

## **Reply to Ref. #1**

First of all we want to thank this reviewer for the positive assessment of our manuscript and the constructive and helpful suggestions!

### **General comments:**

In this paper, Wang et al. presented MAX-DOAS retrievals of NO<sub>2</sub>, SO<sub>2</sub>, and HCHO over Wuxi, a city within the heavily polluted Yangtze River Delta region in eastern China. They compared the MAX-DOAS retrievals with various OMI and GOME-2A/2B products. They also investigated the effects of a priori profiles and aerosols on satellite retrievals. The paper presents an interesting study that should be of interest to satellite trace gas retrieval community, especially the section discussing the effects of vertical profiles and aerosols on retrieval biases. The paper is generally well organized (given the multiple species/products/topics covered) and figures are mostly clear, although some improvement in writing would help. That said, I don't feel that the paper is quite ready for publication in its current form. It is very long with a lot of long, complicated sentences that are not easy to understand. I also feel that some of the 21 multi-panel figures are not completely necessary and can be removed or moved to the supplementary material. Overall, I'd recommend that the authors try to make the paper more concise and focus more on the key points.

### **Author reply:**

Many thanks for the positive assessment!

We made four important modifications to the paper. Firstly we moved section 3.1 about the coincidence criteria into supplement as section 1. Secondly we moved the section 2.1.2 about the cloud effects on MAX-DOAS observations into the supplement as section 2. Thirdly we rewrote the discussion on the aerosol effects in section 3.5 of the revised version (see general comment b from Reviewer 2). Fourthly we rewrote the discussion about the influence of the eCF on the shape factor effects on the AMF (section 3.2 of the revised version). This modification is following the specific comment #6.

### **Specific Comments:**

1) The introduction part may be a bit too long and can be shortened.

#### **Author reply:**

We shortened the introduction by rewriting paragraphs 6 and 7 of the introduction section.

2) The authors used the entire section 3.1 and several figures (Fig. 2-6) in the main text to introduce how temporal/spatial averaging is done to match MAX-DOAS data with satellite data for the comparison. To me, such a lengthy discussion would be justified if the data averaging time and/or spatial averaging radius could be used for other validation/comparison studies. But I doubt that would be the case, given the location of the site and the inhomogeneous surface properties and trace gas loading over the area (Fig. 1). I feel that this section is probably best included in the supplementary material.

#### **Author reply:**

Thanks for the suggestion! We followed your suggestion to move the entire section 3.1 into the supplement. And we added a new paragraph in the beginning of section 3 to describe the main conclusions about the coincident criteria.

- 3) How was eCF calculated? And how was daily mean satellite VCDs calculated? Did the authors consider the size of the each satellite ground footprint?

Author reply:

The effective cloud fraction (eCF) is defined as in Stammes et al. (2008). We added this reference to section 2.2 of the revised version. We directly extract eCF data from the published operational products. We clarified the calculation of the daily and bi-monthly mean satellite VCDs in the beginning of section 3 as “Here it needs to be clarified that the daily and bi-monthly averaged satellite data are the averaged values of all satellite pixels located in the coincidence area around the measurement site (see below). The averaged MAX-DOAS data are the averaged values for all measurements within 2 hours around the satellite overpass time”. Note that for the selection of the satellite data we don’t explicitly use the size of the satellite ground footprint. Instead we use the distance of the center of the pixel to the measurement station. However, we chose distances according to the size of the footprint sizes and the expected gradients of the trace gases. Note that we also exclude the outermost pixels of the OMI swath (i.e. pixel numbers 1–5 and 56–60) as described at the end of section 2.2.

- 4) Section 3.2: the NASA SO<sub>2</sub> product essentially uses the same AMF for all pixels (re-gardless of viewing geometry and other conditions) that may lead to additional errors and affect its correlation with MAX-DOAS retrievals.

Author reply:

Thanks for pointing out this important information! We agree that this simplified assumption might be part of the reason for the worse correlation of the NASA OMI product with MAX-DOAS compared to the BIRA product. We added this information in section 2.2 of the revised manuscript as “A fixed surface albedo (0.05), surface pressure (1013.25 hPa), solar zenith angle (30 °) and viewing zenith angle (0 °) as well as a fixed climatological SO<sub>2</sub> profile over the summertime eastern U.S. are assumed in the PCA retrieval (Krotkov et al., 2008).”. It needs to be noted that the significantly worse R<sup>2</sup> for the OMI NASA product compared to the OMI BIRA product could partly be attributed to the assumed fixed measurement condition (and thus the fixed AMF) in the NASA PCA retrievals. However the similar slopes and MRDs between the two OMI products indicate that the simplification of the NASA PCA retrieval only slightly contributes to the systematic bias of the averaged values.”

- 5) Section 3.2: Fig. 8, 10, 13, these figures may be replaced with a table.

Author reply:

Since figures represent the most direct way to show the dependence of the consistency on the eCF, we prefer to keep these figures.

- 6) Section 3.3: Fig. 14 – one would expect that the AMF calculated from TM4 profile would be on average smaller than AMF calculated using MAX-DOAS profile? Note that Fig. 14e and Fig. 15e show the same sign in AMF difference. The TM4 shape in Fig. 14a shows larger weight than MAX-DOAS in the lowest part of the profile, the IMAGES profile in Fig 15a, on the other hand, shows smaller weight than the MAX-DOAS profile in the in lowest layers.

Author reply:

Great thanks for pointing out this problem. The reviewer is correct with his/her description of the effect of the shape factor on the AMF. However because of the missing information on the shape factor above 4 km (which was not explicitly mentioned in the original manuscript) the SF effect

on the AMF was not correctly explained in the original manuscript. In the revised version (in section 3.2) we firstly clarified how we treat the concentration of TGs above 4km as “It needs to be noted that only the profiles below 4km can be reliably drawn from MAX-DOAS observations. Thus the profile<sub>M</sub> between 4km and the tropopause (a fixed value of 16 km is used in this study) are derived from the corresponding CTM profiles of the individual satellite data sets. Therefore the SF<sub>M</sub> is derived from the combined profile<sub>M</sub> using Eq.3.”. We also modified the explanation about the different effects on the AMF under different cloud conditions. For details, please see the revised manuscript. In addition, in part 4 of section 3.2, we also point out that the lack of information about the profiles above 4km from MAX-DOAS observations is a potential error source in the analysis of SF effects on satellite AMF calculations.

7) Fig. 18, the authors may want to point out that GOME-2/OMI ratio for NO<sub>2</sub> may be much more meaningful than that for SO<sub>2</sub>, given the overall smaller retrieval uncertainty.

Author reply:

Thanks for pointing out this issue! We modified the description in section 3.4 in the revised version as “For NO<sub>2</sub>, the Ratio<sub>Sat</sub> for both GOME-2 instruments show good agreement. Good agreement is also found for the seasonal variation with the MAX-DOAS results, but the absolute values differ. The systematic difference of Ratio<sub>Sat</sub> and Ratio<sub>M-D</sub> can be attributed to the known overestimation of the GOME-2 A/B tropospheric VCD compared to the MAX-DOAS results (see Fig. 12a). This finding also indicates that using GOME-2 and OMI data can lead to wrong conclusions about the diurnal cycles of NO<sub>2</sub>. Also for the other trace gases we investigated the ratios between the different data sets. However, because of the larger uncertainties compared to NO<sub>2</sub>, the conclusions for SO<sub>2</sub> and HCHO should be treated with care. For SO<sub>2</sub>, although Ratio<sub>Sat</sub> shows several deviations from Ratio<sub>M-D</sub>, Ratio<sub>M-D</sub> and Ratio<sub>Sat</sub> are consistent on average and close to unity during a whole year indicating similar SO<sub>2</sub> VCDs around the overpass times of GOME-2 and OMI. For HCHO, on average good agreement between Ratio<sub>Sat</sub> and Ratio<sub>M-D</sub> is found for GOME-2A and GOME-2B (except some outliers of Ratio<sub>Sat</sub>). Interestingly, both Ratio<sub>Sat</sub> and Ratio<sub>M-D</sub> are below unity indicating lower HCHO VCDs in the morning than in the afternoon.”.

8) Fig. 19: did the authors use MAX-DOAS profiles to correct for retrievals to isolate aerosols as a source of error in satellite retrievals?

Author reply:

Yes, the OMI data used in Fig. 19 (Fig. 14 in the revised version) are the modified VCD using MAX-DOAS profiles. Thanks for pointing out the missing information. We clarified it in section 3.5 as “It needs to be noted that the OMI VCDs used in Fig. 14 are the modified values using the SFs derived from MAX-DOAS observations in order to isolate the aerosol effects.”.

9) Fig. 20 and section 3.6: can the authors specify the aerosol optical properties and size distribution assumed in the RTM calculations? Particularly, for the UV wavelengths especially 319 nm?

Author reply:

We modified the description of aerosol properties used in the RTM simulations in section 3.5 of the revised version as “The aerosol optical properties (single scattering albedo of 0.9, asymmetry

parameter of 0.72, and Angstroem parameter of 0.85) are taken from the AERONET observations at the nearby Taihu station (Holben et al. 1998, 2001).” .

10) Section 4: instead of simply repeating the results already presented in the paper, the authors may consider condensing this part or provide some more in-depth discussion.

Author reply:

We made some modification in conclusion part to improve the discussion.

## Reply to Ref. #2

First of all we want to thank this reviewer for the positive assessment of our manuscript and the constructive and helpful suggestions!

### General comments

- a) Wang and co-authors investigate the quality of satellite retrievals of NO<sub>2</sub>, SO<sub>2</sub>, and HCHO over Wuxi in polluted China via a detailed comparison with ground-based col-umn measurements obtained with the MAX-DOAS technique. This technique is sen-sitive to pollution in the lower atmosphere, and Wuxi in the Yangtze River area faces pervasive high levels of pollution from these gases and aerosols. The three years of MAX-DOAS measurements collected in Wuxi thus provide a very interesting data set to test the satellite retrievals, and provide guidance on how to use and possibly improve the retrievals. The authors report that the KNMI OMI NO<sub>2</sub> (DOMINO v2) product agrees very well with the MAX-DOAS NO<sub>2</sub> columns in Wuxi, especially in situations with few clouds. But the KNMI NO<sub>2</sub> products from the GOME-2 sensors tend to be overestimated. Because of this overestimation of GOME-2 NO<sub>2</sub>, also the satellite-derived NO<sub>2</sub> diurnal cycle, while correct in sign, is overestimated. Satellite retrievals of SO<sub>2</sub> and HCHO from BIRA and NASA tend to be underestimated by tens of per-cents relative to the MAX-DOAS measurements. These findings are relevant to the many users of satellite data interested in obtaining a better understanding of Chinese air pollution.

The paper then addresses some of the critical assumptions made in the satellite re-trievals on: the a priori trace gas vertical distribution in the retrievals, the cloud correc-tions made, the aerosol correction, and to what extent this proceeds implicitly via the cloud retrievals that are sensitive to aerosol effects (Leitao et al., 2011; Boersma et al., 2011; Castellanos et al., 2015; Chimot et al., 2016). The comparison of MAX-DOAS and (CTM-derived) a priori profile shapes is a strong and innovative element of the study, and it is interesting to see how replacing the CTM-profiles with the actually observed profiles helps in improving the agreement between MAX-DOAS and satellite retrievals. Profile validation is dearly needed, and this study explores new avenues on how to achieve this, even though the vertical resolution of the MAX-DOAS and model profiles differ substantially. One highlight is that~20% of the SO<sub>2</sub> and HCHO underestimation can be explained by the IMAGES profile shapes insufficiently capturing the enhanced SO<sub>2</sub> and HCHO concentrations in the Wuxi boundary layer.

### Author reply:

Many thanks for the positive assessment!

We made four important modifications for the paper. Firstly we moved section 3.1 about the coincidence criteria into supplement as section 1. Secondly we moved the section 2.1.2 about the cloud effects on MAX-DOAS observations into the supplement as section 2. Thirdly we rewrote the discussion on the aerosol effects in section 3.5 of the revised version (see general comment b). Fourthly we rewrote the discussion about the influence of the eCF on the shape factor effects on the AMF (section 3.2 of the revised version). This modification is following the specific comment #6 from Reviewer 1.

- b) Section 3.6 on aerosol effects on the AMFs is potentially also interesting, but I have serious concerns about the way it has been set up, and the current method does not allow drawing any firm conclusions. The section starts with an analysis of the NO<sub>2</sub> discrepancies (satellite minus MAX-DOAS) as a function of AOD. This is relevant, but it does not become clear whether the discrepancies arise because of high AOD, because of residual clouds, or because of aerosols influencing the cloud fractions. Showing NO<sub>2</sub> discrepancies only for cloud fraction < 0.1 is inconclusive since these cloud

fractions. may represent real clouds, ‘effective’ clouds, or a combination of the two. To properly attribute the NO<sub>2</sub> discrepancies to the effect of the aerosols, the authors should do what they did for Table 2: use MODIS to distinguish the cloud-free, aerosol loaded situations from the situations with residual clouds still present, and focus their analysis on that data cloud-cleared ensemble to rule out the contributions from clouds.

The subsequent box AMF calculations are only just a brief sensitivity study for a limited set of situations that is not representative for the large and robust data ensemble collected by the authors over Wuxi. For instance, only one viewing geometry has been tested (P18, L1). Furthermore, how much box AMFs differ between implicit and explicit aerosol corrections depends strongly on the exact assumption of AOT (profile), particle type, NO<sub>2</sub> profile, albedo (why always 0.1?), as shown in many previous studies (e.g. Leitao et al. [2011]). None of this becomes clear on page 18, yet the conclusion is drawn that “the implicit aerosol correction typically causes larger bias of the satellite TG VCDs than the clear-sky assumption”. This conclusion is based on only a few calculations that do not represent the full range of situations encountered by the retrievals under evaluation. The authors should have been as rigorous as in section 3.3 and replace the implicit aerosol correction by an explicit aerosol correction for the full set of satellite pixels.

Author reply:

Based on the comments of the reviewer, we rewrote the whole section 3.5 in the revised version about aerosol effects. One important point is that although the differences of clear sky AMF, implicit aerosol correction, and explicit aerosol correction have been systematically investigated in the previous studies (i.e. Leitão et al. (2010) and Chimot et al., 2016), here we characterize the aerosol effects for typical aerosol properties (profile, optical properties, and corresponding aerosol induced eCF and CTP) for a polluted region. Also, as mentioned by the reviewer the previous studies indicated that the aerosol effect “depends strongly on the exact assumption of AOT (profile), particle type, NO<sub>2</sub> profile, and albedo”.

Thus we completely re-wrote the whole section, and we extended the RTM simulations to five different satellite observation geometries (listed in Table 2 of the manuscript) following the suggestions of the reviewer. The new results are shown in Fig. 17 of the revised manuscript. The new simulations indicate that the aerosol effects depends on the observation geometries, however the main conclusion on the effects of clear sky AMF and implicit aerosol corrections are consistent for different geometries.

For the discussion on Fig. 14 in the revised version (Fig. 19 in the original version) about the analysis of the NO<sub>2</sub> discrepancies as a function of AOD, we agree with the reviewer that for cloud fractions < 0.1 residual clouds can not certainly be excluded. Therefore we also used an additional criterium of CTP>900hPa, which can exclude residual cirrus clouds. Considering specific low altitude clouds (with either small OD and large geometric coverage or high OD and small geometric coverage) we performed additional simulation studies, which are described in the section 4 of the supplement. Our main conclusion is that for the selected cases the effect of residual clouds is negligible. We added this information to the main text of the manuscript.

In addition we excluded the part about the six pure aerosol pollution days, because we can not draw any general conclusion from these cases. We added a new figure (Fig. 15) showing aerosol-induced eCF and CTP derived from the OMI cloud retrieval as a function of the corresponding AOD derived from MAX-DOAS.

The reviewer asked the question “why always 0.1?”. Here we updated the text as follows “The surface albedo is set to 0.1 for NO<sub>2</sub> and 0.05 for SO<sub>2</sub> and HCHO simulations based on the averaged value of the surface reflectivity data base derived from OMI by Kleipool et al. (2008) over Wuxi station.” And we redo the RTM simulations with these surface albedo for more observation geometries using McArtim RTM.

c) The paper is too long. The section on the coincidence criteria can be shortened considerably. Other studies have investigated these issues, and the findings are probably specific for the Wuxi circumstances anyway. I recommend to move much of section 3.1, including the figures, to the supplementary material and focus on the final criteria, and then refer the reader for justification of

these criteria to the supplement. Also sections 3.3 can be shortened; I'm not sure if for each retrieval the discrepancies as a function of cloud fraction need to be discussed (and shown) at length.

Author reply:

Many thanks for the suggestion! We followed your suggestion to move the entire section 3.1 into the supplement. And we added a new paragraph at the beginning of section 3 to describe the main conclusions about the coincidence criteria.

- d) The systematic dependence of the HCHO spectral fitting uncertainty on the retrieved VCD for GOME-2 is intriguing, and deserves more attention. Why is this exactly? Why would this be different than for OMI? The authors should clarify these issues. Then their decision to only validate OMI HCHO retrievals with fitting uncertainties  $< 7 \cdot 10^{15}$  molec.cm<sup>-2</sup> is questionable, since setting this threshold basically excludes half the data, not just some outliers or misfits. The authors may report that validation results for this sub-set are better than for the full set, as long as those results are also reported, because users of OMI HCHO data typically use all data, not just the sub-set retrieved with SCD uncertainties  $< 7 \cdot 10^{15}$  molec.cm<sup>-2</sup>.

Author reply:

Unfortunately, at the moment we can't give any confirmed explanation on dependence of the HCHO spectral fitting uncertainty on the retrieved VCD for GOME-2 and the differences compared to OMI. We clarified this in the revised manuscript.

Concerning the filter of the fit error, the Fig. 6b in the revised version (Fig. 11 in the previous version) shows the comparisons of the linear regression parameters for the data before and after the filtering. We also add a new Fig. S12 in the supplement to show the effect of the fit error on the daily averaged data. The two comparisons demonstrate that the filter only considerably improves the correlation coefficient, but hardly changes the slopes and y-intercepts. Thus we conclude that it will not impact the conclusion on the systematic bias of the OMI HCHO products. The point is clarified in the revised manuscript. Furthermore, as mentioned in the paper, data with large uncertainty need to be excluded for a further investigation on cloud and aerosol effects. Otherwise the effects will be overwhelmed by the large uncertainties.

### Specific comments

- 1) P3, L17-20: here it should be stressed that methodological assumptions on how clouds and aerosols should be accounted for in the AMF calculation matter, e.g. Lin et al. [2015].

Author reply:

We add this finding and the reference to the text.

- 2) P4, L1: studies investigating the shape factor are not "rare". There are many studies investigating the quality and effect of a priori profiles on retrievals and emission estimates; e.g. Boersma et al. [2004]; Hains et al., [2010]; Heckel et al. [2011]; Barkley et al. [2012]; Vinken et al. [2014]. Regardless, studying the impact of the shape factor remains relevant because profile measurements are indeed 'rare'.

Author reply:

We corrected the sentence as "Here it is important to note that many studies already investigated the quality and effect of a-priori SFs on satellite retrievals (i.e. Boersma et al., 2004; Hains et al., 2010; Heckel et al., 2011) and demonstrated that the SF effect on the tropospheric AMFs can dominate the systematic errors of tropospheric satellite products especially in highly polluted (especially urban and industrial) regions (Boersma et al., 2011, Theys et al., 2015 and De Smedt et al., 2015), Nevertheless, because profile measurements are rare, the SF effect is still not well understood in many regions."

- 3) P4, L35 and P5, L1-3: the argument in favour of the implicit aerosol correction in the Boersma-2011 paper is made for substantial AOD when particles are mostly scattering, i.e. not unlike cloud droplets. Castellanos et al. [2015] clearly showed that for absorbing particles and high AOD, the implicit aerosol correction breaks down. So the sentence that Castellanos demonstrated that for elevated biomass burning aerosols, the implicit correction does a good job is completely out of place. Their study showed that the implicit aerosol correction compares well with an explicit aerosol correction for low-modest AOD and  $SSA > 0.95$ . For high AOD and lower SSA, the implicit aerosol correction breaks down, but these situations occur less frequently than the former.

Author reply:

Many thanks for this hint! We modified the sentence as “For mostly scattering aerosols at high altitudes the implicit aerosol correction can largely account for the aerosol effect on the TG products (Boersma et al., 2011). However in some important cases (for low altitude aerosols with high AOD and small SSA) the implicit correction might even increase the errors of the AMF Castellanos et al. (2015).”.

