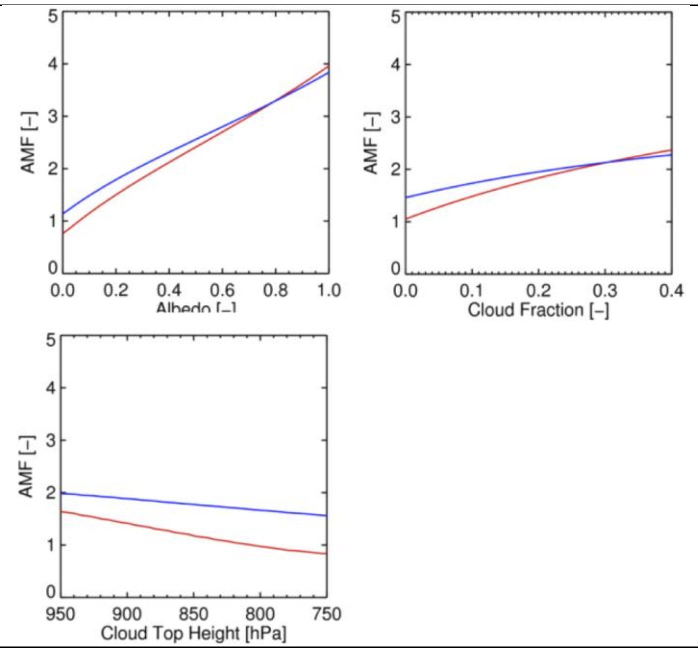

Figure 1: Sensitivity tests showing the variation in the AMF and its error, owing to changes in albedo, cloud fraction (CF), and cloud-top height (CTP).

Red line correspond to a continental GOME-2 observation over southeast US (geolocation = ~35N, 80W, albedo ~ 0.07, cloud fraction = 0.2, cloud top height ~ 990 hPa).

Blue line corresponds to a oceanic GOME-2 observation over the Atlantic Ocean US (geolocation = ~35N, 25W, albedo ~ 0.09, cloud fraction = 0.2, cloud top height ~ 935 hPa).

Default albedo, CF and CTP parameters are fixed unless being varied in each specific test. The AMF variations (left column) due albedo, CF and CTP are consistent with other studies [e.g., De Smedt 2011].

Specific minor comments:

P1113, lines 13-18: the description of w(z) is incomplete because no mention is made of how w(z) is computed for the cloudy part of the pixel. This omission should be repaired.

We have added the appropriate formula defining the scattering weights to avoid any confusion. The scattering weights are calculated in the same manner for both cloud-free and cloudy conditions using the LIDORT RTM. This has been made clear in the text.

P1114, L21: suggest to add with observation times "and viewing geometries" different from

Section text has been changed to reflect the use of the GOME-2 albedo data.

P1115, L18-20: can artificially enhance the retrieval of tropospheric columns

Changed text to: "…enhance retrieved tropospheric vertical columns…"

Section 3: I was surprised not to read about including O3 as a potential AMF dependence. Does it need to be done or not?

Lee at al. [2009] is the first paper (that we are aware of) that scaled ozone profiles in AMF calculations. In their study, which focused on SO_2, scaling the US ozone profile changed AMFs by up to 30% under certain conditions. However, as shown in section 5.7 of the manuscript, the scaling of the ozone profile has a relatively minor influence on our AMFs owing to the 340 nm wavelength used for HCHO, compared to the SO_2 wavelengths of 313 nm and 319 nm used for OMI and SCIAMACHY. We have added this information to Section 3.

P1116, L14-17: some other (SAO) retrievals do not need a background correction (K. Chance, personal communication, AGU 2014), why is it needed here? How large are the biases in the slant columns?

HCHO has a weak spectral signature, and hence most satellite HCHO retrievals have needed some form of correction typically using a reference sector technique approach. Good examples include: Palmer et al. [2006], Barkley et al, [2008, 2013], De Smedt et al., [2008, 2012] and Marais et al. [2012].

A previous version of the NASA OMI HCHO data product, produced by Kelly Chance's SAO group, did contain HCHO columns without such correction applied. However, in order to use that data in a quantitative manner individual studies had to perform their own reference sector corrections, see e.g,, Marias et al. [2012], Fortems-Cheiney et al., [2012] or Barkley et al., [2013]. In addition, the OMI HCHO column data beyond 2008 were essentially unusable due to noise. Note, the SAO group have now subsequently released an upgraded OMI HCHO retrieval that does now include a reference sector correction [Gonzalez Abad et al., 2015].

Our previous paper, Hewson et al. [2013], discusses these biases and shows the applied polynomial reference sector correction. Typically biases are of order +/- 10^{16} molecules cm^{-2}, again consistent with other HCHO retrievals.

P1118, L9: I don't see how 340 nm AMFs are consistent with 360 nm LER estimates. Why is a wavelength-dependency correction not applied to scale the LERs from 360 to 340 nm?

We accept that using 360 nm LER is not ideal to calculate the 340 nm AMFs, hence the ***upgrade*** *to the OMI 342 nm data (in the original manuscript) and now the new GOME-2 340 nm data (in the revised manuscript). Our previous studies have not scaled the TOMS data to 340 nm [e.g., Barkley et al., [2012, 2013], Hewson et al. [2013]) but rather have used the data 'as is'. Hence we do not adjust the TOMS albedo in order to be consistent with work gone before, and to facilitate comparison.*

P1119, L4-5: to what types do the SSA values quoted correspond?

Now added to manuscript.

P1123, L24-25: the BRDF effect for HCHO AMFs at 340 nm is probably even less relevant than for NO2, given the stronger Rayleigh scattering at 340 nm screening surface effects much stronger than at 440 nm.

Now acknowledged in the text.

P1127, L14-17: it should be clarified how aerosols affect the FRESCO+ effective cloud fraction retrievals.

Now acknowledged in Section 3 of the main text.

P1129, L3-9: this whole part is unclear to me. It is supposed to describe how you tested for aerosol effects on the HCHO AMF, but you lost me.

We have rephrased parts of this section to make it clear to the reader.

P1140, Table 2: I think a distinction should be made between the global retrievals listed here by Boersma et al. and Valks et al., and the regional retrievals by Lin et al. and Russell et al. Furthermore, for surface albedo, the DOMINO v2.0 uses the 440 nm values from the Kleipool et al. [2008] climatology, and not the 479.5 nm values.

Distinction now added to table and wavelength changed to 440 nm.

P1141, Table 3: the cloud approach is missing from Table 3. It should be incorporated.

Now included in Table 3.