- 4) P5, L31: it should be ‘heavy fog’.

Author reply:

corrected

- 5) P6, L4-6: it should be clarified if the difference between the geometrical approximation and profile integration is systematic, or that the discrepancies are variable in both directions.

Author reply:

We clarified it as “Our previous study (Wang et al., 2016) demonstrated that the tropospheric trace gas VCDs from the full profile inversion are in general much more accurate than those from the geometric approximation. The discrepancy of VCDs between the two methods is systematic and can be mainly attributed to the errors of the geometric approximation, for which the errors can be up to 30% depending on the observation geometry, and the properties of aerosols and TGs. ”.

- 6) P6, L12: Capital S missing in ‘sky’.

Author reply:

Corrected.

- 7) P6, L30: what is the source of information for the  $68 \times 14 \text{ km}^2$  pixel size at OMI swath edges?

Author reply:

Many thanks for this hint! We changed the values to  $150 \times 13 \text{ km}^2$ , see:

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote Sens.*, 44, 1093–1101, 2006b.

- 8) P7, L5-7: it would be appropriate to refer to Dirksen et al. [2011] here when discussing the data assimilation procedure to estimate the stratospheric background  $\text{NO}_2$ . Similar to OMI  $\text{SO}_2$  from BIRA, DOMINO v2 can be regarded as the ‘proxy’ algorithm for the upcoming TROPOMI mission.

Author reply:

We added the reference to Dirksen et al. [2011] and also clarified “The retrieval algorithm for DOMINO v2 forms the basis of  $\text{NO}_2$  retrievals for the upcoming TROPospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor mission (Veefkind et al., 2012).”

9) P8, L13: suggest to state ‘similar data assimilation procedures’.

Author reply:  
Corrected.

10) P9, L26: what is meant with the ‘statistical uncertainty of the satellite data’?

Author reply:  
We delete “statistical”.

11) P11, LL27-28

Author reply:

12) P12, L11-14: with underestimations of 50%, it is rather odd to conclude that GOME- 2A products are “most accurate” for cloud fractions below 30%. Also the ‘recommendation’ to use SO<sub>2</sub> observations with cloud fractions below 10% is far fetched. One might as well recommend to not use any SO<sub>2</sub> data over the Yangtze area at all in view of the large, systematic biases shown in this study.

Author reply:  
We modified the description as “Thus we conclude that the cloud effects on both GOME-2A products are appreciable for eCF > 30%. For the GOME-2B BIRA data, an obvious decrease of R<sup>2</sup> and slope is found for eCF > 10%, while for eCF>30% largely variable MRDs are found. Thus clouds can considerably impact the GOME-2B BIRA product for eCF > 10%, and more significantly for eCF > 30%. ”.

13) P12, L29: ‘because of the weaker degradation’ than OMI or GOME-2A? Please clarify.

Author reply:  
We changed the text to “because of the weaker degradation of GOME-2B during the short time after launch compared to OMI and GOME-2A.”.

14) P13, L1: dependencies.

Author reply:  
Corrected.

15) P13, L4: when suggesting that HCHO products should be used for cloud fractions < 0.3, the authors should be more aware that their recommendation is based on the situation for Wuxi, which is not necessarily representative for situations with enhanced HCHO concentrations elsewhere (just think about the high aerosol loadings). Also, if they make such a recommendation, they should discuss it in the context of what the algorithm providers actually recommend for appropriate use of their data, and what has typically been done in successful applications of the OMI HCHO data.

Author reply:  
We modified the sentence as “In general cloud effects on the HCHO products become substantial for eCF > 30% for the three satellite instruments. However it needs to be noted that our findings are derived for one location (Wuxi) and might not be fully representative for other locations. The use of the HCHO products with eCF < 40% is recommended by the retrieval algorithm developer (De Smedt et al., 2015).”.

16) P14, L11: ‘latitude range’ should be altitude range, and ‘larges’ should be ‘largest’.

Author reply:  
Corrected.

17) P14, L11-14: it would be fair to clearly conclude here that the TM4 a priori profile shapes agree well with the MAX-DOAS NO<sub>2</sub> profiles in an average sense.

Author reply:  
We added this finding.

18) P15, L15: please provide more detail on the months in the x-axis of Figure 17; we now only have tick marks for month 5 and 11. Some more specific indication for the bi-monthly averages would be useful.

Author reply:  
We modified the figure accordingly. Note that the Fig. 17 in the previous version is Fig. 12 in the revised version.

19) P15, L23-24: please clarify why the TM4 NO<sub>2</sub> columns are so much lower than those from the measurements. Later on page 16, same for SO<sub>2</sub> modelled by IMAGES; why is HCHO from IMAGES doing a good job whereas SO<sub>2</sub> is not?

Author reply:  
We added that “The significant underestimation of the TM4 NO<sub>2</sub> VCDs could be due to many factors, most importantly the limited spatial model resolution, which is especially relevant for species with strong horizontal gradients such as NO<sub>2</sub> and SO<sub>2</sub> (see Figure 1), but also possible errors in the emissions, transport and/or chemical mechanism. The determination of the specific contributions of the different error sources should be the subject of future studies.” in the revised version. We also mention that the results of the IMAGES model for SO<sub>2</sub> and HCHO need further investigations in the future.

20) P16, L22-24: it would be appropriate to refer to Boersma et al., JGR, 2008 here. That study was the first to investigate the diurnal cycle of NO<sub>2</sub> with satellite measurements. Also some more explanation on what causes the diurnal changes in NO<sub>2</sub>, SO<sub>2</sub>, and HCHO columns is needed here.

Author reply:  
We added the reference and now mention that “The diurnal variations can be attributed to the complex interaction of the primary and secondary emission sources, depositions, atmospheric chemical reactions, and transport processes.”.

21) P17, L18: some more information is needed on the ‘clear-sky AMF’ that is applied in SO<sub>2</sub> and HCHO retrievals for cloud fractions < 0.1. How is such an AMF calculated – in an atmosphere with Rayleigh scattering only? Or is there some aerosol background assumed in the radiative transfer calculations?

Author reply:  
‘clear-sky AMF’ means in an atmosphere with Rayleigh scattering only. We clarified this in the revised version.

22) P19, L24: please clarify what is meant with “cloud effects become significant”. Do you mean that the discrepancies between MAX-DOAS and satellite columns are larger when cloud fractions are larger?

Author reply:  
Yes. We already clarified it in the sentence before that sentence as “The consistency (correlations and systematic bias) of satellite data with MAX-DOAS results deteriorates with increasing eCF.”.

23) P19, L33-34: suggest to be more specific here and state that IMAGES profiles and TM4 profiles have been compared against MAX-DOAS profiles.

Author reply:

We added this information in the revised version.

24) P20, L21-22: the sentence “NO<sub>2</sub> satellite products systematically overestimate the magnitude of NO<sub>2</sub> diurnal variation” is misleading. The diurnal variation is overestimated because the GOME-2 retrievals are too high, but OMI is in agreement with MAX-DOAS. Suggest to rephrase accordingly.

Author reply:

We changed the text to “The systematic difference of RatioSat and RatioM-D can be attributed to the known overestimation of the GOME-2 A/B tropospheric VCD compared to the MAX-DOAS results (see Fig. 12a). This finding also indicates that using GOME-2 and OMI data can lead to wrong conclusions about the diurnal cycles of NO<sub>2</sub>.”.

25) P20, L30-35: this part is too strong-worded and should be rephrased after the authors have addressed my concerns about section 3.6. The current sensitivity study provides too little ground to base these conclusions on.

Author reply:

We modified the section 3.5 in the revised version (section 3.6 in the previous version). Thus the relevant conclusion part is re-written as:

Finally we studied aerosol effects on the OMI products over Wuxi station based on the MAX-DOAS observations. We find that the underestimation of the TG VCDs derived from satellite observations for mainly cloud-free observations compared to the MAX-DOAS observations systematically increases with AOD. We also investigate the aerosol effect based on RTM simulations. Here it is also possible to separate the aerosol effect into two contributions: a) the effect of using a clear sky AMF instead of an AMF taking explicitly into account the aerosol effects, and b) the effect of aerosols on the cloud retrievals, which are used in the satellite TG retrievals (implicit aerosol correction). We find that for the measurements affected by high aerosol loads in Wuxi, in general the effect of the implicit cloud correction on the retrieved TG VCDs is much stronger than the difference of a clear sky AMF compared to an AMF taking explicitly into account the aerosol extinction. We also showed that for eCF <10% and CTP >900hPa the effect of residual clouds can be neglected if aerosol extinction is explicitly taken into account. Moreover, the observed underestimation of the OMI NO<sub>2</sub> VCD for large AOD can be well explained by the error caused by the implicit aerosol correction. Therefore it could be reasonable to apply the clear-sky AMFs in the satellite retrievals of TG tropospheric VCDs in case of CTP > 900hPa and eCF<10% if explicit aerosol information is not available.

# Validation of OMI, GOME-2A<sub>2</sub> and GOME-2B tropospheric NO<sub>2</sub>, SO<sub>2</sub> and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of a-priori profiles effect and aerosols effect on the satellite products

5 Yang Wang<sup>1</sup>, Steffen Beirle<sup>1</sup>, Johannes Lampel<sup>1,2</sup>, Mariliza Koukoulis<sup>3</sup>, Isabelle De Smedt<sup>4</sup>, Nicolas Theys<sup>4</sup>, Ang Li<sup>5</sup>, Dexia Wu<sup>5</sup>, Pinhua Xie<sup>5,6,7</sup>, Cheng Liu<sup>8,6,5</sup>, Michel Van Roozendael<sup>4</sup>, Trissevgeni Stavrakou<sup>4</sup>, Jean-François Müller<sup>4</sup>, and Thomas Wagner<sup>1</sup>

<sup>1</sup> Max Planck Institute for Chemistry, Mainz, Germany

<sup>2</sup> Institute of Environmental Physics, University of Heidelberg, Heidelberg, Germany

10 <sup>3</sup> Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>4</sup> Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

<sup>5</sup> Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China

<sup>6</sup> CAS Center for Excellence in Urban Atmospheric Environment, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China

15 <sup>7</sup> School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei, China

<sup>8</sup> School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China

*Correspondence to:* Yang Wang (y.wang@mpic.de), Ang Li (Angli@aiofm.ac.cn), Cheng Liu (chliu81@ustc.edu.cn)

## 20 **Abstract.**

Tropospheric vertical column densities (VCDs) of NO<sub>2</sub>, SO<sub>2</sub> and HCHO derived from Ozone Monitoring Instrument (OMI) on AURA and Global Ozone Monitoring Experiment 2 aboard METOP-A (GOME-2A) and METOP-B (GOME-2B) are widely used to characterize the global distributions, trends, dominating sources of the trace gases and for comparisons with chemical transport models (CTM). We use tropospheric VCDs and vertical profiles of NO<sub>2</sub>, SO<sub>2</sub> and HCHO derived from  
25 MAX-DOAS measurements from 2011 to 2014 in Wuxi, China, to validate the corresponding products derived from OMI, and GOME-2A/B by different scientific teams (daily and bimonthly averaged data). Prior to the comparison, we investigate the effects of the the spatial and temporal coincidence criteria for MAX-DOAS and satellite data on the comparison results. are determined by a use of sensitivity studies using different spatial and temporal averaging conditions. We find that the distance of satellite data from the location of the MAX-DOAS station is the dominating effect, and we make suggestions for the spatial (20km for OMI NO<sub>2</sub> and SO<sub>2</sub> products and 50km for OMI HCHO and all GOME-2A/B products) and temporal averaging (2 hours around satellite overpass time). We also investigate the effect of eCloud effects s on both MAX-DOAS and satellite observations are also investigated. Our results indicate that the discrepancies between satellite and MAX-DOAS results increase with increasing effective cloud fractions and are dominated by the cloud effect on the satellite products. ~~Our comparison results~~ In comparisons with MAX-DOAS, we found ~~indicate~~ a systematic underestimation of all SO<sub>2</sub> (40% to  
30 57%) and HCHO products (about 20%), and an overestimation of the GOME-2A/B NO<sub>2</sub> products (about 30%) ~~but well~~

consistent DOMINO version 2 NO<sub>2</sub> product (~~it is only slightly underestimated by 1%~~). To better understand the reasons for the differences, we evaluated the a-priori profile shapes ~~derived from CTM and~~ used in OMI retrievals (derived from CTM) by comparisons with those derived from MAX-DOAS observations. ~~A significant differences are~~ found for SO<sub>2</sub> and HCHO profile shapes derived from the IMAGES model, whereas on average good agreement was found ~~well consistency is found~~ for the NO<sub>2</sub> profile shapes derived from the TM4 model ~~in an average sense. we recalculated the AMFs for satellite observations based on the shape factors (SFs) derived from MAX-DOAS. The recalculated satellite VCDs agree better with the MAX-~~ he modified satellite VCDs ~~by a use of~~ based on the profile shapes ~~derived~~ from MAX-DOAS observations agree better with the MAX-DOAS VCDs than ~~those from the~~ VCDs from the original ~~products data sets~~ by up to 10%, 47% and 35% for NO<sub>2</sub>, SO<sub>2</sub> and HCHO, respectively. ~~The improvement is strongest for periods with large trace gas VCDs. Finally~~ Furthermore, we investigate the effect of aerosols on the satellite retrievals. For OMI observations of NO<sub>2</sub>, a systematic underestimation is found for a large AOD, which is mainly attributed to effect of the aerosols on the cloud retrieval and subsequent application of a cloud correction scheme (implicit aerosol correction). In contrast, the effect of aerosols on the clear sky AMF (explicit aerosol correction) ~~(implicit aerosol correction)~~ has a smaller effect. For SO<sub>2</sub> and HCHO observations selected in the same way, no clear aerosol effect is found, probably because for the considered data no cloud correction is applied (but also because of the larger scatter). From our findings we conclude that for ~~for the common aerosol and trace gas scenarios in Wuxi. For the scenarios, the implicit correction (with corresponding effective cloud fraction (eCF) below 10% and cloud top pressure (CTP) of 1000 to 850 hPa) could cause a larger negative bias of tropospheric VCDs than the clear sky assumption compared to explicit aerosol corrections. And the bias can amount to up to about 45%, 15% and 35% for NO<sub>2</sub>, HCHO, and SO<sub>2</sub>, respectively~~ We find an increasing underestimation of the OMI NO<sub>2</sub>, SO<sub>2</sub> and HCHO products with increasing AOD by up to 8%, 12% and 2%, respectively. One reason for this finding is that aerosols systematically affect the satellite cloud retrievals and can lead to apparent effective cloud fractions of up to 10% and apparent cloud top pressures of down to 830 hPa for the typical urban region in Wuxi. We show that in such cases the implicit aerosol correction could cause a strong underestimation of tropospheric VCDs by up to about 45%, 77% and 100% for NO<sub>2</sub>, SO<sub>2</sub> and HCHO, respectively. ~~Therefore it could be reasonable to apply the clear sky AMFs in the satellite retrievals of TG tropospheric VCDs observations in case for observations of~~ with CTP > 900hPa and eCF < 10% the application of a clear sky AMF might be a good option if explicit aerosol information is not available. ~~In addition~~ Another finding of our study is that ~~in comparisons with MAX-DOAS, the diurnal variations ( the ratio of morning to afternoon ratios) of NO<sub>2</sub> VCDs can be considerably~~ For such conditions it might be better to apply AMFs for clear sky conditions than AMFs based on the satellite cloud retrievals.

~~We find that the satellites systematically overestimated~~ if results from different sensors and/or retrievals ~~by combining the~~ (e.g. OMI and GOME-2) are used, ~~observations due to the overestimation of GOME-2 NO<sub>2</sub> VCDs~~ the magnitude of the diurnal variations of NO<sub>2</sub> and HCHO, whereas ~~well~~ less deviations ~~consistency with MAX-DOAS is found~~ for HCHO and SO<sub>2</sub> VCDs are found ~~satellite observations.~~ ~~No significant weekly cycle for all trace gases is found by either the satellites or the MAX-DOAS measurements.~~

## 1 Introduction

Nitrogen oxides ( $\text{NO}_x \equiv \text{NO}_2 + \text{NO}$ ), sulphur dioxide ( $\text{SO}_2$ ), and formaldehyde (HCHO) play critical roles in the tropospheric chemistry through various gas phase and multi-phase chemical reactions (Seinfeld and Pandis, 1998). In an urban and industrialized region, anthropogenic emissions from traffic, domestic heating, factories, power plants and biomass burning significantly elevate the concentrations of these (and other) trace gases (TGs) in the boundary layer (Environmental Protection Agency, 1998; Seinfeld and Pandis, 1998). There is strong evidence that aerosol particles formed through photochemistry of  ~~$\text{NO}_x$  nitrogen oxides~~,  $\text{SO}_2$ , and VOCs significantly contribute to haze pollution events occurring frequently around megacities and urban agglomerations in China, like the Jing-Jin-Ji region and the Yangtze River Delta region (Crippa et al., 2014; Huang et al., 2014; Jiang et al., 2015; Fu et al., 2014). The aerosols also impact the local radiative forcing through direct (e.g. McCormic and Ludwig, 1967) and indirect effects (Lohmann and Feichter, 2005). Understanding global and regional distributions and temporal variations of the TGs, and further identifying and quantifying their dominant sources can provide a firm basis for a better understanding of the formation mechanisms of haze pollution and for the development of mitigation strategies.

Since 1995 a series of sun-synchronous satellites, such as ERS-2, ENVISAT, AURA, METOP-A and METOP-B, ~~were~~ was launched carrying UV/vis/NIR spectrometers with moderate spectral resolution, which allowed scientists to determine ~~the~~ global distributions of several important tropospheric trace gases including  $\text{NO}_2$ , HCHO and  $\text{SO}_2$  for the first time. The first instrument was the Global Ozone Monitoring Experiment (GOME) (Burrows et al., 1999), followed by the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) (e.g. Bovensmann et al., 1999), the Ozone Monitoring Instrument (OMI) (Levelt et al., 2006a, b), and the GOME-2A and GOME-2B instruments (Callies et al., 2000; Munro et al., 2006, 2016). The OMI and GOME-2A/B instruments are still in operation. A large number of studies developed retrieval algorithms to acquire the tropospheric vertical column densities (VCD) of  $\text{NO}_2$  (e.g. Boersma et al., 2004, 2007 and 2011; Richter et al., 2005; Beirle et al., 2010 and Valks et al., 2011),  $\text{SO}_2$  (e.g. Krueger et al., 1995; Eisinger and Burrows, 1998; Carn et al., 2004; Krotkov et al., 2006; Richter et al., 2006 and 2009; Yang et al., 2007; Lee et al., 2009; Nowlan et al., 2011; Rix et al., 2012; Li et al., 2013; Theys et al., 2015), and HCHO (Chance et al., 2000; Palmer et al., 2001; Wittrock et al., 2006a; De Smedt et al., 2008, 2012 and 2015; Kurosu, 2008; Millet et al., 2008; Hewson et al., 2013; González Abad et al., 2015) for all the satellite instruments. In this validation study we include several products, which are published recently and widely used: for  $\text{NO}_2$  the near-real-time OMI DOMINO v2.0 (Boersma et al., 2007 and 2011) and the GOME-2A/B TM4NO2A (Boersma et al., 2004); for  $\text{SO}_2$  the operational OMSO2 OMI product (Li et al., 2013) published by National Aeronautics and Space Administration (NASA), the O3M-SAF operational GOME-2A product published by the German Aerospace Centre (DLR) (Rix et al., 2012 and Hassinen et al., 2016), and the OMI, ~~and~~ GOME-2A/B products developed by BIRA (Theys et al., 2015); for HCHO the OMI and GOME-2A/B products developed by BIRA (De Smedt et al., 2008, 2012 and 2015). Many users already benefit from these products for several ~~atmospheric~~ applications, e.g. detection and quantification of emissions, identification of transport processes and chemical transformations, and for the

comparison with model simulations (e.g. Beirle et al., 2003 and 2011; Martin et al., 2003; Richter et al., 2005; van der A et al., 2008; Herron-Thorpe et al., 2010; Gonzi et al., 2011 ~~and~~; Barkley et al., 2012; Koukouli et al., 2016).

Although several studies have paid efforts to improve ~~the~~ satellite retrievals, still significant differences compared to ground based measurements were reported by several validation studies, e.g. a systematic underestimation of the tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO was obtained for OMI by > 30% in or near Beijing, China (Ma et al., 2013; Theys, et al., 2015 and De Smedt et al., 2015; Jin, et al., 2016). The satellite retrieval errors are mainly attributed to the slant column retrievals, the stratospheric correction (for NO<sub>2</sub>) and the tropospheric air mass factor (AMF) calculations. The AMF uncertainties are related to several factors, such as the surface albedo, the cloud and aerosol properties, methodological assumptions on how clouds and aerosols should be accounted (Lin et al., 2015), the a-priori ~~(relative)~~-profile shape (also referred to as the shape factor (SF) in the following) as well as interpolation errors of the discrete look-up table entries (Lin et al., 2014). Thus validation studies for satellite products using independent ground-based measurements are essential to quantify uncertainties, identify dominant error sources and to further improve the satellite retrieval algorithms.

Since about 15 years ago, the Multi Axis - Differential Optical Absorption Spectroscopy (MAX-DOAS) technique (Hönninger and Platt, 2002; Bobrowski et al., 2003; Van Roozendaal et al., 2003; Hönninger et al., 2004; Wagner et al., 2004 and Wittrock et al., 2004), is applied to retrieve tropospheric vertical profiles of TGs and aerosols from spectra of scattered UV/Visible sunlight measured at different elevation angles (e.g. Frieß et al., 2006, 2011 and 2016; Wittrock et al., 2006b; Irie et al., 2008 and 2011; Clemer et al., 2010; Li et al., 2010 and 2012; Vlemmix et al., 2010, 2011 and 2015b; Wagner et al., 2011; Yilmaz, 2012; Hartl and Wenig, 2013 and Wang et al., 2013a and b). MAX-DOAS observations provide valuable information that can be applied for ~~the a~~ quantification of air pollutants (e.g. Li et al., 2012; Hendrick et al., 2014; Wang et al., 2014a; Wang et al., 2016) ~~and for a the~~ validation of tropospheric satellite products (e.g. Irie et al., 2012 and 2016; Ma et al., 2013; Kanaya et al., 2014; Theys, et al., 2015 and De Smedt et al., 2015; Jin et al., 2016), and an evaluation of ~~results of~~ chemical transport model (CTM) simulations (e.g. Vlemmix et al., 2015a). The tropospheric vertical profiles are also valuable for the evaluation of SFs used in the satellite AMF calculations. Here it is important to note that ~~errors of the tropospheric AMFs usually dominate the systematic errors of tropospheric satellite products especially in highly polluted (especially urban and industrial) regions (Boersma et al., 2011; Theys et al., 2015 and De Smedt et al., 2015)~~, many studies already investigated the quality and effect of a-priori SFs on satellite retrievals (i.e. Boersma et al., 2004; Hains et al., 2010; Heckel et al., 2011) and demonstrated that the ~~errors~~ SF effect ~~of~~ on the tropospheric AMFs ~~usually could~~ can dominate the systematic errors of tropospheric satellite products especially in highly polluted (especially urban and industrial) regions (Boersma et al., 2011; Theys et al., 2015 and De Smedt et al., 2015), Nevertheless, because profile measurements are rare, ~~the SF effect is still not~~ but studies on the effect of the SF on the satellite retrievals are still rare well understood in many regions. In this study the SF effect ~~of the SF~~ on the tropospheric AMF will be investigated using the vertical profiles of the TGs derived from the MAX-DOAS observations in Wuxi, China from 2011 to 2014 (Wang et al., 2016).

Wuxi is located about 130 km north-west of Shanghai belonging to the most industrialized part of the Yangtze River delta (YRD) region. YRD including Shanghai City and four nearby provinces is the largest economic region in China and

heavily industrialized and can be considered as the largest metropolitan area in Asia with the population of about 150 millions. ~~The air pollution due to strong anthropogenic pollutant emissions in this region threatens the health of the inhabitants and has been of great concern in the atmospheric and environmental science community as well as for the public.~~

Several studies already used satellite products of the pollutants to quantify the corresponding emissions (Ding et al., 2015; Han et al., 2015; Bauwens et al., 2016) in this region. However validation studies for the satellite products in this region are still sparse. Chen et al. (2009), Irie et al. (2012), Kanaya et al. (2014), and Chan et al. (2015) validated the satellite NO<sub>2</sub> tropospheric VCD products using MAX-DOAS (or zenith-sky DOAS) measurements in Rudong, Hefei, and Shanghai. So far there are no validation reports for SO<sub>2</sub> and HCHO products in the YRD region. However several validation studies have been carried out in other regions of China (e.g. Theys et al., 2015; De Smedt et al., 2015; and Jin et al., 2016).

In this study we validate daily (2 hours around the satellite overpass time) and bi-monthly averaged tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub>, and HCHO derived from OMI and GOME-2 using the MAX-DOAS observations in Wuxi. ~~To minimise the influence of different air masses detected by MAX-DOAS and satellite instruments, coincidence-Coincidence criteria should be used for both data sets. In this study we investigate the influence of the temporal and spatial coincidence criteria of satellite and MAX-DOAS observations are discussed. So far only few~~ Previous studies (Ma et al., 2013 and Jin et al., 2016)

already ~~show~~ presented comparison studies and discussed several aspects limiting the ~~degenerated~~ consistency between satellite and MAX-DOAS observations. Concerning the impact of clouds on both MAX-DOAS and satellite retrievals, we separately evaluate the cloud effects on ~~two~~ both observations ~~sides~~. ~~evaluated the cloud effect on the tropospheric TG products. Thus in this study the comparisons for daily average data are performed for different effective cloud fraction (eCF) intervals (Stammes et al., 2008; Wang et al., 2008). Because clouds could also impact the MAX-DOAS results, it is necessary to evaluate the cloud effects on MAX-DOAS and satellite products separately. This issue will also be discussed in this study. We also investigate~~ The weekly cycles observed by satellite instruments and ratios of morning and afternoon values (representing diurnal ~~eyeles~~ variations) ~~observed~~ acquired by combining GOME-2 and OMI observations ~~the satellite instruments~~ are also evaluated by comparison with similar ratios calculated from the corresponding MAX-DOAS results..

Aerosol information is not considered in radiative transfer models (RTM) in-used for the AMF calculations of ~~for~~ most tropospheric satellite products (one exception is the OMI NO<sub>2</sub> product (POMINO) provided by the Peking University over China (Lin et al., 2014)), but recently such aerosol effects have drawn more and more attention. Shaiganfar et al. (2011), Ma et al. (2013), and Kanaya et al. (2014) found negative biases of the OMI tropospheric NO<sub>2</sub> VCDs between 26 and 50 % over areas with high aerosol pollution through the validation by MAX-DOAS observations. But aerosol effects on the satellite retrievals are still not well understood. The aerosols effects can be generally separated sorted into two types contributions: a) the effect of aerosols on the satellite AMF compared to AMFs for a pure Rayleigh atmosphere (absence of explicit aerosol correction), and b) information in AMF calculations and aerosol contamination of cloud products. The the effect of aerosols on the retrieval of cloud products second type is also (often referred to as “implicit aerosol correction”, ~~due to the possible compensation of aerosol induced clouds for the aerosol effects on the satellite AMF calculations~~ (Boersma et al., 2011; Castellanos et al., 2015; Chimot et al., 2016). The two types contributions of of aerosols effects on the satellite retrievals are

discussed in this study based on the aerosol and TG profiles of MAX-DOAS in Wuxi and by comparing the satellite TG VCDs to the corresponding results from the MAX-DOAS observations. Leitão et al. (2010) performed simulation studies and compared the satellite NO<sub>2</sub> AMFs for clear sky with different aerosol scenarios. They found that the influence of aerosols on the satellite AMFs depends mainly on the relative vertical distributions of aerosols and TGs. Recently several studies reported that the OMI and GOME-2 cloud retrievals (eCF and cloud top pressure (CTP)) (Stammes et al., 2008; Wang et al., 2008) are indeed sensitive to the presence of (strong loads of) aerosols (Boersma et al., 2011; Lin et al., 2014; Wang et al., 2015; Chimot et al., 2016). They also claimed that for some cases of heavy aerosol loads the cloud correction can (partly) account for the aerosol effects on the satellite AMFs (referred to as implicit aerosol correction). For example, Castellanos et al. (2015) demonstrated that for biomass burning aerosols extending to high altitudes (about 2 km), the implicit correction can well correct the aerosol effect on the OMI tropospheric NO<sub>2</sub> product. Here it is important to note that the aerosol around the heavily polluted urban region typically resides close to the surface, showing often an overlap with the trace gas profiles. Elevated aerosol layers can e.g. occur if long range transport, e.g. from biomass burning contributes to the local aerosol load (Wang et al., 2016). Simulation studies by Lin et al., 2014 and Chimot et al., 2016 showed that the impact of the implicit correction is quite dependent on the vertical profiles of aerosols and the TGs. Thus, in many cases, the implicit correction might even increase the errors of the AMF. In this study the tropospheric aerosol extinction profiles acquired from MAX-DOAS measurements are used to evaluate the aerosol effects on the satellite observations (not only for NO<sub>2</sub>, but also for SO<sub>2</sub> and HCHO) around heavily polluted urban regions.

The paper is organized as follows: in section 2 we describe MAX-DOAS observations in Wuxi and the satellite products involved in this study. ~~We also discuss the cloud effect on the MAX-DOAS results.~~ In section 3 we compare NO<sub>2</sub>, SO<sub>2</sub> and HCHO VCDs derived from MAX-DOAS with those from the satellite instruments. We investigate in particular ~~the impact of the coincidence criteria and~~ the effects of clouds, SF<sub>6</sub> and aerosols on the satellite retrievals. In section 4 the conclusions are given.

## 2 MAX-DOAS measurements and satellite data sets

### 2.1 MAX-DOAS instrument and data analysis ~~in Wuxi~~

#### 2.1.1 ~~MAX-DOAS instrument and data analysis~~

A MAX-DOAS instrument developed by Anhui Institute of Optics and Fine Mechanics (AIOFM) (Wang et al., 2015 and 2016) is located on the roof of a 11-story building in Wuxi City (Fig. 1 a-1), China (31.57° N, 120.31° E, 50 m a.s.l.) and operated by the Wuxi CAS Photonics Co. Ltd from May 2011 to Dec 2014. Wuxi City is located in the YRD region which is typically affected by high loads of NO<sub>2</sub>, SO<sub>2</sub> and HCHO (Fig. 1 a-2b, a-3c, a-4d). The DOAS method (Platt and Stutz, 2008) and the PriAM profile inversion algorithm (Wang et al., 2013a/b and 2016) are applied to derive the vertical profiles of aerosol extinction (AEs) and volume mixing ratios (VMRs) of NO<sub>2</sub>, SO<sub>2</sub> and HCHO from scattered UV/visible sunlight

recorded by the MAX-DOAS instrument at five elevation angles (5 °, 10 °, 20 °, 30 ° and 90 °). The telescope of the instrument is pointed to the north. The data analysis and the results derived from the MAX-DOAS measurements are already described in our previous study (Wang et al., 2016). In that study we ~~also~~ compared the MAX-DOAS results with collocated independent techniques including an AERONET sun photometer, a visibility meter, and a long path DOAS. The comparisons were done for different cloud conditions as derived from a cloud classification scheme based on the MAX-DOAS observations (Wagner et al., 2014 and Wang et al., 2015). One important conclusion of that study is that meaningful trace gas profiles can be retrieved not only for clear skies, but also for most cloudy conditions (except ~~heavily~~-heavy fog or haze and optically thick clouds). Thus here we use all MAX-DOAS trace gas profiles obtained for these sky conditions (Wang et al., 2016). Here it is important to note that differently from previous studies (e.g. Ma et al., 2013; Jin et al., 2016), we derive the tropospheric VCDs of the TGs by an integration of the vertical profiles, but not by the so-called geometric approximation (e.g. Brinksma et al., 2008). Our previous study (Wang et al., 2016) demonstrated that the tropospheric trace gas VCDs from the full profile inversion are in general much more accurate than those from the geometric approximation. The discrepancy of VCDs between the two methods is systematic and can be mainly attributed to the errors of the geometric approximation, for which the errors can be up to 30% depending on the observation geometry, and the properties of aerosols and TGs. ~~ometries of sun and measurements, and scenarios of aerosols and TGs.~~

### **~~2.1.2 Cloud effect on MAX-DOAS tropospheric VCDs around the satellite overpass time~~**

~~In the validation procedure the MAX DOAS VCDs are averaged over a time period of  $\pm$ one hour around the satellite overpass time. Typically about ten MAX DOAS elevation sequences are recorded during that period, during which the cloud conditions can change. This effect is probably most important for the presence of broken cloud cover. Thus in order to evaluate the cloud effect on MAX DOAS results, we compare the average MAX DOAS VCDs derived from all measurements in  $\pm$ 1 hour around the satellite overpass time with those from the measurements under clear sky conditions only. Sky conditions are derived from MAX DOAS measurements (Wang et al., 2015). The OMI overpass time of 13:30 local time (LT) is selected for the investigation of this effect, and similar features are expected for observations around the GOME-2 overpass time. Fig. 2a, b and c show scatter plots and linear regressions of the average MAX DOAS VCDs from all the measurements in  $\pm$ 1 hour around the satellite overpass time against those under clear sky conditions for NO<sub>2</sub>, SO<sub>2</sub> and HCHO, respectively. Almost 1:1 linear regression lines and correlation coefficients ( $R^2$ ) (the Pearson's product moment correlation coefficient is applied in this paper) close to unity are found for all three species. To quantify the systematic differences of the TG VCDs, the corresponding mean differences (and standard deviations) are displayed for eCF<10% and eCF>10%, respectively. In general larger standard deviations are found for all three species for eCF>10%, indicating that larger deviations are related to larger eCF. Mean differences of  $0.15 \times 10^{15}$  molecules cm<sup>-2</sup>,  $0.02 \times 10^{15}$  molecules cm<sup>-2</sup> and  $0.05 \times 10^{15}$  molecules cm<sup>-2</sup> (corresponding to 0.8%, 0.05% and 0.4% of the average VCDs) are found for NO<sub>2</sub>, SO<sub>2</sub> and HCHO, respectively, indicating that the cloud effect on MAX DOAS results is probably negligible for the satellite validations. Here it should be noted that the shown comparison results represent only situations, for which clear and cloudy~~

~~conditions occur during the two-hour period around the satellite overpass time. Thus we cannot rule out that the errors for measurements under continuous cloud cover are larger. However situations of continuous cloud cover are not relevant for this validation study, because for such conditions no meaningful satellite results can be obtained.~~

## 2.2 NO<sub>2</sub>, SO<sub>2</sub> and HCHO products derived from OMI

5 The OMI instrument (Levelt et al., 2006a, b) aboard the sun-synchronous EOS Aura satellite was launched in July 2004. It achieves daily global coverage with a spatial resolution of 24×13 km<sup>2</sup> in nadir and ~~about 68150×14-13~~ km<sup>2</sup> at the swath edges (Levelt et al., 2006b). The overpass time is around 13:30 LT. In this study, we validate the operational level 2 (Boersma et al., 2007 and 2011) tropospheric NO<sub>2</sub> VCD (DOMINO version 2) obtained from the TEMIS website (<http://www.temis.nl>). The NO<sub>2</sub> SCDs are retrieved in the 405–465 nm spectral window using a DOAS algorithm and are  
10 converted to NO<sub>2</sub> tropospheric VCDs using tropospheric AMFs from a look-up table, which is generated using the DAK radiative transfer model (RTM) (Stammes, 1994), after the stratospheric column was subtracted. SFs of NO<sub>2</sub> for the AMF ~~simulations-calculations~~ are obtained from the TM4 CTM (Williams et al., 2009) for individual measurements and can be downloaded from the TEMIS website. TM4 assimilations run at a resolution of 2° × 3° (lat × lon) and 35 vertical levels up to 0.38 hPa and are spatially interpolated to the OMI pixel center (Boersma et al., 2007 and 2011; Dirksen et al., 2011). The  
15 ~~effective cloud fraction (eCF) (Stammes et al., 2008; Wang et al., 2008)eCF and CTP~~ cloud top pressure (CTP) (Acarreta et al., 2004) are obtained from the OMCLD02 cloud product based on the O<sub>4</sub> absorption band at 477 nm assuming a Lambertian cloud with an albedo of 0.8 (Acarreta et al., 2004). ~~The retrieval algorithm for DOMINO v2 can form~~ forms the basis of NO<sub>2</sub> retrievals for the upcoming TROPospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor mission (Veefkind et al., 2012).

20 Two data sets of tropospheric SO<sub>2</sub> VCDs derived from OMI observations are validated in this study. One is the operational level 2 OMSO2 planetary boundary layer (PBL) SO<sub>2</sub> data set (assuming SO<sub>2</sub> mostly in the PBL) provided via the NASA website (<http://avdc.gsfc.nasa.gov>). In the following this product is simply referred to as “OMI NASA”. For the PBL SO<sub>2</sub> product, the VCD is derived from the OMI-measured radiances between 310.5 and 340 nm using a principal component analysis (PCA) algorithm (Li et al., 2013). ~~A fixed surface albedo (0.05), surface pressure (1013.25 hPa), as well as solar zenith angle (30°) and viewing zenith angle (0°) as well as a fixed climatological SO<sub>2</sub> profile over the summertime eastern U.S. are assumed in the PCA retrievals. A a fixed climatological SO<sub>2</sub> profile over the summertime eastern U.S. profile~~ (Krotkov et al., 2008) ~~is is used in the PBL retrievals for all OMI measurements (Krotkov et al., 2008)~~. The second product is a data set extracted by a new OMI SO<sub>2</sub> retrieval algorithm developed by BIRA (Theys et al., 2015). In the following this product is simply referred to as “OMI BIRA”. It ~~will form~~ forms the basis of the algorithm for the operational level-  
25 2 SO<sub>2</sub> product to be derived from the upcoming ~~TROPospheric Monitoring Instrument (TROPOMI)~~ instrument ~~aboard the Sentinel-5 Precursor mission (Veefkind et al., 2012)~~. SO<sub>2</sub> SCDs are ~~first~~ retrieved in a window between 312–326 nm using the DOAS technique and ~~then a background corrected~~ correction for possible biases ~~is applied~~. The SO<sub>2</sub> SCDs are converted to VCDs using ~~AMFs from a~~ AMF a look-up table, which is generated using the Linearized Discrete Ordinate Radiative

30

Transfer (LIDORT) version 3.3 RTM (Spurr et al., 2001 and 2008). SFs for SO<sub>2</sub> are obtained from the IMAGES CTM (Müller and Brasseur, 1995) for individual measurements at a horizontal resolution of 2° × 2.5° and at 40 vertical unevenly distributed levels extending from the surface to the lower stratosphere (44 hPa) (Stavrou et al., 2013 and 2015). Like for the OMSO<sub>2</sub> data set the cloud information is obtained from the OMCLDO<sub>2</sub> cloud product.

5 The HCHO data set validated in this study is the OMI HCHO tropospheric VCD level 2 data retrieved by a DOAS algorithm v14 developed at BIRA-IASB (De Smedt et al., 2015). This algorithm will also be applied to the upcoming TROPOMI instrument. HCHO SCDs are retrieved in the spectral window between 328.5–346 nm using the DOAS technique. After applying a background corrections, HCHO SCDs are converted to tropospheric VCDs using AMFs from a look-up table generated by LIDORT with HCHO SFs obtained from the IMAGES CTM for individual measurements (Stavrou et al.,  
10 2015). Also for this product the cloud information is obtained from the OMCLDO<sub>2</sub> cloud product.

Here one important aspect should be noted: different AMF strategies are used in the DOMINO v2 NO<sub>2</sub> product and the BIRA SO<sub>2</sub> and HCHO products for eCF < 10%. For the NO<sub>2</sub> product the eCF and CTP are explicitly considered in the AMF simulations while for the SO<sub>2</sub> and HCHO products the clear sky AMFs are applied. These differences will be especially important for measurements in the presence of high aerosol loads (see section 3.65). For eCF > 10%, a cloud correction based  
15 on the independent pixel approximation (IPA) (Cahalan et al., 1994) is applied for the three TG retrievals. It should also be noted that observations of the outermost pixels (i.e. pixel numbers 1–5 and 56–60) and pixels affected by the so called “row anomaly” (see <http://www.temis.nl/docs/omiwarning.html>) were removed before the comparisons.

### 2.3 NO<sub>2</sub>, SO<sub>2</sub> and HCHO products derived from GOME-2

The GOME-2A and B instruments (Callies et al., 2000; Munro et al., 2006, 2016) are aboard the sun-synchronous  
20 Meteorological Operational Satellite platforms MetOp-A and MetOp-B, respectively. MetOp-A (launched on 19 October 2006) and MetOp-B (launched on 17 September 2012) operate in parallel with the same equator crossing time of 09:30 LT. Before 15 July 2013 GOME-2A had ~~the a~~ swath width of 1920 km, corresponding to a ground pixel size of 80 km × 40 km and a global coverage within 1.5 days. Since 15 July 2013, the GOME-2A swath width was changed to 960 km with a ground pixel size of 40 km × 40 km. The GOME-2A settings before 2013 are also applied to GOME-2B.

25 In this study, we validate the operational level 2 tropospheric NO<sub>2</sub> VCDs derived from the TM4NO<sub>2</sub>A version 2.3 product (Boersma et al., 2004) for GOME-2A ~~and GOME-2~~ and B obtained from the TEMIS website. The NO<sub>2</sub> SCDs are retrieved in the 425–450 nm spectral window at BIRA with QDOAS (<http://uv-vis.aeronomie.be/software/QDOAS/>). The tropospheric NO<sub>2</sub> VCDs are ~~obained~~ obtained from SCDs using the similar data assimilation procedures as for the DOMINO v2 product. However, for the GOME-2 products the eCF and CTP are retrieved by the improved Fast Retrieval Scheme for Clouds from  
30 the Oxygen A-band algorithm (FRESCO+) based on the measurements of the oxygen A-band around 760 nm (Wang et al., 2008) again assuming a Lambertian cloud.

Two SO<sub>2</sub> products derived from GOME-2A observations are included in the study. The first one is the operational level 2 O3M-SAF SO<sub>2</sub> product derived from GOME-2A observations (Rix et al., 2012 and Hassinen et al., 2016). In the following the product is simply referred to as “GOME-2A DLR”. This product is provided via the EUMETSAT product navigator (<http://navigator.eumetsat.int>) or the DLR EOWEB system (<http://eoweb.dlr.de>). The SO<sub>2</sub> SCDs are retrieved using the DOAS technique in the wavelength range between 315 and 326 nm. For the conversion of SCDs to VCDs, the AMFs are acquired from a ~~AMF~~ look-up table, ~~which is~~ generated using LIDORT 3.3. For the AMF computation, three types of SFs are assumed as Gaussian distributions with a FWHM of 1.5km around three central heights of 2.5km, 6km and 15km. Because for the SO<sub>2</sub> concentrations at Wuxi mostly anthropogenic pollutions is relevant, only the SO<sub>2</sub> product corresponding to the central height of 2.5 km is included in the validation study. The cloud information is obtained from GOME-2 measurements by the OCRA and ROCINN algorithms (Loyola et al., 2007) based on oxygen A-band observations at around 760 nm. The second product is provided from BIRA using the same retrieval algorithm as for the OMI BIRA SO<sub>2</sub> product, referred to as “GOME-2A BIRA”. The same algorithm is also used to acquire the SO<sub>2</sub> data from GOME-2B observations. The product is referred to as “GOME-2B BIRA” in the following. The cloud properties used in the two products are derived from GOME-2A/B observations using the FRESCO+ algorithm.

The HCHO tropospheric VCD level 2 products derived from GOME-2A and B observations (De Smedt et al., 2012 and 2015) are validated in this study. The same retrieval approach as for the OMI BIRA HCHO product is applied, but the cloud properties are derived from GOME-2A/B observations using the FRESCO+ algorithm.

### 3 Validation of the satellite data sets

In this section the daily and bi-monthly averaged NO<sub>2</sub>, SO<sub>2</sub>, and HCHO VCDs from OMI and GOME-2 are validated by comparisons with the tropospheric VCDs derived from the MAX-DOAS observations. Here it needs to be clarified that the daily and bi-monthly averaged satellite data are the averaged values of all the satellite pixels located in the coincidence area around the measurement site (see below). And the averaged MAX-DOAS data are the averaged values for those all measurements within 2 hours around the satellite overpass time. Also the diurnal and weekly cycles from satellite observations are compared with those from the corresponding MAX-DOAS observations. Finally the influence of the SF and the effects of aerosols on the OMI products are discussed. The SFs from the CTM used for the OMI AMF calculations are compared to the SFs derived from MAX-DOAS.

Averaging of individual satellite and/or MAX-DOAS observations can be advantageous for several reasons. First, especially for observations with rather large statistical uncertainties (in particular for satellite observations of SO<sub>2</sub> and HCHO), the merging of several observations can substantially reduce these uncertainties. Second, the effect of spatial gradients across satellite pixels can be partly accounted for by averaging MAX-DOAS measurements over a period around the satellite overpass time. However, for the averaging of satellite and MAX-DOAS data reasonable coincidence selection criteria need to be determined, can be different for the different TGs and satellite sensors. ~~for the following comparisons of satellite and~~

~~MAX-DOAS observations. MAX-DOAS results need to be averaged over a time period around the satellite overpass time to (partly) compensate the effect of horizontal gradients of the TG concentrations. Meanwhile satellite observations need to be extracted based on the distance of satellite pixel centre to the MAX-DOAS measurement site and averaged on each day to minimise the random noise of the satellite data for daily comparisons.~~ The effects of the selection criteria, in particular the ~~average~~ time period used for the MAX-DOAS measurements and the distance of the selected satellite observations from the measurement site ~~on the validation activities~~ are evaluated and discussed in detail in section S1 in the supplement. On general finding is that ~~In general the effect of~~ the effect of the chosen time period is negligible compared to the effect of the chosen distance. Therefore it is reasonable to arbitrarily use 2 hours around the satellite overpass time, namely 12:30 LT to 14:30 LT for the comparisons with OMI and from 08:30 LT to 10:30 LT for the comparisons with GOME-2A/B. The ~~coincidence~~ distances around the measurement site, for which satellite observations are averaged are ~~of~~ chosen differently for the different satellite products ~~are discreetly determined~~ based on the sensitivity studies in section S1 in the supplement. In the following comparison activities, the OMI NO<sub>2</sub> and SO<sub>2</sub> (HCHO) data are selected for satellite pixels with ~~the~~ distances of <20km (<50km) from the Wuxi station. ~~And~~ The GOME-2A/B data of the three species are selected for ~~the~~ distances of <50km. It should be noted that these findings are derived for a polluted site in China. For other locations and conditions, ~~different coincidence criteria might be best suited.~~

### ~~3.1 Effects of variations of the coincidence criteria on the validation~~

~~Because of the large ground pixel size of the satellite observations, MAX-DOAS results are averaged over a time period around the satellite overpass time to (partly) compensate the effect of horizontal gradients of the TG concentrations. In principle the time period is a function of the satellite pixel size, the wind speed and the life time of the trace gases. Although some factors change frequently, here we use one fixed time period for the long term comparisons for simplicity. In this study, we test the effect on the satellite validation for four time periods including 1 hour, 2 hours, 3 hours and 4 hours around the satellite overpass time. Scatter plots of the average MAX-DOAS data over three time periods (1 hour, 3 hours and 4 hours) against those over 2 hours are shown in Fig. 3. The correlation coefficients are close to unity for all time periods. However, the slopes become systematically smaller for larger time periods (up to 10%) because of temporal smoothing. The results of the linear regressions and mean relative differences from the comparisons are also shown in Fig. 5a and will be discussed below together with the effect of the selected coincidence area of the satellite products.~~

~~In principle for the satellite validation the satellite pixel closest to the MAX-DOAS instrument need to be selected. However, in order to minimise the random noise of the satellite data, it is useful to calculate the average of several satellite observations close to the measurement site (see e.g. Irie et al., 2012 and Ma et al., 2013). As selection criterion, a distance between the centre of the satellite pixel and the measurement site can be specified. This optimum distance depends on many factors, such as the satellite ground pixel size, the selected time period over which the MAX-DOAS results are averaged, the expected horizontal gradients of the trace gas and the statistical uncertainty of the satellite data. A distance of < 20 km has~~

been used for NO<sub>2</sub> comparisons (e.g. Ma et al., 2013 and Chan et al., 2015), 100 km for HCHO (De Smedt et al., 2015) and SO<sub>2</sub> (Theys et al., 2015). Irie et al. (2012) already found that the correlations and slopes of the linear regressions of the NO<sub>2</sub> tropospheric VCDs from OMI and GOME 2A against those from MAX DOAS observations depend systematically on the distance to the MAX DOAS station:

5 We test the effect of the variation of the distance between 10 km to 75 km on the comparison between the satellite data (OMI and GOME 2) and the MAX DOAS data for all three TGs. The areas for the four radii (10km, 20km, 50km and 75km) and the pixel sizes of OMI and GOME 2 are shown in the earth view image downloaded from the Google Earth service in Fig. 1 b 1. For distances larger than 20 km, the cities of Suzhou, Changzhou, Huzhou and Nantong are included in the area. Because of transport of the pollutants between the cities and the different residence times, different horizontal distributions of the NO<sub>2</sub>, SO<sub>2</sub> and HCHO VCDs are found around Wuxi as shown in Fig. 1 b 2, b 3 and b 4, respectively. HCHO has a smoother distribution than SO<sub>2</sub>, which is smoother than NO<sub>2</sub>. The satellite data for pixels with the distances of 0-10km, 10-20km, 20-50km and 50-75km to the MAX DOAS station are compared with the MAX DOAS results.

10 We compare both the results for individual satellite pixels and daily averages for the four radii with the average MAX DOAS data over 2 hours around the satellite overpass time. The comparisons for OMI NO<sub>2</sub>, SO<sub>2</sub> and HCHO for pixels with distances of 0-10km, 10-20km, 20-50km and 50-75km are shown in Fig. 4a, b and c, respectively (the comparisons for pixels with the distances of <10km, <20km, <50km and <75km are shown in Fig. S1 in the Supplement). We use the SO<sub>2</sub>-OMI product from BIRA for this study, because it shows in general a higher correlation with the MAX DOAS data. We found that the linear regressions for the daily averaged data are quite similar to those for the individual pixel data. Only the correlation coefficients are higher. The results of the linear regressions and the mean relative differences for the two distance categories as indicated in Fig. 4 and in Fig. S1 in the Supplement are shown in Fig. 5 b and c, respectively. The slopes decrease with increasing distance for the three gases. The decrease of the slopes (from 0.75 to 0.49 and R<sup>2</sup> from 0.66 to 0.29) are stronger for NO<sub>2</sub> than for SO<sub>2</sub> and HCHO. This finding is consistent with the typically stronger horizontal inhomogeneity of NO<sub>2</sub>. The mean differences for HCHO show almost no dependence on the distance. This finding can be explained by the more homogenous distribution of HCHO compared to NO<sub>2</sub> and SO<sub>2</sub>. A significant decrease of the slopes from 0.73 to 0.50 and the R<sup>2</sup> from 0.65 to 0.44 is found for NO<sub>2</sub> with increasing distance over 20km. A decrease of the slope is also found for SO<sub>2</sub> for the distances larger than 20km. From these findings we conclude that 20km is a reasonable distance to select OMI NO<sub>2</sub> and SO<sub>2</sub> data for conditions similar to those at Wuxi. In contrast, for HCHO we select a distance of 50 km. Although for such distances the slope is smaller than for shorter distances, we find nearly identical mean differences. Because of this finding and the rather high noise of the HCHO satellite data we select a distance of 50 km, for which the number of available measurements largely increases. The comparison of Fig. 5a and b indicates that the effect of time periods used for averaging the MAX DOAS results on the validation study is much smaller than the effect of distances for selecting the satellite data. Thus we apply the time period of 2 hours around the satellite overpass time in this study.

30 Similar results for GOME 2 data as those for OMI shown in Fig. 5 are shown in Fig. 6. The O3M SAF GOME 2A SO<sub>2</sub> product from DLR is used for this sensitivity study. Also for the GOME 2 SO<sub>2</sub> data set the effect of the horizontal

coincidence criterion is larger than the effect of the time period for the averaging of the MAX-DOAS data is found. Thus also 2 hours around the satellite overpass time will be used for GOME-2 comparisons in this study. The largest changes of the slopes for the three trace gases are found around the distance of 10km, but the results for the selection criterion of 0-10km should be treated with care because of the low number of available measurements. The changes of the slopes for distances larger than 20km are smaller than 0.06 for NO<sub>2</sub> and 0.04 for HCHO, but are larger for SO<sub>2</sub>. However, the results of the linear regressions for SO<sub>2</sub> should again be treated with care because of the rather low correlation coefficients. From these results we select 50km as a reasonable distance for GOME-2 data of NO<sub>2</sub>, SO<sub>2</sub> and HCHO.

In summary, in the following validation studies, the MAX-DOAS results are selected within the period from 12:30 LT to 14:30 LT for the comparisons with OMI and from 08:30 LT to 10:30 LT for the comparisons with GOME-2A/B. The OMI NO<sub>2</sub> and SO<sub>2</sub> (HCHO) data are selected for satellite pixels with the distance of <20km (<50km) from the Wuxi station. The GOME-2A/B data of the three species are selected for the distances <50km.

### 3.2.1 Daily comparisons

The daily averaged satellite data for measurements within the chosen distances (see section 3.1) are compared with the daily averaged MAX-DOAS data within 2 hours around the satellite overpass time. To characterize the cloud effect on the comparisons, the comparisons are performed for different eCF bins of 0-10%, 10-20%, 20-30%, 30-40%, 40-50% and 50-100% for NO<sub>2</sub> and SO<sub>2</sub>, and for eCF bins of 0-10%, 10-30%, 30-50% and 50-100% for HCHO. In addition— Note that the cloud effects on the MAX-DOAS results are discussed in detail in section S2 of the supplement. We found the cloud effect on MAX-DOAS results is negligible for the satellite validation activities.

#### 1) NO<sub>2</sub>

Figures 7a2a, b and c display scatter plots (and the parameters from the linear regressions) of the daily averaged NO<sub>2</sub> tropospheric VCDs derived from OMI, GOME-2A and GOME-2B products versus those derived from the corresponding MAX-DOAS measurements for eCF < 10%. Systematically—Generally higher correlation coefficients (R<sup>2</sup>) for OMI than for GOME-2A/B are found. The systematic biases of the satellite data with respect to the MAX-DOAS data are quantified by the mean relative difference (MRD) calculated following Eq. 1:

$$\text{MRD} = \frac{\sum_1^n (V_{s_i} - V_{M_i}) / V_{M_i}}{n} \quad (1)$$

Here  $V_{s_i}$  and  $V_{M_i}$  represent the averaged TG VCDs from satellite observations and MAX-DOAS measurements on day  $i$ , respectively;  $n$  is the total number of the available days. The MRD is only 1% for OMI, and 27% and 30% for GOME-2A and GOME-2B, respectively.

The R<sup>2</sup>, slopes and intercepts of the linear regressions, and the MRD as well as the number of available days for the three satellite products are shown for the five eCF bins in Fig. 83. For OMI, R<sup>2</sup> decreases with increasing eCF; the slopes significantly change for eCF > 50% and the MRD drops to -40% for eCF > 40%. For GOME-2A, a steep decrease of R<sup>2</sup> for eCF > 30% is found. For GOME-2B, a generally lower R<sup>2</sup> is found for eCF > 30%; the MRD indicates an increasing

systematic overestimation for  $eCF > 30\%$ . Thus we conclude that the cloud effect on OMI and GOME-2A/B  $NO_2$  data becomes critical-significant for  $eCF > 40\%$  and  $30\%$ , respectively.

## 2) $SO_2$ :

Figures 9a4a, b, c, d, and e display scatter plots of the daily averaged  $SO_2$  tropospheric VCDs derived from the OMI NASA, OMI BIRA, GOME-2A DLR, GOME-2A and B BIRA products versus those derived from the corresponding MAX-DOAS measurements for  $eCF < 10\%$ .  $R^2$  and slopes are more close to unity for the OMI BIRA product than for the other products. The MRDs indicate a similar systematic underestimation ( $-40\%$  to  $-52\%$ ) by all products. There are fewer negative values in the OMI BIRA product than in other satellite products. It needs to be noted that the significantly worse  $R^2$  for the OMI NASA product than compared to for the OMI BIRA product could partly be attributed to the assumed one-fixed measurement condition and  $SO_2$  profile (and thus the fixed AMF) in the NASA PCA retrievals. However the similar slopes and MRDs between the two OMI products indicate that the simplification of the NASA PCA retrieval only slightly contributes to the systematic bias of the averaged values.

The  $R^2$ , slopes and intercepts of the linear regressions, the MRD as well as the number of the available days obtained for the five satellite  $SO_2$  products are shown for the five  $eCF$  bins in Fig. 495. For the OMI BIRA product, a significant decrease of  $R^2$  occurs for  $eCF > 10\%$  together with a decrease of the slopes and the MRD. A steep increase of the MRD is found for  $eCF > 40\%$ . Thus-Therefore we conclude cloud effects on that the OMI BIRA  $SO_2$  data are become considerable most accurate for  $eCF \leq 10\%$ , while they might be still used for  $eCF$  of  $10\%$  to  $40\%$  with a  $20\%$  larger systematic negative bias than those for  $eCF < 10\%$ . For the OMI  $SO_2$  NASA data,  $R^2$ , slope, and MRD significantly decrease for  $eCF > 20\%$ .  $R^2$  for both GOME-2A data are low ( $< 0.09$ ) for all  $eCF$  bins, thus from the linear regressions cannot is not helpful to yield no meaningful information on the cloud effect can be derived. Almost constant MRDs are found for both GOME-2A  $SO_2$  products for  $eCF < 30\%$ . For  $eCF > 30\%$  largely varying MRD are found, especially for the GOME-2A BIRA products. Thus we conclude that the cloud effects on both GOME-2A products are most are appreciable accurate for  $eCF \leq 30\%$ . For the GOME-2B BIRA data, an obvious decrease of  $R^2$  and slope is found for  $eCF > 10\%$ , while for  $eCF > 30\%$  largely variable MRDs are found. Thus for clouds can considerably impact the GOME-2B BIRA product -for we recommend to use observations with  $eCF \leq 10\%$ .  $SO_2$  VCDs for  $eCF < 30\%$  might also be used, but are subject to larger uncertainties, and more significantly for  $eCF > 30\%$ .

## 3) HCHO:

Because of the rather low-small atmospheric absorption of HCHO, the DOAS fit errors often dominate the total uncertainty of the HCHO satellite data (De Smedt et al., 2015). Thus systematic effects, e.g. caused by clouds, are more difficult to identify and quantify than for  $NO_2$  and  $SO_2$ . Figure 11 shows that the scatter plots of OMI HCHO VCDs for individual pixels versus those derived from MAX-DOAS observations for  $eCF < 30\%$  is shown in Fig. 6. One important finding is that the  $R^2$  for data with a fit error  $< 7 \times 10^{15}$  molecules  $cm^{-2}$  is better than the  $R^2$  for all data (see Fig. 6b). And the A similar phenomenon result is also found obtained for the daily averaged OMI HCHO VCDs (see Fig. S12 in the supplement) indicating -This indicates that the fit error dominates the random noise-uncertainty of satellite HCHO tropospheric VCDs.

~~Meanwhile~~In contrast, the slopes of the linear regressions for the OMI data before and after the filtering are quite similar as shown in Fig. 6b and supplementary Fig. S12. Thus the data screening ~~will not~~has no considerably impact on the analysis of the systematic bias of OMI HCHO products. Considering that ~~the~~the mean fit error of the HCHO VCDs is  $7 \times 10^{15}$  molecules  $\text{cm}^{-2}$  for OMI data, ~~Thus~~for further comparisons, we exclude the HCHO VCDs with ~~the~~fit error  $> 7 \times 10^{15}$  molecules  $\text{cm}^{-2}$  for OMI. However for the GOME-2A/B products, the filter for the fit error is not applied because in contrast to the OMI HCHO data we find a systematic dependence of the fit error on the retrieved HCHO tropospheric VCD (see Fig. ~~S2-S13~~ in the supplement). ~~The different phenomenon~~findings with respect ~~of~~to the HCHO fit errors ~~between~~for OMI and GOME-2 ~~is~~are not clearly understood and should be addressed ~~. It might be valuable for a~~in further investigations.

If the additional filter of the fit error for the OMI product is applied, 48% of the total number of HCHO data is left for comparisons. In order to still include sufficient numbers of data, we use broader eCF bins (0-10%, 10%-30%, 30%-50% and 50%-100%). Figures ~~12a~~7a, c and d display scatter plots of the satellite daily averaged data versus the MAX-DOAS data for eCF  $< 10\%$  for OMI, GOME-2A, and GOME-2B data, respectively. We found the best consistency for the GOME-2B product probably because of the weaker degradation of ~~the GOME-2B instrument~~during the short time after launch ~~than~~compared to OMI and GOME-2A. Nevertheless also other unknown reasons might play a role. One interesting finding is the better correlation of the OMI products for the eCF bin of 10% to 30% (see Fig. ~~12b~~7b) compared to the eCF  $< 10\%$ . However, for eCF of 10% to 30% also a larger MRD of -34% (see Fig. ~~13~~8) is found, which might be attributed to the special effect of clouds, namely the clear sky AMFs used in the retrievals for eCF  $< 10\%$  (see the last paragraph of section 2.2).

The dependencies of the results of the linear regressions and the MRDs on the eCFs are shown in Fig. ~~13~~8 for the three satellite instruments. For ~~the OMI product~~a decrease of  $R^2$  occurs for eCF  $> 30\%$ , while for GOME-2A and GOME-2B, low  $R^2$  are already found for eCF  $> 10\%$ . Gradually increasing absolute values of the MRDs for all ~~the~~satellite instruments are found for increasing eCF. ~~We suggest~~In general cloud effects on ~~that the~~ HCHO products become significantsubstantial ~~for~~for eCF  $\leftrightarrow 30\%$  ~~should be used~~for the three satellite instruments. However it needs to be noted that our findings are derived for one location (Wuxi) and might not be fully representative for other locations. The use of the HCHO products with eCF  $< 40\%$  areis recommended ~~to be used~~by the retrieval algorithm developer (De Smedt et al., 2015).

### **3.3.2 Errors of ~~Shape~~shape Factors factors from CTM and the effect on satellite VCD products**

The SF is an input for the calculation of satellite AMF, which is needed to convert the SCD to VCD (Palmer et al., 2001). Different retrieval algorithms acquire the SFs in different ways, mostly from a CTM for individual measurements or assuming a fixed SF (see section 2.2 and 2.3). The MAX-DOAS measurements acquire the vertical profiles of  $\text{NO}_2$ ,  $\text{SO}_2$  and HCHO from the ground up to the altitude of about 4km (depending on the measurement conditions), in which the tropospheric amounts of the TGs is mostly concentrated. Thus the profiles derived from MAX-DOAS observations are valuable to evaluate the SFs used in the satellite retrievals and their effect on the AMFs and VCDs. ~~Because the averaging~~

kernels and SFs for individual satellite measurements are available only for the DOMINO NO<sub>2</sub>, BIRA SO<sub>2</sub> and BIRA HCHO products derived from OMI observations, the three products are used to evaluate the effect of the SF in this section.

For the three selected products, the calculation of satellite tropospheric AMFs follows the same way introduced in Palmer et al. (2001) as Eq. 2:

$$5 \quad \text{AMF} = \int_{\text{ground}}^{\text{tropopause}} \text{BAMF}(z) \text{SF}(z) dz \quad (2)$$

~~Where~~ Here BAMF(z) is the box AMF, which characterizes the measurement sensitivity as a function of altitude (z). The integration is done from the ground to the tropopause. The SFs of the TGs are obtained from different CTM (TM4 for NO<sub>2</sub>, IMAGES for SO<sub>2</sub> and HCHO, see section 2.2). The profiles (profile<sub>M</sub>) derived from MAX-DOAS can be converted to SF (SF<sub>M</sub>) using Eq. 3:

$$10 \quad \text{SF}_M(z) = \frac{\text{profile}_M(z)}{\text{VCD}_M} \quad (3)$$

where VCD<sub>M</sub> is the tropospheric VCD derived by an integration of the corresponding profile<sub>M</sub>. It needs to be noted that only the profiles below 4km can be reliably drawn from MAX-DOAS observations. Thus the profile<sub>M</sub> between 4km and the tropopause (a fixed value of 16 km is used in this study) are derived from the corresponding CTM profiles of the individual satellite data sets. Therefore the SF<sub>M</sub> is derived from the combined profile<sub>M</sub> using Eq.3. It needs to be note that considering that only the profiles below 4km can be drawn from MAX-DOAS observations, the profile<sub>M</sub> in the layer from 4km up to tropopause (a fixed value of 16 km is used in this study) are derived from the corresponding CTM profiles corresponding to the individual satellite data. Therefore the SF<sub>M</sub> is derived from the combined profile<sub>M</sub> using Eq.3.

A similar relationship connects the BAMFs and averaging kernels (Eskes and Boersma, 2003):

$$15 \quad \text{AK}(z) = \frac{\text{BAMF}(z)}{\text{AMF}} \quad (4)$$

20 The SF<sub>M</sub> can replace the SF from CTM (SF<sub>C</sub>) to recalculate the AMF using Eq. 2. A similar study was recently conducted by Theys et al. (2015) and De Smedt et al. (2015) for OMI BIRA SO<sub>2</sub> and HCHO products over the Xianghe area. They demonstrated the improvements of the consistency between OMI VCDs and MAX-DOAS VCDs when using the SF<sub>M</sub> for the AMF calculation of the satellite products by 20%-50%. In our study we follow the same procedure.

### 1) NO<sub>2</sub>

25 The averaged NO<sub>2</sub> SF<sub>C</sub> for the measurements under clear sky with eCF < 10%, is compared to SF<sub>M</sub> in the altitude range of up to 4km in Fig. ~~14a9a~~. The differences between the averaged SF<sub>C</sub> and SF<sub>M</sub> shown in Fig. ~~14b-9b~~ indicate that in the layer below 4km the NO<sub>2</sub> SF<sub>C</sub> is considerably larger than SF<sub>M</sub> below 0.4 km and smaller than SF<sub>M</sub> in the layer below and above 0.4 km in the layer below 4km, respectively. The differences in the altitude range above 4km SF<sub>C</sub> is slightly larger than SF<sub>M</sub> are also shown in (see supplementary Fig. S14a). ~~is slightly larger than SF<sub>M</sub> in the altitude range~~ The OMI VCDs (VCD<sub>CTM</sub>) ~~from the DOMINO NO<sub>2</sub> product retrieved with based on SF<sub>C</sub> (directly derived from the DOMINO NO<sub>2</sub> product) —and the modified OMI VCDs (VCD<sub>SM</sub>) (based on SF<sub>M</sub> and the NO<sub>2</sub> SCDs which are derived from the DOMINO NO<sub>2</sub> product, but based on SF<sub>M</sub>)~~ SF<sub>M</sub> are plotted against the VCDs derived from MAX-DOAS observations in Fig. ~~14e9c~~. Very similar

results for both  $VCD_{CTM}$  and  $VCD_{SM}$  are found. In Fig. 14e-9e the relative differences of the AMFs using either  $SF_C$  ( $AMF_{CTM}$ ) or  $SF_M$  ( $AMF_{MAX-DOAS}$ ) are shown. The differences are calculated in two ways: either the relative differences are first calculated for individual measurements, and then the individual relative differences are averaged. Alternatively first the AMFs of the individual measurements are averaged, and then the relative differences are calculated.

The results in Fig. 14e-9e show that also for both calculations very similar results are obtained. For  $eCF < 10\%$  the relative differences are only 0.3%. The small differences can be explained by the compensation effect of the negative and positive differences between  $SF_C$  and  $SF_M$  near the surface and at high altitudes, respectively, contributes to the negligible SF effect on the AMF.

For different eCF bins, the relative differences of  $AMF_{CTM}$  and  $AMF_{MAX-DOAS}$  increase systematically with increasing eCF. For  $eCF < 10\%$  the relative differences are only 0.3%. The compensation of the negative and positive difference between  $SF_C$  and  $SF_M$  near the surface and at high altitudes contributes to the negligible SF effect on the AMF. The stronger effect of the SF on the AMF under cloudy sky conditions This phenomenon finding can be explained by the fact that the box AMF below the cloud decrease strongly the partial AMF above 4km. (see. This is The partial AMFs below and above 4km are shown in supplementary Fig. S14c). The partial  $AMF_{CTM}$  is always larger than the partial  $AMF_{MAX-DOAS}$  above 4km due to the larger  $SF_C$  than compared to  $SF_M$ . And the difference is increases substantially larger for a larger with increasing eCF. Meanwhile the weighting contribution of the partial AMF above 4km into the total tropospheric AMF become significant for increases a large with increasing eCF due to the strong decrease of the partial AMF below 4km with an increase of eCF. And the decrease of the lower partial AMF attributes to the decrease of BAMF along an increase of eCF (see Fig. 9d). However no substantial dependence of BAMF on eCF occurs in the altitudes above 4km (see Fig. S14b in the supplement). the latitude range with the largest differences between  $SF_C$  and  $SF_M$ . Overall the overestimation of the partial  $AMF_{CTM}$  compared to the partial  $AMF_{MAX-DOAS}$  above 4km become critical under cloudy conditions.

In general the TM4  $NO_2$  a-priori profile shapes agree well with the MAX-DOAS profiles on average, and the agreement with MAX-DOAS VCDs by replacing  $SF_C$  with  $SF_M$  in the AMF calculation is only slightly improved for all the eCF bins a small eCF. For a large eCF,  $VCD_{SM}$  is systematically larger than  $VCD_{CTM}$  by 20% on average (see Fig. 3-), It is consistent with the AMF differences shown in Fig. 14e9e.

## 2) $SO_2$

The results shown in Figures 105a and b indicate that in the layer below 4km for  $eCF < 10\%$ , the  $SO_2$   $SF_C$  is considerably smaller than  $SF_M$  below 1 km and larger than  $SF_M$  in the layer below and above 1 km in the altitude range of up to 4km, respectively. And as can be seen in the supplementary Fig. S15a,  $SF_C$  is in general slightly larger than  $SF_M$  in the altitude range above 4km in general. Since the BAMFs increase with altitude (Fig. 14d10d)  $SO_2$   $AMF_{CTM}$  are on average larger than  $AMF_{MAX-DOAS}$  by 18% (Fig. 150e). In contrast to  $NO_2$ , the  $SO_2$   $VCD_{SM}$  agrees better with the MAX-DOAS VCDs than  $VCD_{CTM}$ , i.e.  $R^2$  and slope increase from 0.47 to 0.60 and from 0.55 to 0.90, respectively (see Fig. 15e10c). Also the

systematic bias of  $VCD_{SM}$  is smaller than that of  $VCD_{CTM}$ , i.e. the MRD is -26% for  $VCD_{SM}$  and -40% for  $VCD_{CTM}$  (see black and red curves in Fig. 405).

3) For different eCF bins, the differences between  $SO_2$   $SF_C$  and  $SF_M$  (Fig. 45b10b) as well as BAMFs (Fig. 10d) are slightly different from each other in the altitude range below 4km. However an obvious dependence of ~~two~~ both quantities above 4km on eCF can be seen in supplementary Fig. S15a and b. The overestimation of  $SF_C$  compared to  $SF_M$  above 4km ~~mainly~~ increases with ~~an increase of~~ increasing eCF, ~~and the BAMF above 4km is also larger~~ and the BAMF above 4km ~~is also larger~~ increases for a large with increasing eCF. Therefore the dependences of ~~two~~ both quantities on eCF dominates the different levels of ~~agreement~~ discrepancy of the partial  $AMF_{CTM}$  and partial  $AMF_{MAX-DOAS}$  above 4km ~~under~~ for different cloud conditions as shown in supplementary Fig. S15c. Furthermore, the ~~upper~~ partial AMF above 4km dominates the total tropospheric AMF for a large eCF due to the decrease of the lower partial AMF with ~~an increase~~ increasing eCF (see supplementary Fig. S15c). ~~Therefore the~~ These dependencies of ~~differences of upper~~ partial  $AMF_{CTM}$  and partial  $AMF_{MAX-DOAS}$  above 4km on eCF ~~dominates~~ also explain the dependencies of ~~and the~~ BAMFs for large eCF are larger and smaller than those for low eCF at high and low altitudes, respectively. ~~Also the relative~~ differences between the total tropospheric  $AMF_{CTM}$  and  $AMF_{MAX-DOAS}$  ~~depend~~ on eCF as shown in Fig. 10e, ~~with larger differences for large eCF.~~ However, in general the dependence ~~of the~~ differences on eCF is smaller than that for  $NO_2$ . ~~Here it is interesting to note that~~ In addition the results shown in Figure 10 ~~also showed~~ a better consistency between the  $SO_2$   $VCD_{SM}$  and the MAX-DOAS VCDs than for the  $VCD_{CTM}$  can be seen in Fig. 5 for all the eCF bins.

#### 4) ~~3)~~ HCHO

The results shown in ~~Figure Fig.~~ Fig. 46a-11a and b indicate that in the altitude range below 4km for eCF < 10% the  $HCHO_{SF_C}$   ~~$SF_C$~~  is considerably smaller below 1.7 km and larger than  $SF_M$  ~~in the layer below and above 1.7 km~~ in the altitude range below 4km, respectively. As can be seen in the ~~in~~ supplementary Fig. S16a ~~nd~~ the  $SF_C$  almost equals ~~as~~ the  $SF_M$  above 4km as shown in supplementary Fig. S16a for eCF < 10%. Since the BAMF increases with altitude (Fig. 46d11d) the  $HCHO$   $AMF_{CTM}$  is on average larger than  $AMF_{MAX-DOAS}$  by 11% (Fig. 46e11e). Like for  $SO_2$  the  $VCD_{SM}$  agrees better with the MAX-DOAS VCD than  $VCD_{CTM}$ , i.e.  $R^2$  and slope increase from 0.15 to 0.21 and from 0.44 to 0.61, respectively (see Fig. 46e11c). Also the systematic bias of  $VCD_{SM}$  is smaller than that of  $VCD_{CTM}$ , i.e. the MRD is -10% for  $VCD_{SM}$  and -18% for  $VCD_{CTM}$  (see Fig. 438).

For different eCF bins, larger differences between  $AMF_{CTM}$  and  $AMF_{MAX-DOAS}$  are found for a towards larger eCF (see Fig. 11e). Similar to  $NO_2$  and  $SO_2$ , ~~this finding~~ ~~phenomenon is dominated~~ is caused by the partial AMFs above 4km ~~the~~. The dependences of differences between  $HCHO_{SF_C}$  and  $SF_M$  (Fig. 11b) as well as BAMFs (Fig. 11d) in the altitude range below 4km on eCF are insignificant. However both quantities above 4km obviously depend on eCF (see supplementary Fig. S16a and b). The overestimation of  $SF_C$  compared to  $SF_M$  above 4km increases with increasing eCF, and the BAMF above 4km also increases with increasing eCF. Therefore the dependences of both quantities on eCF dominates the different levels of agreement of the partial  $AMF_{CTM}$  and partial  $AMF_{MAX-DOAS}$  above 4km for different cloud conditions as shown in

supplementary Fig. S16c. Furthermore, the partial AMF above 4km dominates the total tropospheric AMF for large eCF due to the decrease of the ~~lower~~ partial below 4km AMF with increasing eCF (see supplementary Fig. S16c). These dependencies of partial  $AMF_{CTM}$  and partial  $AMF_{MAX-DOAS}$  above 4km on eCF also explain the dependencies of differences between the total tropospheric  $AMF_{CTM}$  and  $AMF_{MAX-DOAS}$  on eCF as shown in Fig. 11e. ~~The dependences of differences between HCHO SF<sub>e</sub> and SF<sub>M</sub>, and BAMFS on the eCFs; the dependences of differences between HCHO SF<sub>e</sub> and SF<sub>M</sub>, and BAMFS on the eCFs are found similar to those of SO<sub>2</sub> (see Fig. 16b and d). Again, the large relative differences between  $AMF_{CTM}$  and  $AMF_{MAX-DOAS}$  are found for large eCF (see Fig. 16e).~~ In addition, Figure ~~figure 16e~~ 8 shows that for all the eCF bins the consistency between VCD<sub>SM</sub> and the MAX-DOAS VCD is better than for VCD<sub>CTM</sub>.

#### 5) 4) Uncertainties of the SF from MAX-DOAS

The previous study on Wuxi MAX-DOAS observations (Wang et al., 2016) demonstrated that the profile retrievals are not sensitive to altitudes above 1-2km, where the retrieved profiles are strongly constrained to the a-priori profiles. Thus the SFs at high altitudes could be underestimated by MAX-DOAS retrievals. This effect could be considerable especially for SO<sub>2</sub> and HCHO, because they typically extend to higher altitudes than NO<sub>2</sub> (Xue et al., 2010, Junkermann, 2009 and Wagner et al., 2011). Because BAMFs of satellite observations are normally larger at high altitudes, the uncertainties of SFs from MAX-DOAS could cause an underestimation of  $AMF_{MAX-DOAS}$ , which further ~~could~~ causes an overestimation of VCD<sub>SM</sub>. ~~In~~ Since ~~addition~~ the profiles above 4km are not available from MAX-DOAS observation, they are taken ~~and derived~~ from the corresponding CTM simulations for the ~~individual~~ different satellite data sets in this study. ~~The assumption~~ This procedure can contribute to an unknown error ~~to~~ in the analysis of SF effects on satellite AMF and VCD calculations.

#### 3.4.3 Comparisons of the bimonthly mean VCD

We calculate bi-monthly averaged tropospheric VCDs for eCF < 30% for the coincident observations of the satellite instruments and MAX-DOAS (and also from the CTM simulations for the OMI products) from 2011 to 2014. The results for NO<sub>2</sub>, SO<sub>2</sub> and HCHO are shown in Fig. ~~47~~ 12. The numbers of available days for each satellite products are also shown in the bottom panels of each subfigure.

##### 1) NO<sub>2</sub>

For OMI good agreements with the MAX-DOAS VCDs are found both for the DOMINO and the improved VCDs using SFs from MAX-DOAS observations with ~~a~~ slightly better agreement ~~can be seen for~~ of the improved VCDs. GOME-2A and ~~GOME-2B~~ VCDs are systematically larger than the MAX-DOAS VCDs by about  $5 \times 10^{15}$  molecules cm<sup>-2</sup> on average. The overestimation could ~~be~~ ~~be~~ attributed ~~to~~ the errors of the NO<sub>2</sub> SFs from TM4 (Pinardi et al., 2013). Systematic differences between the GOME2-A and ~~GOME-2B~~ VCDs are found, which can be partly explained by the different swath widths of both sensors after 15 July 2013. For the same reason also better agreement between GOME-2A and MAX-DOAS VCDs is found after summer 2013. The NO<sub>2</sub> VCDs simulated by TM4 for the OMI DOMINO ~~v2~~ product are much smaller than those ~~observed by~~ from satellite and MAX-DOAS ~~observations~~. However the model data show a similar seasonality as the

observational data. The significant underestimation of the TM4 NO<sub>2</sub> VCDs could be due to many factors, most importantly the limited spatial model resolution, which is especially relevant for species with strong horizontal gradients such as NO<sub>2</sub> and SO<sub>2</sub> (see Fig. 1), but also possible errors in the emissions, transport and/or chemical mechanism. The determination of the specific contributions of the different error sources should be the subject of future studies.~~errors of limited errors of the chemical~~The determination of

## 2) SO<sub>2</sub>

For SO<sub>2</sub> large differences between the absolute values of the satellite and MAX-DOAS results are found, but all data sets show a similar seasonality with minima in summer and maxima in winter. The best agreement with MAX-DOAS results is found for the OMI BIRA VCD<sub>SM</sub>, which displays an almost identical magnitude of the SO<sub>2</sub> annual variation (while still showing a large bias). Interestingly, a much better agreement is found for the modified OMI SO<sub>2</sub> than for the OMI BIRA using the SF from the CTM. However the MAX-DOAS results are still significantly higher than the modified OMI products by about  $10 \times 10^{15}$  molecules cm<sup>-2</sup> on average. Several reasons could contribute to the differences: 1) the horizontal gradient of SO<sub>2</sub> (see Fig. 1) and the MAX-DOAS pointing direction to the North can contribute to the differences of about  $3 \times 10^{15}$  molecules cm<sup>-2</sup>. 2) The SO<sub>2</sub> cross section at 203K is applied in the current version of the OMI BIRA product. It was found that the temperature dependence of the SO<sub>2</sub> cross sections (Bogumil et al. 2003) should also be considered using e.g. a post-correction method (BIRA-IASB, 2016). The correction can increase SO<sub>2</sub> VCDs by up to  $10 \times 10^{15}$  molecules cm<sup>-2</sup> with the highest absolute changes in winter. 3) The surface albedo used in the retrieval of the OMI BIRA product is taken from the climatological monthly ~~minimum Lambertian equivalent reflector (minLER)~~ data from Kleipool et al. (2008) at 328 nm. -We expect an uncertainty of the albedo of about 0.02. This will translate to an error of 15-20% of the SO<sub>2</sub> VCDs. 4) some unknown local emissions near the station might be underestimated by the satellites, but seen by the MAX-DOAS.

The BIRA GOME-2A/B and DLR GOME-2A data are well consistent with each other, but show large differences to the corresponding MAX-DOAS results. The SO<sub>2</sub> VCDs simulated by IMAGES are systematically lower than the MAX-DOAS observations and show only a low amplitude of the seasonal variation. Same as for TM4 NO<sub>2</sub>, the discrepancy of the IMAGES SO<sub>2</sub> VCDs needs a further investigation to understand in future studies.

## 3) HCHO

Relatively good agreements between the satellite and MAX-DOAS observations of HCHO are found for all data sets (except GOME-2A before summer 2013). For OMI a better agreement is found for the modified VCDs than for the original product, with a larger improvements of the OMI VCD in summer. GOME-2A/B products are consistent with each other but strongly underestimate the HCHO VCDs, especially in summer. It is interesting to note that the CTM results have a better consistency with the MAX-DOAS results than the OMI data. The much better consistency of the IMAGES HCHO VCDs than compared to the SO<sub>2</sub> VCDs with MAX-DOAS measurements compared to the SO<sub>2</sub> VCDs is also worth a further investigation in the future. ~~In addition~~It should be noted that GOME-2A data before summer 2013 show the largest

disagreement with the MAX-DOAS data. The reason for this ~~finding-phenomenon~~ is not clear, but might be related to the different swath width in that period.

### 3.5.4 Diurnal variations characterized by combining the GOME-2A/B and OMI observations and the weekly cycle

Because of the morning and afternoon overpass time of GOME-2 and OMI, respectively, several studies (e.g. [Boersma et al., 2008](#); [Lin et al., 2010](#); [De Smedt et al., 2015](#)) investigated the differences of both data sets to characterize the diurnal variations of the TGs. The diurnal variations can be attributed to the complex interaction of the primary and secondary emission sources, depositions, atmospheric chemical reactions, and transport processes. In this section we perform a similar study, but include also MAX-DOAS data coincident to the satellite observations. We calculate the ratios between the bi-monthly mean tropospheric VCDs from GOME-2A/B and OMI ( $\text{Ratio}_{\text{Sat}}$ ) for each species and the corresponding ratios from the MAX-DOAS observations ( $\text{Ratio}_{\text{M-D}}$ ). The results are shown in Fig. ~~18~~13. The averaged  $\text{Ratio}_{\text{Sat}}$  and  $\text{Ratio}_{\text{M-D}}$  over the whole period are listed in Table 1. For  $\text{NO}_2$ , the  $\text{Ratio}_{\text{Sat}}$  for both GOME-2 instruments show good agreement. Good agreement is also found for the seasonal variation with the MAX-DOAS results, but the absolute values differ. The systematic difference of  $\text{Ratio}_{\text{Sat}}$  and  $\text{Ratio}_{\text{M-D}}$  can be attributed to the known overestimation of the GOME-2 A/B tropospheric VCD compared to the MAX-DOAS results (see Fig. 12a). This finding also indicates that using GOME-2 and OMI data can lead to wrong conclusions about the diurnal cycles of  $\text{NO}_2$ . Also for the other trace gases we investigated the ratios between the different data sets. However, because of the larger uncertainties compared to  $\text{NO}_2$ , the conclusions for  $\text{SO}_2$  and HCHO should be treated with care. For  $\text{SO}_2$ , although  $\text{Ratio}_{\text{Sat}}$  shows ~~also~~ several deviations from  $\text{Ratio}_{\text{M-D}}$ . ~~This finding also indicates that using GOME-2 and OMI data can lead to wrong conclusions about the diurnal cycles of  $\text{NO}_2$ . Also for the other trace gases we investigated the ratios between the different data sets. However, because of the larger uncertainties compared to  $\text{NO}_2$ , the conclusions for  $\text{SO}_2$  and HCHO should be treated with care. For  $\text{SO}_2$ , although  $\text{Ratio}_{\text{Sat}}$  shows also several deviations from  $\text{Ratio}_{\text{M-D}}$ .~~  $\text{Ratio}_{\text{M-D}}$  and  $\text{Ratio}_{\text{Sat}}$  are consistent on average and close to unity during a whole year indicating ~~, implying~~ similar  $\text{SO}_2$  VCDs around the overpass times of GOME-2 and OMI, ~~but  $\text{Ratio}_{\text{Sat}}$  shows also several positive and negative deviations from unity.~~ For HCHO, ~~in on general~~ average, good well good agreements between  $\text{Ratio}_{\text{Sat}}$  and  $\text{Ratio}_{\text{M-D}}$  ~~are is found~~ for GOME-2A and GOME-2B ~~is found~~ (except some outliers of  $\text{Ratio}_{\text{Sat}}$ ). Interestingly, both  $\text{Ratio}_{\text{Sat}}$  and  $\text{Ratio}_{\text{M-D}}$  are below unity indicating lower HCHO VCDs in the morning than in the afternoon.

We evaluate the weekly cycles of the VCDs of the TGs observed by satellite instruments and the corresponding MAX-DOAS. The weekly cycles are shown in the Fig. ~~S3-S17~~ in the supplement. In general only both GOME-2 instruments and corresponding MAX-DOAS measurements observed considerable weekly cycles for  $\text{NO}_2$ .

### 3.6.5 Aerosol effects on the satellite results

In this section the aerosols effects on the satellite products are investigated. The OMI products (for  $\text{SO}_2$  the OMI BIRA product is used ~~for  $\text{SO}_2$~~ ) are used for this study because of their ~~marked~~ better consistency with the MAX-DOAS results compared to the products of the other satellite instruments. In Fig. 14 the absolute (top) and relative (bottom) differences of

the TG VCDs between OMI and MAX-DOAS observations for individual OMI pixels are plotted against the AODs at 360nm derived from the MAX-DOAS observations (Wang et al. 2016). It needs to be noted that the OMI VCDs used in Fig. 14 are the modified values using the SFs derived from MAX-DOAS observations in order to isolate the aerosol effects. The left subfigures show the comparisons for the data with eCF<10%, for which a potential cloud contamination ~~are mostly excluded~~ is minimised. However the eCF filter can not exclude all clouds, and thus ~~i.e.~~ observations with thin cirrus clouds or other clouds with small geometric cloud fraction might still be included in the comparison. Therefore CTP > 900 hPa is used to further exclude residual clouds from the comparisons. ~~and~~ The comparisons for the data with eCF <10% and CTP>900 hPa are shown in the ~~right~~ center part ~~subfigures~~ of Fig. 14. Finally also observations with small TG VCDs ( $\text{NO}_2 < 2 \times 10^{16}$  molecules  $\text{cm}^{-2}$ ,  $\text{SO}_2 < 2 \times 10^{16}$  molecules  $\text{cm}^{-2}$ , and  $\text{HCHO} < 1 \times 10^{16}$  molecules  $\text{cm}^{-2}$  ~~please give individual thresholds~~) are skipped to minimise the influence of non-polluted observations on the comparison. The results after applying all three filters are shown in the right part of Fig. 14.

A systematically increasing underestimation of the OMI VCDs compared to MAX-DOAS VCDs with ~~an increase of~~ increasing AOD can be seen for ~~the three trace gases~~  $\text{NO}_2$  and  $\text{SO}_2$ . This indicates the effects of aerosols on the satellite products. ~~However~~, here one aspect needs to be considered here. Besides aerosols, also residual (low altitude) clouds might still have an effect on the comparison results. In order to quantify their potential effect, ~~clarified that some low clouds with a small eCF can not be excluded using eCF and CTP thresholds, thus they could be included in the plots of Fig. 14 and impact the analysis of aerosol effects.~~ However we performed RTM simulations ~~in~~ (for details see section 3 ~~of~~ in the supplement) to evaluate the difference of TG AMFs ~~between~~ with either aerosols ~~and~~ with residual clouds. As residual clouds we chose either homogeneous optically thin clouds covering the whole satellite pixel or optically thick clouds covering only a small geometric fraction of the satellite pixel. For both types of clouds the extinction profiles were chosen to match the ~~Note that the residual clouds and aerosols which lead to the same~~ radiance and  $\text{O}_4$  SCDs at 477nm ~~are compared~~ of the aerosol cases. We found that ~~Their~~ the differences of the TG-AMFs for aerosols and residual clouds are generally smaller than 10% for  $\text{NO}_2$  and 5% for  $\text{SO}_2$  and HCHO. It should be noted that the actual effect of residual clouds is in general much smaller, because usually aerosols and clouds are present at the same time. Thus we conclude that residual clouds have a negligible effect on the comparison results shown in Fig. 14 ~~Therefore the influence of residual clouds in Fig. 14 is not significant.~~

The dependence on AOD shown in ~~In~~ Fig. 14 is strongest for  $\text{NO}_2$ . Besides the larger uncertainties of the HCHO and  $\text{SO}_2$  retrievals, this is probably mainly related to the fact that ~~besides the larger underestimation of OMI TG VCDs for a larger AOD, a systematically larger underestimation for a larger eCF and CTP is also found for  $\text{NO}_2$ , but not for  $\text{SO}_2$  and HCHO.~~ Here it needs to be note that eCF and CTP are consistently used in the retrieval of DOMINO  $\text{NO}_2$  product in contrast to the DOMINO  $\text{NO}_2$  product, for the OMI BIRA  $\text{SO}_2$  and HCHO products no cloud correction is performed, i.e.: ~~however differently~~ a clear-sky AMF (namely in an atmosphere with for a Rayleigh scattering ~~only~~ atmosphere) is applied ~~in the retrieval of the OMI BIRA  $\text{SO}_2$  and HCHO products in cases of eCF<10%. Because currently explicit aerosol information is not considered in satellite retrievals of the official trace gas and cloud products.~~

Aerosols effects affect on the satellite TG retrievals in two ways could be different due to the two different strategies: First they affect the The aerosol effect in the first strategy is related to the fact that the OMI and GOME 2 cloud retrievals of eCF and CTP and thus the TG AMFs if a cloud correction is applied in the satellite retrievals. If a Lambertian cloud model is used the effect of this implicit aerosol correction depends systematically on the aerosol properties, are sensitive to the presence of (strong loads of) aerosols (Boersma et al., 2011; Lin et al., 2014; Wang et al., 2015; Chimot et al., 2016). The aerosol induced cloud products could contribute positive and negative effects on the satellite retrievals. The positive effect refers to that the cloud products can (partly) account for the aerosol effects on the satellite AMFs, especially for mostly scattering aerosols at high altitudes (referred to as the implicit aerosol correction can largely account) for the aerosol effect on the TG products (Boersma et al., 2011). However the impact of the implicit correction quite depends on the vertical profiles of aerosols and the TGs as well as the optical properties of aerosols (Lin et al., 2014 and Chimot et al., 2016). In some important cases (for low altitude aerosols with high AOD and small SSA) the implicit correction might even increase the errors of the AMF. For instance, Castellanos et al. (2015) demonstrated that the implicit aerosol correction can work well for the DOMINO v2 NO<sub>2</sub> product for elevated biomass burning aerosols (corresponding to effective CTP between 900 and 800hPa, around 2km) in case of modest aerosol optical depths (AODs) (<0.6) and single scattering albedo (SSA) >0.95, but break down for high AOD and lower SSA.

Besides the aerosol effect on the cloud retrievals and cloud correction schemes, aerosols also directly affect the of AMF compared to AMFs for pure Rayleigh scattering conditions, and in aerosol scenarios. It is discussed in Leitão et al. (2010) and Chimot et al., 2016 found that using RTM simulations. They found that the influence of aerosols on the satellite retrievals mainly depends on the relative vertical distributions of aerosols and TGs.

To further quantify both aerosol effects on the satellite retrievals we performed RTM simulations for typical Therefore the errors with respect to two strategies depend on the typical scenarios of aerosols and TGs in a specific location in Wuxi. In the following in order to characterize the systematic errors with respect to the two strategies in the typical anthropogenic aerosol pollution situation over Wuxi station, we characterize the typical aerosols, TGs, and aerosol induced OMI eCF and CTP. And then RTM simulations of BAMF and AMF are performed for the typical scenarios.

In Fig. 15 the OMI eCF and CTP (for eCF<10% and CTP>900hPa) are plotted against the AOD around at 360nm derived from MAX-DOAS observation in Fig. 15. (Similar plots for the AOD at 340nm derived from the nearby Taihu AERONET station (Holben et al. 1998, 2001) are also shown in the supplementary Fig. S18). The results indicate a systematic increase of eCF and CTP with an increase of increasing AOD, but also a large variability of scatter eCF, especially for AOD <1. Figure 15 also indicates CTP is close to surface pressure (similar with the typical aerosol layer height) for a large AOD and widely variable for AOD <1. The systematic increase of eCF and CTP to aerosols with AOD are consistent with the model simulations in Chimot et al., 2016. The variability of eCF and CTP can be attributed to different observation geometries as well as uncertainties of the cloud retrievals (e.g. instrument related to measurement uncertainties, and/or the variability of atmospheric and surface properties). Also the frequency distributions of eCF and CTP are

shown in Fig. 15. It demonstrates that although half data is assigned to  $eCF < 2\%$ , another half is in the  $eCF$  range of 2% to 10%. And the frequency of a large CTP is systemically high. Considering the frequencies and variability and frequency distribution of  $eCF$  and CTP,  $eCF$  of 5% and 10% as well as CTP of 900hPa and 1000hPa are used in the following for the RTM simulations to estimate the errors caused by aerosols ~~budget of implicit aerosol corrections~~. As aerosol properties we ~~The frequency of AOD shown in Fig. 15 indicates AOD is in the range of 0.2 to 1.7 with a maximum frequency in the range of 0.4 to 1.0~~ chose AOD values of 0.8 and 1.5, which represent typical and high aerosol loads at Wuxi, respectively. As vertical profile we chose an average profile derived from MAX-DOAS measurements under clear sky conditions ~~Our previous paper (Wang et al., 2016), which is shown in supplementary Fig. S19, reported that the aerosol profiles over Wuxi station representing a box like shape near the surface and an exponential decrease above 0.5 to 1 km. About 70% of aerosols accumulate in air layers below 1 km in general. The averaged profile derived from all the measurements under cloud free sky conditions are shown in supplementary Fig. S19. In order to evaluate the average and maximum aerosol effects, the averaged AOD of 0.8 and typical high AOD of 1.5 are used for the explicit aerosols in the following RTM simulations. The same profile shape with the averaged aerosol profile, which is shown in the supplementary Fig. S19, is used for the two aerosol scenarios.~~ The aerosol optical properties (single scattering albedo of 0.9, asymmetry parameter of 0.72, and Angstrom parameter of 0.85) are taken from the AERONET observations at the nearby Taihu station (Holben et al. 1998, 2001). ~~profile shapes (shape factors (SFs)) of As height profiles of w~~ We use either shape factors (SFs) derived from the Wuxi MAX-DOAS observations ~~and~~ or from the CTM simulations, which are also used for the satellite retrievals. ~~under cloud free sky conditions are used in the following simulations to calculate AMFs.~~ The SFs of the TGs are shown in the supplementary Fig. S19. The surface albedo is set to 0.1 for  $NO_2$  and 0.05 for  $SO_2$  and HCHO simulations based on the averaged value of the surface reflectivity data base derived from OMI by Kleipool et al. (2008) over Wuxi station. Temperature and pressure profiles are derived from US standard atmosphere data base. The RTM simulations are performed for five typical satellite observation geometries shown in Table 2.

~~For that purpose we focus on the OMI products (the OMI BIRA product for  $SO_2$ ) because of their marked consistency with MAX DOAS results. We selected satellite observations for  $eCF < 10\%$ , for which a potential cloud contamination is small. Moreover, especially over polluted regions like Wuxi,  $eCF$  larger than zero often indicates the effect of aerosols rather than that of clouds. In Fig. 19a 1, b 1 and c 1 the differences of the TG VCDs between OMI and MAX DOAS observations for individual OMI pixels are plotted against the aerosol optical depths (AODs) derived from the MAX DOAS observations (Wang et al. 2016). We find the increasing negative bias with increasing AOD indicating the effect of aerosols on the satellite retrievals. Moreover, for  $NO_2$  we find that the strongest negative biases are obtained for large CTP, indicating the presence of aerosols rather than of clouds. To skip measurements which are probably affected by remaining clouds, in Fig. 19 a 2, b 2 and c 2, we only show data for  $eCF < 10\%$  and  $CTP > 900$  hPa. We find that the stronger negative biases are generally related to a larger  $eCF$ , especially for  $NO_2$ . In summary we conclude that the OMI TG VCDs tends to underestimate the true TG VCDs with increasing AOD. Here it is important to note that in contrast to the DOMINO  $NO_2$  product, a clear sky AMF is applied in the retrieval of the OMI BIRA  $SO_2$  and HCHO products for  $eCF < 10\%$ . For the~~

DOMINO-NO<sub>2</sub> product, the AMF is calculated assuming a Lambertian cloud using the simultaneously derived eCF and CTP (see below). Since aerosols affect these cloud products, this correction is often referred to as ‘implicit aerosol correction’ (Chimot et al., 2016).

To further characterize the influence of applying either the clear sky AMF or the implicit aerosol correction in the following AMFs based on typical conditions of aerosols and trace gases are calculated. In a first step we characterize the typical aerosol induced eCF and CTP over the Wuxi station.

For that purpose we select six clear days with substantial aerosol pollution. We checked that the selected days were indeed cloud free based on RGB images from the MODIS instrument operated on the Aqua satellite with an overpass time eight minutes later than OMI. The MODIS images are obtained from the MODIS Rapid Response website, NASA/GSFC ([http://aeronet.gsfc.nasa.gov/cgi-bin/bamgomas\\_interactive](http://aeronet.gsfc.nasa.gov/cgi-bin/bamgomas_interactive)) (Kaufman, 2002). In addition to the MODIS images we also checked time series of the AOD from MAX DOAS and the nearby Taihu AERONET station (Holben et al. 1998, 2001). The MODIS images and time series of the AODs are shown in the Fig. S4 in the Supplement. In Table 2, the daily averaged AODs derived from MAX DOAS observations, and the eCF and CTP derived from OMI observations are shown for the six days. The aerosol induced eCF and CTP range from 4% to 9% and from 830 to 995 hPa, respectively. The averaged vertical aerosol extinction profiles for cloud free sky conditions are shown in Fig. S5.

For one five typical nadir satellite observation geometries shown in Table 2 (40° SZA, 180° RAA and 30° VZA), we The TG simulated BAMFs and AMFs were simulated for NO<sub>2</sub> at 435nm, HCHO at 337 nm and SO<sub>2</sub> at 319 nm using the RTM SCIATRAN-McArtim 2.23 (DeutschmannRozanov et al., 20052011).

The simulations are performed for four scenarios: 1) pure Rayleigh scattering conditions (BAMF<sub>clear-sky</sub> and AMF<sub>clear-sky</sub>); 2) including the typical explicit MAX DOAS aerosol profiles with the AOD of 0.8 and 1.5 (BAMF<sub>explicit</sub> and AMF<sub>explicit</sub>); 3) including Lambertian clouds at the surface (CTP of about 1000hPa) with an aerosol induced eCF of 510% and 5% and CTP of 1000hPa (near the surface) (BAMF<sub>low-cloud</sub> and AMF<sub>low-cloud</sub>); 4) including Lambertian clouds at 1km (CTP of about 900hPa) with an aerosol induced eCF of 10% and 5% and CTP of 900hPa (cloud height of about 1km) (BAMF<sub>high-cloud</sub> and AMF<sub>high-cloud</sub>). For the simulations with the explicit aerosol profiles, a SSA of 0.9, an asymmetry factor (g) (Henyey and Greenstein, 1941) of 0.72, and an Ångström factor of 0.85 are used and characterized by the measurements of the nearby Taihu AERONET station (Holben et al. 1998, 2001). The latter two cases 3 and 4 represent the implicit aerosol correction. Note that we use and the same cloud model (Lambertian surface reflector with an albedo of 0.8) as with that used as in in the official OMI cloud and TG retrievals products is used. The surface albedo is set to 0.11 for NO<sub>2</sub> and 0.05 for SO<sub>2</sub> and HCHO simulations based on the averaged value of the surface reflectivity data base derived from OMI by Kleipool et al. (2008) over Wuxi station. Temperature and pressure profiles are derived from US standard atmosphere data base.

In Fig. 20a the resulting BAMFs for the different TGs simulated for the four scenarios at the g1 observation geometry (40° SZA, 180° RAA and 30° VZA) are shown in Fig. 16a for the different TGs and aerosol and cloud assumptions are shown. Note that the results of scenario 3 and 4 with eCF of 10% are shown. And the relative differences of the BAMFs of clear sky and clouds compared to those with the explicit aerosols of AOD of either 0.8 or 1.5 are shown in Fig. 14b and c.

respectively. For all TGs, ~~similar results are obtained: compared to the clear sky the BAMFs for the explicit aerosol simulations are decrease~~ are higher close to the surface and ~~increased~~ lower for higher altitudes than the explicit aerosol BAMFs, which is caused by the additional aerosol scattering. ~~The phenomenon can be attributed to stronger upward scattering of photos by aerosols than air molecules. And~~ The BAMFs near the surface ~~For~~ for the cloud scenarios are either larger ('low cloud scenario') or smaller ('high cloud scenario') ~~are larger largely increased BAMFs are found for at all altitudes than the aerosol AMFs, especially near the surface. The phenomenon is attributed to the much larger upward reflectance of the Lambertian clouds than the real aerosols.~~ For both cloud scenarios the BAMFs are higher than the aerosol BAMFs at higher altitudes. ~~the 'high cloud scenario' the strong upward reflectance of Lambertian cloud plane strongly decreases and increases the effective light path below and above the cloud.~~ Overall the differences of the BAMFs for the cloud scenarios compared to the aerosol BAMFs ~~is~~ are larger than the differences between the clear sky BAMFs and aerosol BAMFs. ~~Compared to the BAMFs of explicit aerosols, the differences of BAMFs of two Lambertian clouds are significantly larger than the clear sky BAMFs, especially for the "low cloud scenario". And larger differences can be drawn~~ are found for the comparisons with a ~~the aerosols of high AOD of 1.5 than AOD of 0.8.~~ ~~For the 'high cloud scenario' largely increased BAMFs are found for high altitudes, whereas the BAMFs close to the surface are similar or slightly lower than the clear sky BAMFs. The differences of the BAMFs compared to the clear sky BAMFs are shown in Fig. 20 b and c.~~

Finally, ~~we calculate~~ The AMFs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO for the four scenarios are calculated using ~~for~~ the ~~four~~ corresponding ~~simulated~~ BAMFs ~~using~~ and typical SFs (shown in ~~supplementary~~ Fig. S159) derived from MAX-DOAS ~~results~~ measurements ~~or~~ and CTM simulations by Eq. 2. The relative differences of the AMFs for clear sky and for two cloud scenarios compared to the AMFs for explicit aerosol simulations for five different satellite observation geometries (listed in Table 2) are shown in Fig. 17. Fig. 17a and b shows the ~~comparisons with the explicit aerosol AMFs of~~ results for AOD of 0.8 and 1.5, respectively. It can be seen that the implicit aerosol correction can lead to large deviations, especially for the 'low cloud scenario'. The deviation for 'high cloud scenario' is close to the deviation of clear sky AMF, and even smaller in some cases, due to the compensation of the partial AMF below and above the cloud plane. Here it should be noted that for aerosol layers reaching to higher altitudes the errors of the high cloud scenario will in general increase. ~~And the deviation of~~ For the 'low cloud scenario' the deviation increases with increasing ~~is lower for~~ eCF. ~~of 5% than for 10%. The phenomenon is due to the fact that fewer Lambertian clouds are included in the AMF calculation with a decrease of eCF.~~ As already seen for the BAMFs, ~~Moreover both~~ also for the AMFs the deviations of the clear sky AMFs and Lambertian cloud AMFs ~~are larger for the comparisons with the aerosols of~~ increase with AOD ~~of 1.5 than 0.8.~~ Overall both ~~the errors of the clear sky AMF and the implicit aerosol correction increase with an increase of AOD.~~ Overall ~~the~~ ~~And the implicit aerosol correction can cause~~ the systematic biases introduced by the implicit aerosol correction (~~of AMFs of~~ 3% to 85% for NO<sub>2</sub>, -4% to 26% for HCHO, and -2% to 45% for SO<sub>2</sub>), ~~which~~ are significantly larger ~~than bias of~~ than those for the clear sky AMFs ~~of~~ (-5% to 50% for NO<sub>2</sub>, -12% to -5% for HCHO, and -9% to 1% for SO<sub>2</sub>). One important finding is that the stronger overestimation of the NO<sub>2</sub> AMF for the 'high cloud scenario' than for the 'low cloud scenario' as well as for eCF of 10%

than 5% can well explain the observed dependence of the magnitude of the underestimation of the OMI NO<sub>2</sub> VCD on the CTP and eCF as shown in Fig. 14. Therefore we conclude that for measurements with strong aerosol loads at Wuxi, the ~~in general implicit aerosol correction in general leads to can lead to~~ larger biases of TG VCDs than the use of a clear sky AMF ~~over Wuxi station, especially for CTP>900hPa and eCF>5%.~~ The derived AMFs are shown in Fig. 21. For most cases the ~~best agreement with the AMFs derived for the explicit aerosol profiles is found for the clear sky AMFs. In contrast, assuming an implicit aerosol correction can lead to large deviations, especially for the low cloud scenario. Overall similar results are found for the SF derived from MAX-DOAS or CTM.~~

~~These findings are consistent with the aerosol effects on the OMI DOMINO NO<sub>2</sub> data shown in Fig. 19. In summary we conclude that for aerosol loads like those over Wuxi the implicit aerosol correction typically causes larger bias of the satellite TG VCDs than the clear sky assumption. Thus if no explicit aerosol information is available, we recommend to apply the clear sky AMFs for eCF<10%, especially for CTP>900hPa.~~

#### 4 Conclusions

Tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub>, and HCHO derived from OMI, GOME-2A/B observations are validated using MAX-DOAS measurements in Wuxi, China from May 2011 to Dec 2014. The tropospheric VCDs and vertical profiles of aerosols and trace gases derived from the Wuxi MAX-DOAS observations using the PriAM OE-based algorithm are applied in this validation study.

~~Before the data sets are compared in a systematic way, the effects of the spatial and temporal coincidence criteria for the MAX-DOAS results and the satellite data are evaluated in detail. We find that the temporal scale over which the MAX-DOAS data are averaged has only a small effect on the comparison results. In contrast, the spatial scale over which the satellite data are averaged has a strong effect for the three species. However, a smaller effect is found for HCHO than for NO<sub>2</sub> and SO<sub>2</sub>, which is explained by the weaker horizontal gradient of the HCHO distribution. Based on our results we recommend using OMI products within distances to the MAX-DOAS station of 20km for NO<sub>2</sub> and SO<sub>2</sub>, and 50km for HCHO. For GOME 2A/B, which has a larger ground pixel size, we recommend to use data within distances of 50km for NO<sub>2</sub>, SO<sub>2</sub> and HCHO.~~

We compare the daily averaged tropospheric VCDs from the satellite products with the corresponding MAX-DOAS results under clear sky conditions (eCF<10%). For NO<sub>2</sub>: good agreement ( $R^2$  of 0.73 and systematic bias of 1%) is found for the DOMINO v2 product. For both GOME-2 products (TM4NO2A) much weaker correlation ( ~~$R^2$~~  of 0.33 for GOME-2A and 0.2 for GOME-2B) is found with the same systematic bias of about 30%. For SO<sub>2</sub>: the OMI BIRA product has a much better correlation coefficient ( $R^2$  of 0.47) than the OMI NASA product ( ~~$R^2$~~   $\approx$ 0.12), the GOME-2A BIRA product ( ~~$R^2$~~   $\approx$ 0.07), the GOME-2A DLR product ( $R^2$   $\approx$ 0.09) and the GOME-2B BIRA product ( $r^2$   $\approx$ 0.28). ~~All of these products systematically underestimate the SO<sub>2</sub> tropospheric VCDs by about 40% to 5760%.~~ For HCHO: the best agreement is found for the GOME-

2B product with  $R^2$  of 0.53 and a systematic bias of -12%. The OMI and GOME-2A products have lower  $R^2$  of 0.17 and 0.18 with the same systematic bias of about -20%, respectively.

In general, we expect that the VCDs from MAX-DOAS observations have much lower uncertainties than those from satellite observations. However we should also consider the total uncertainties of the MAX-DOAS VCDs of  $\text{NO}_2$ ,  $\text{SO}_2$  and HCHO of about 25%, 31% and 54%, respectively (Wang et al., 2016). Moreover, MAX-DOAS has low sensitivity to high altitudes,

~~normally~~ above about 1-2km. This can cause an underestimation of the VCDs retrieved from MAX-DOAS. The strength of this effect depends on the vertical distribution of the species, the atmospheric visibility, and the observation geometry of the MAX-DOAS ~~instrument~~measurement. In this study we do not discuss these issues in more detail. This should be done in further studies. Nevertheless, the sensitivity of MAX-DOAS observations to the boundary layer is much larger than ~~of~~for satellite observations. This is the altitude range in which the pollutants are usually accumulated. Thus it is reasonable to assume that the systematic differences between both data sets are mainly attributed to the errors of the satellite observations.

The cloud effects on the MAX-DOAS results and satellite products are discussed. ~~Under partial cloud coverage, the cloud effects on the MAX-DOAS results are negligible.~~ The consistency (correlations and systematic bias) of satellite data with MAX-DOAS results deteriorates with increasing eCF. The cloud effects become significant for eCF > 40% for the OMI DOMINO  $\text{NO}_2$  product, >30% for the GOME-2A\B  $\text{NO}_2$  products, > 10% for the OMI BIRA  $\text{SO}_2$  product, >20% for the OMI NASA  $\text{SO}_2$  product, >30% for the GOME-2A/B BIRA  $\text{SO}_2$  products and >30% for all HCHO products. Here it should be noted that except optically thick clouds and fog, Under partial cloud coverage, the cloud effects on the MAX-DOAS results are negligible. It should also be noted~~Note~~ that these findings are obtained ~~conclusions are~~ for the original satellite products, namely using SF from CTM or assumed fixed SF. In addition, the different thresholds of eCF could also be related to the properties of the different cloud products. This effect is not discussed in this paper, and is valuable to be further studied. In general, it should be noted that these results are representative for conditions like in Wuxi, and might be different for other locations.

In the OMI DOMINO  $\text{NO}_2$ , OMI BIRA  $\text{SO}_2$  and HCHO products, the a-priori SFs of the trace gases are obtained from CTM. We compare these SFs (derived from TM4 for  $\text{NO}_2$ , and IMAGES for  $\text{SO}_2$  and HCHO) with those derived from MAX-DOAS observation and find substantial differences. We investigate the effect of using the MAX-DOAS SFs in the satellite retrievals. Under clear sky conditions, including the SFs from MAX-DOAS changes the  $\text{SO}_2$  and HCHO AMFs by about 18% and 11%, respectively, but has almost no impact on the  $\text{NO}_2$  AMFs. We find that the modified satellite VCDs show much better agreement with the MAX-DOAS results (showing considerably higher correlation coefficients  $R^2$  and smaller systematic biases) than the original satellite data. The improvement is the strongest for periods with large trace gas VCDs, namely  $\text{NO}_2$  and  $\text{SO}_2$  in winter and HCHO in summer. In ~~the~~that period,  $\text{NO}_2$ ,  $\text{SO}_2$  and HCHO VCD change by up to 10%, 47% and 35%, respectively. We also found that the effect of using the MAX-DOAS SFs in the satellite retrievals ~~has the strongest effect~~increases for increasing eCF. This finding is mainly caused by ~~the shielding effect of clouds on the satellite observations~~the partial satellite AMF above 4km and the significant reduction of the partial satellite AMF below 4km. In addition, the low sensitivity of MAX-DOAS above about 1 – 2 km could ~~cause an underestimate~~underestimation of the

MAX-DOAS SFs of the trace gases at higher altitudes (above 1-2km), especially for SO<sub>2</sub> and HCHO. This effect could cause the underestimation of the AMFs and an overestimation of the VCDs by using the MAX-DOAS SFs. In addition the missing information on the TG profiles above 4km from MAX DOAS observations could also contribute additional errors to the analysis of SF effects.

5 We also compare the bi-monthly mean satellite products to the corresponding MAX-DOAS results. The relative seasonal variations ~~of the bi-monthly mean of the~~ NO<sub>2</sub>, SO<sub>2</sub>, HCHO tropospheric VCDs from the different satellite products agree well with the corresponding MAX-DOAS results. From the comparison of the absolute VCDs we obtain the following results:

The best consistency is found for the OMI DOMINO NO<sub>2</sub> product. A systematic overestimation of the NO<sub>2</sub> VCDs is found for GOME-2A\B NO<sub>2</sub> products. All SO<sub>2</sub> satellite products show similar SO<sub>2</sub> VCDs and a systematic underestimation of

10 about  $20 \times 10^{156}$  molecules cm<sup>-2</sup>. Based on the studies on the OMI BIRA product, the systematic underestimation could be attributed to a combined effect of errors of SFs, horizontal gradients of the SO<sub>2</sub> distribution, the temperature dependence of the SO<sub>2</sub> cross section, and uncertainties from the surface albedo and local emissions. The OMI NASA<sub>2</sub> and the GOME-2A

BIRA and NASA-DLR SO<sub>2</sub> products show a larger random variability than the OMI and GOME-2B BIRA SO<sub>2</sub> products. All OMI and GOME-2A/B products systematically underestimate the tropospheric HCHO VCDs by about  $5 \times 10^{15}$  molecules cm<sup>-2</sup>,

15 while showing a similar seasonality as the MAX-DOAS results. The biases found for the bimonthly averaged satellite TG VCDs are consistent ~~quantification of the biases of satellite TG VCDs compared to MAX DOAS results are drawn from the bimonthly comparisons and~~ with those found for ~~from~~ the daily comparisons.

We compared the diurnal variations (ratios of morning and afternoon values) of TGs by combining GOME-2A/B (morning overpass) with OMI (afternoon overpass) observations with the corresponding MAX-DOAS observations. Generally higher

20 ~~For-NO<sub>2</sub> higher values and lower HCHO values in the morning are found in the morning, while for HCHO higher values are found in the afternoon~~ are acquired. ~~And For-SO<sub>2</sub>~~ no significant diurnal cycle was found for SO<sub>2</sub>. Well ~~-consistent diurnal variations of HCHO and SO<sub>2</sub> between satellite and MAX-DOAS observations have been~~ derived. The combined satellite observations systematically overestimate the magnitude of the NO<sub>2</sub> diurnal variation compared to MAX-DOAS due

to the overestimation of the NO<sub>2</sub> VCDs by GOME-2~~For the MAX DOAS data similar results were obtained, but the NO<sub>2</sub> satellite products systematically overestimate the magnitude of NO<sub>2</sub> diurnal variation compared to the MAX DOAS data.~~ In

25 addition ~~No~~ significant weekly cycle was found for the three ~~trace-gase~~ TGs in the satellite and MAX-DOAS data.

Finally we studied ~~the~~ aerosol effects on the OMI products ~~over the~~ Wuxi station based on the MAX-DOAS observations. We find that the underestimation of the TG VCDs derived from satellite observations for mainly cloud-free observations compared to the MAX-DOAS observations systematically increases with AOD. We also investigate the aerosol effect based

30 on RTM simulations. Here it is also possible to separate the aerosol effect into two contributions: ~~of aerosol and TG profiles.~~ Aerosol effects on satellite tropospheric AMFs can be separated in to two types: a) the ~~discrepancies~~ effect of using a clear sky AMF instead of an AMF taking explicitly into account the aerosol effects, and b) ~~or~~ the effect of aerosols on the cloud retrievals, which are used in the satellite TG retrievals ~~using the~~ (implicit aerosol correction) ~~compared to using an explicit aerosol correction.~~

We find that for the measurements affected by high aerosol loads in Wuxi, in general the effect of the

implicit cloud correction on the retrieved TG VCDs is much stronger than the difference of a clear sky AMF compared to an AMF taking explicitly into account the aerosol scattering extinction. We also showed that for eCF <10% and CTP >900hPa the effect of residual clouds can be neglected if aerosol scattering extinction is explicitly taken into account. Moreover, the observed underestimation of the OMI NO<sub>2</sub> VCD for large AOD can be well explained by the error caused by the implicit aerosol correction.

The averaged effect is evaluated based on a typical scenario of aerosols and TGs in a specific location. Therefore we firstly characterized the aerosol and TG scenarios as well as the corresponding aerosol induced OMI eCF and CTP. Secondly the two types of aerosol effects are evaluated by a use of RTM simulations for different satellite observation geometries. Our results indicate that the implicit correction generally cause a larger negative bias of tropospheric VCDs than the clear sky assumption for the typical aerosol scenario in Wuxi, especially in case of CTP > 900hPa. The error of the implicit aerosol correction depends on the profiles of aerosols and TGs, observation geometries, and satellite cloud products. And the dependence of the underestimation of OMI NO<sub>2</sub> VCD on the CTP and eCF can be well explained by the error of the implicit aerosol correction. We found an increasing underestimation of OMI NO<sub>2</sub>, SO<sub>2</sub> and HCHO products with increasing AOD by up to 8%, 12% and 2%, respectively. The aerosol effects on the different satellite products are different, because different strategies for the calculation of AMFs are used: for the OMI DOMINO NO<sub>2</sub> product an implicit aerosol correction is applied based on the OMI cloud products. In contrast, for the BIRA SO<sub>2</sub> / HCHO products AMFs for clear sky are used for eCF <10%. We investigated the aerosol effect on the cloud products (eCF and CTP) on six cloud free days with pure aerosol pollutions. Aerosol induced eCF and CTP between 4% and 9% and between 830 and 995 hPa are found, respectively. Our results indicate that the implicit correction could cause a strong underestimation of tropospheric VCDs by up to about 45%, 77% and 100% for NO<sub>2</sub>, SO<sub>2</sub> and HCHO, respectively. For conditions with eCF <10% and CTP >900 hPa the AMFs based on the cloud products can lead even to larger errors than the AMFs based on the clear sky assumption. Thus Therefore -it is could be recommended reasonable to apply the clear-sky AMFs in the satellite retrievals of TG tropospheric VCDs in such cases if in case of CTP > 900hPa and eCF <10% if explicit aerosol information is not available, is not available.

**Acknowledgements:** We thank Wuxi CAS Photonics Co. Ltd for their contributions in operating the MAX-DOAS instrument. We thank the Institute of Remote Sensing / Institute of Environmental Physics, University of Bremen, Bremen, Germany for their freely accessible RTM SCIATRAN program. We thank Goddard Space Flight Center, NASA for their freely accessible archive of AERONET and MODIS data. We thank Prof. Ma Ronghua in Nanjing Institute of Geography & Limnology Chinese Academy of Sciences for his effort to operate the Taihu AERONET station. We thank the Royal Netherlands Meteorological Institute for their freely accessible archive of OMI tropospheric NO<sub>2</sub> data. This work was supported by Max Planck Society-Chinese Academy of Sciences Joint Doctoral Promotion Programme, and Monitoring and Assessment of Regional air quality in China using space Observations, Project Of Long-term sino-european co-Operation (MarcoPolo), FP7 (Grant No: 606953) and National Natural Science Foundation of China (Grant No.: 41275038).

## References

- Acarreta, J.R., De Haan, J.F. and Stammes, P.: Cloud pressure retrieval using the O<sub>2</sub> - O<sub>2</sub> absorption band at 477 nm. *Journal of Geophysical Research: Atmospheres*, 109(D5), 2004.
- 5 Barkley, M.P., Kurosu, T.P., Chance, K., De Smedt, I., Van Roozendael, M., Arneth, A., Hagberg, D. and Guenther, A.: Assessing sources of uncertainty in formaldehyde air mass factors over tropical South America: Implications for top - down isoprene emission estimates. *Journal of Geophysical Research: Atmospheres*, 117(D13), 2012.
- Bauwens, M., Stavrou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Wiedinmyer, C., Kaiser, J. W., Sindelarova, K., and Guenther, A.: Nine years of global hydrocarbon emissions based on source inversion of OMI  
10 formaldehyde observations, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-221, in review, 2016.
- Beirle, S., Platt, U., Wenig, M., and Wagner, T.: Weekly cycle of NO<sub>2</sub> by GOME measurements: a signature of anthropogenic sources, *Atmos. Chem. Phys.*, 3, 2225–2232, doi:10.5194/acp-3-2225-2003, 2003.
- Beirle, S. and Wagner, T.: Tropospheric vertical column densities of NO<sub>2</sub> from SCIAMACHY, available at:<http://www.sciamachy.org/products/NO2/NO2vcv09MPIAD.pdf>(last access: October 2012), 2010.
- 15 Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space, *Science*, 333, 1737–1739, 2011.
- BIRA-IASB: S5P/TROPOMI SO<sub>2</sub> ATBD, Issue 1.0.0, available at: [http://www.tropomi.eu/sites/default/files/files/S5P-BIRA-L2-ATBD-SO<sub>2</sub>\\_400E\\_TROPOMI\\_v1p0p0\\_20160205.pdf](http://www.tropomi.eu/sites/default/files/files/S5P-BIRA-L2-ATBD-SO2_400E_TROPOMI_v1p0p0_20160205.pdf), (last access: 26 July 2014), 2016.
- Bobrowski, N., Honninger, G., Galle, B. and Platt, U.: Detection of bromine monoxide in a volcanic plume,  
20 *Nature*, 423, 273–276, 2003.
- Boersma, K.F., H.J. Eskes and E.J. Brinksma, Error Analysis for Tropospheric NO<sub>2</sub> Retrieval from Space, *J. Geophys. Res.* 109 D04311, doi:10.1029/2003JD003962, 2004
- Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsel, E. J.: Near-real time retrieval of tropospheric NO<sub>2</sub> from OMI, *Atmos. Chem. Phys.*, 7, 2103-2118, doi:10.5194/acp-7-2103-2007, 2007.
- 25 [Boersma, K. F., Jacob, D. J., Eskes, H. J., Pinder, R. W., and Wang, J.: Intercomparison of SCIAMACHY and OMI tropospheric NO<sub>2</sub> columns: Observing the diurnal evolution of chemistry and emissions from space, \*J. Geophys. Res.\*, 113, D16S26, doi:10.1029/2007JD008816, 2008](#)
- Boersma, K.F., H.J. Eskes, R. J. Dirksen, R. J. van der A, J. P. Veefkind, P. Stammes, V. Huijnen, Q. L. Kleipool, M. Sneep,  
30 J. Claas, J. Leitao, A. Richter, Y. Zhou, and D. Brunner, An improved retrieval of tropospheric NO<sub>2</sub> columns from the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 4, 1905-1928, 2011
- Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY

- pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, *J. Photoch. Photobio. A*, 157, 167–184, 2003.
- Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Nođ, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIA-MACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, 56, 127–150, 1999.
- 5 Brinksma, E. J., Pinardi, G., Volten, H., Braak, R., Richter, A., Sch önhardt, A., van Roozendaal, M., Fayt, C., Hermans, C., Dirksen, R. J., Vlemmix, T., Berkhout, A. J. C., Swart, D. P. J., Oetjen, H., Wittrock, F., Wagner, T., Ibrahim, O. W., de Leeuw, G., Moerman, M., Curier, R. L., Celarier, E. A., Cede, A., Knap, W. H., Veeffkind, J. P., Eskes, H. J., Allaart, M., Rothe, R., Pipers, A. J. M., and Levelt, P. F.: The 2005 and 2006 DANDELIONS NO<sub>2</sub> and aerosol intercomparison campaigns, *J. Geophys. Res.*, 113, D16S46, 2008.
- 10 Burrows, J.P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weiffenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.U. and Eisinger, M.: The global ozone monitoring experiment (GOME): Mission concept and first scientific results, *Journal of the Atmospheric Sciences*, 56(2), pp.151-175, 1999
- Cahalan, R. F., W. Ridgway, W. J. Wiscombe, S. Gollmer, and Harshvardhan: Independent pixel and Monte Carlo estimates of stratocumulus albedo, *J. Atmos. Sci.*, 3776, 1994.
- 15 Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – MetOp’s Second Generation Sensor for Operational Ozone Monitoring, *ESA Bulletin*, 102, 28–36, 2000.
- Carn, S.A., Krueger, A.J., Krotkov, N.A. and Gray, M.A.: Fire at Iraqi sulfur plant emits SO<sub>2</sub> clouds detected by Earth Probe TOMS. *Geophysical research letters*, 31(19), 2004.
- Castellanos, P., Boersma, K. F., Torres, O., and de Haan, J. F.: OMI tropospheric NO<sub>2</sub> air mass factors over South America: effects of biomass burning aerosols, *Atmos. Meas. Tech.*, 8, 3831-3849, doi:10.5194/amt-8-3831-2015, 2015.
- ~~Castellanos, P., Boersma, K. F., and van der Werf, G. R.: Satellite observations indicate substantial spatiotemporal variability in biomass burning NO<sub>x</sub> emission factors for South America, *Atmos. Chem. Phys.*, 14, 3929-3943, doi:10.5194/acp-14-3929-2014, 2014.~~
- 25 Celarier, E.A., Brinksma, E.J., Gleason, J.F., Veeffkind, J.P., Cede, A., Herman, J.R., Ionov, D., Goutail, F., Pommereau, J.P., Lambert, J.C. and Van Roozendaal, M.: Validation of Ozone Monitoring Instrument nitrogen dioxide columns. *Journal of Geophysical Research: Atmospheres*, 113(D15), 2008
- Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., and Jacob D. J.: Satellite observations of formaldehyde over North America from GOME, *Geophys. Res. Lett.*, 27, 3461–3464, 2000.
- Chan, K. L., Hartl, A., Lam, Y. F., Xie, P. H., Liu, W. Q., Cheung, H. M., Lampel, J., Pöhler, D., Li, A., Xu, J., Zhou, H. J., Ning, Z. Wenig, M. O.: Observations of tropospheric NO<sub>2</sub> using ground based MAX-DOAS and OMI measurements during the Shanghai World Expo 2010, *Atmospheric Environment*, 119, 45-58, 2015
- 30 Chen, D., Zhou, B., Beirle, S., Chen, L. M., and Wagner, T.: Tropospheric NO<sub>2</sub> column densities deduced from zenith-sky DOAS measurements in Shanghai, China, and their application to satellite validation, *Atmos. Chem. Phys.*, 9, 3641-3662, 2009.

- Chimot, J., Vlemmix, T., Veefkind, J. P., de Haan, J. F., and Levelt, P. F.: Impact of aerosols on the OMI tropospheric NO<sub>2</sub> retrievals over industrialized regions: how accurate is the aerosol correction of cloud-free scenes via a simple cloud model?, *Atmos. Meas. Tech.*, 9, 359-382, doi:10.5194/amt-9-359-2016, 2016.
- Clemer, K., Van Roozendaal, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Sperr, R., Wang, P., and De Maziere, M.: Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, *Atmos. Meas. Tech.*, 3, 863–878, 2010.
- Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., El Haddad, I., and Prevot, A. S.: High secondary aerosol contribution to particulate pollution during haze events in China, *Nature*, 514, 218–222, 2014.
- De Smedt, I., Müller, J.-F., Stavrou, T., van der A, R., Eskes, H., and Van Roozendaal, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, *Atmos. Chem. Phys.*, 8, 4947–4963, doi:10.5194/acp-8-4947-2008, 2008.
- De Smedt, I., Van Roozendaal, M., Stavrou, 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, doi:10.5194/amt-5-2933-2012, 2012.
- De Smedt, I., Stavrou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendaal, M.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, *Atmos. Chem. Phys.*, 15, 12519-12545, doi:10.5194/acp-15-12519-2015, 2015.
- Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., Platt, U., Prados-Román, C., Pukl, J., Wagner, T. and Werner, B.: The Monte Carlo atmospheric radiative transfer model McArtim: Introduction and validation of Jacobians and 3D features, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(6), 1119-1137, 2011.
- Dirksen, R. J., Boersma, K. F., Eskes, H. J., Ionov, D. V., Bucsela, E. J., Levelt, P. F., and Kelder, H. M.: Evaluation of stratospheric NO<sub>2</sub> retrieved from the Ozone Monitoring Instrument: intercomparison, diurnal cycle and trending, *J. Geophys. Res.*, 116, D08305, doi:10.1029/2010JD014943, 2011.
- Ding, J., van der A, R. J., Mijling, B., Levelt, P. F., and Hao, N.: NO<sub>x</sub> emission estimates during the 2014 Youth Olympic Games in Nanjing, *Atmos. Chem. Phys.*, 15, 9399-9412, doi:10.5194/acp-15-9399-2015, 2015.
- Eisinger, M. and Burrows, J.P.: Tropospheric sulfur dioxide observed by the ERS-2 GOME instrument. *Geophysical Research Letters*, 25(22), pp.4177-4180, 1998.
- Environmental Protection Agency: National air quality and emissions trends report 1998, Rep. EPA 454/R-00-003, 1998.
- Eskes, H. J., and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, *Atmos. Chem. Phys.*, 3, 1285-1291, 10.5194/acp-3-1285-2003, 2003.

- Frieß U., Monks, P. S., Remedios, J. J., Rozanov A., Sinreich R., Wagner T., and Platt, U.: MAX-DOAS O4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies, *J. Geophys. Res.*, 111, D14203, 2006, doi:10.1029/2005JD006618.
- Frieß U., Sihler H, Sander R. The vertical distribution of BrO and aerosols in the Arctic: Measurements by active and passive differential optical absorption spectroscopy, *J. Geophys. Res.*, 116, D00R04, 2011.
- Frieß U., Klein Baltink, H., Beirle, S., Clàmer, K., Hendrick, F., Henzing, B., Irie, H., de Leeuw, G., Li, A., Moerman, M. M., van Roozendaal, M., Shaiganfar, R., Wagner, T., Wang, Y., Xie, P., Yilmaz, S., and Zieger, P.: Intercomparison of aerosol extinction profiles retrieved from MAX-DOAS measurements, *Atmos. Meas. Tech. Discuss.*, doi:10.5194/amt-2015-358, 2016.
- 10 Fu, G. Q., Xu, W. Y., Yang, R. F., Li, J. B., and Zhao, C. S.: The distribution and trends of fog and haze in the North China Plain over the past 30 years, *Atmos. Chem. Phys.*, 14, 11949-11958, doi:10.5194/acp-14-11949-2014, 2014.
- González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, *Atmos. Meas. Tech.*, 8, 19-32, doi:10.5194/amt-8-19-2015, 2015.
- 15 Gonzi, S., Palmer, P.I., Barkley, M.P., De Smedt, I. and Van Roozendaal, M.: Biomass burning emission estimates inferred from satellite column measurements of HCHO: Sensitivity to co - emitted aerosol and injection height. *Geophysical Research Letters*, 38(14), 2011.
- Hains, J. C., Boersma, K. F., Kroon, M., Dirksen, R. J., Cohen, R. C., Perring, A. E., Bucsela, E., Volten, H., Swart, D. P. J., Richter, A., Wittrock, F., Shoenhardt, A., Wagner, T., Ibrahim, O. W., Van Roozendaal, M., Pinardi, G., Gleason, J. F., Veefkind, J. P., and Levelt, P.: Testing and improving OMI DOMINO tropospheric NO<sub>2</sub> using observations from the DANDELIONS and INTEX-B validation campaigns, *J. Geophys. Res.*, 115, D05301, doi:10.1029/2009JD012399, 2010.
- 20
- Han, K. M., Lee, S., Chang, L. S., and Song, C. H.: A comparison study between CMAQ-simulated and OMI-retrieved NO<sub>2</sub> columns over East Asia for evaluation of NO<sub>x</sub> emission fluxes of INTEX-B, CAPSS, and REAS inventories, *Atmos. Chem. Phys.*, 15, 1913-1938, doi:10.5194/acp-15-1913-2015, 2015.
- 25 Hartl, A. and Wenig, M. O.: Regularisation model study for the least-squares retrieval of aerosol extinction time series from UV/VIS MAX-DOAS observations for a ground layer profile parameterisation, *Atmos. Meas. Tech.*, 6, 1959-1980, doi:10.5194/amt-6-1959-2013, 2013.
- Hassinen, S., Balis, D., Bauer, H., Begoin, M., Delcloo, A., Eleftheratos, K., Gimeno Garcia, S., Granville, J., Grossi, M., Hao, N., Hedelt, P., Hendrick, F., Hess, M., Heue, K.-P., Hovila, J., Jønh-Sørensen, H., Kalakoski, N., Kauppi, A., Kiemle, S., Kins, L., Koukouli, M. E., Kujanpää J., Lambert, J.-C., Lang, R., Lerot, C., Loyola, D., Pedergnana, M., Pinardi, G., Romahn, F., van Roozendaal, M., Lutz, R., De Smedt, I., Stammes, P., Steinbrecht, W., Tamminen, J., Theys, N., Tilstra, L. G., Tuinder, O. N. E., Valks, P., Zerefos, C., Zimmer, W., and Zyrichidou, I.: Overview of the O3M SAF
- 30

- GOME-2 operational atmospheric composition and UV radiation data products and data availability, *Atmos. Meas. Tech.*, 9, 383-407, doi:10.5194/amt-9-383-2016, 2016.
- Heckel, A., Kim, S.-W., Frost, G. J., Richter, A., Trainer, M., and Burrows, J. P.: Influence of low spatial resolution a priori data on tropospheric NO<sub>2</sub> satellite retrievals, *Atmos. Meas. Tech.*, 4, 1805–1820, doi:10.5194/amt-4-1805-2011, 2011.
- 5 Herron-Thorpe, F. L., Lamb, B. K., Mount, G. H., and Vaughan, J. K.: Evaluation of a regional air quality forecast model for tropospheric NO<sub>2</sub> columns using the OMI/Aura satellite tropospheric NO<sub>2</sub> product, *Atmos. Chem. Phys.*, 10, 8839-8854, doi:10.5194/acp-10-8839-2010, 2010.
- 10 Hewson, W., Bösch, H., Barkley, M. P., and De Smedt, I.: Characterisation of GOME-2 formaldehyde retrieval sensitivity, *Atmos. Meas. Tech.*, 6, 371–386, doi:10.5194/amt-6-371-2013, 2013.
- Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y., Nakajima, T., Lavenue, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, 66, 1–16, 1998.
- 15 Holben, B. N., Tanre, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J., Chatenet, B., Lavenue, F., Kaufman, Y. J., Van de Castle, J., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karnieli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, Z.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, *J. Geophys. Res.*, 106, 12067–12097, 2001
- Hönninger, G. and Platt, U.: Observations of BrO and its vertical distribution during surface ozone depletion at Alert, *Atmos. Environ.*, 36, 2481–2489, 2002.
- 20 Hönninger, G., Friedeburg, C. von and Platt, U., Multi axis differential optical absorption spectroscopy (MAX-DOAS), *Atmos. Chem. Phys.*, 4, 231–254, 2004
- Huang, R.-J., Zhang, Y., Bozzetti, C., Ho, K.-F., Cao, J., Han, Y., Dällenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., Zotter, P., Wolf, R., Pieber, S. M., Bruns, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., and Prévôt, A. S. H.: High secondary aerosol contribution to particulate pollution during haze events in China, *Nature*, 514, 218–222, 2014.
- 25 Irie, H., Kanaya, Y., Akimoto, H., Iwabuchi, H., Shimizu, A., and Aoki, K.: First retrieval of tropospheric aerosol profiles using MAX-DOAS and comparison with lidar and sky radiometer measurements, *Atmos. Chem. Phys.*, 8, 341–350, 2008.
- Irie, H., Takashima, H., Kanaya, Y., Boersma, K. F., Gast, L., Wittrock, F., Brunner, D., Zhou, Y., and Van Roozendael, M.: Eight-component retrievals from ground-based MAX-DOAS observations, *Atmos. Meas. Tech.*, 4, 1027–1044, 2011.
- 30 Irie, H., Boersma, K. F., Kanaya, Y., Takashima, H., Pan, X., and Wang, Z. F.: Quantitative bias estimates for tropospheric NO<sub>2</sub> columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia, *Atmos. Meas. Tech.*, 5, 2403-2411, doi:10.5194/amt-5-2403-2012, 2012.
- Irie, H., Muto, T., Itahashi, S., Kurokawa, J.I. and Uno, I.: Turnaround of Tropospheric Nitrogen Dioxide Pollution Trends in China, Japan, and South Korea, *SOLA*, 12(0), 170-174, doi:10.2151/sola.2016-035, 2016.

- Jiang, C., Wang, H., Zhao, T., Li, T., and Che, H.: Modeling study of PM<sub>2.5</sub> pollutant transport across cities in China's Jing–Jin–Ji region during a severe haze episode in December 2013, *Atmos. Chem. Phys.*, 15, 5803-5814, doi:10.5194/acp-15-5803-2015, 2015.
- Jin, J., Ma, J., Lin, W., Zhao, H., Shaiganfar, R., Beirle, S. and Wagner, T.: MAX-DOAS measurements and satellite validation of tropospheric NO<sub>2</sub> and SO<sub>2</sub> vertical column densities at a rural site of North China. *Atmospheric Environment*, 133, pp.12-25, 2016.
- Junkermann, W.: On the distribution of formaldehyde in the western Po-Valley, Italy, during FORMAT 2002/2003, *Atmos. Chem. Phys.*, 9, 9187-9196, doi:10.5194/acp-9-9187-2009, 2009.
- Kanaya, Y., Irie, H., Takashima, H., Iwabuchi, H., Akimoto, H., Sudo, K., Gu, M., Chong, J., Kim, Y. J., Lee, H., Li, A., Si, F., Xu, J., Xie, P.-H., Liu, W.-Q., Dzhola, A., Postolyakov, O., Ivanov, V., Grechko, E., Terpigova, S., and Panchenko, M.: Long-term MAX-DOAS network observations of NO<sub>2</sub> in Russia and Asia (MADRAS) during the period 2007–2012: instrumentation, elucidation of climatology, and comparisons with OMI satellite observations and global model simulations, *Atmos. Chem. Phys.*, 14, 7909-7927, doi:10.5194/acp-14-7909-2014, 2014.
- Kaufman, Y. J., Tanré D. and Boucher, O.: A satellite view of aerosols in the climate system. *Nature*, 419, 215-223, 2002
- Khokhar, M.F., Frankenberg, C., Van Roozendaal, M., Beirle, S., Köhl, S., Richter, A., Platt, U. and Wagner, T., Satellite observations of atmospheric SO<sub>2</sub> from volcanic eruptions during the time-period of 1996–2002. *Advances in Space Research*, 36(5), pp.879-887, 2005.
- Kleipool, Q. L., Dobber, M.R., de Haan, J. and Levelt, P.F.: Earth surface reflectance climatology from 3 years of OMI data. *Journal of Geophysical Research: Atmospheres*, 113(D18), 2008.
- Koukouli, M.E., Balis, D.S., Theys, N., Hedelt, P., Richter, A., Krotkov, N., Li, C. and Taylor, M.: Anthropogenic sulphur dioxide load over China as observed from different satellite sensors. *Atmospheric Environment*, 145, 45-59, 2016.
- Krotkov, N.A., Carn, S.A., Krueger, A.J., Bhartia, P.K. and Yang, K.: Band residual difference algorithm for retrieval of SO<sub>2</sub> from the Aura Ozone Monitoring Instrument (OMI). *Geoscience and Remote Sensing, IEEE Transactions on*, 44(5), pp.1259-1266, 2006.
- Krotkov, N.A., McClure, B., Dickerson, R.R., Carn, S.A., Li, C., Bhartia, P.K., Yang, K., Krueger, A.J., Li, Z., Levelt, P.F. and Chen, H.: Validation of SO<sub>2</sub> retrievals from the Ozone Monitoring Instrument over NE China. *Journal of Geophysical Research: Atmospheres*, 113(D16), 2008
- Krueger, A.J., Walter, L.S., Bhartia, P.K., Schnetzler, C.C., Krotkov, N.A., Sprod, I.T. and Bluth, G.J.S.: Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments. *Journal of Geophysical Research: Atmospheres*, 100(D7), pp.14057-14076, 1995.
- Kurosu, T. P.: OMHCHO README FILE, available at: [https://www.cfa.harvard.edu/atmosphere/Instruments/OMI/PGEReleases/READMEs/OMHCHO\\_README.pdf](https://www.cfa.harvard.edu/atmosphere/Instruments/OMI/PGEReleases/READMEs/OMHCHO_README.pdf) (last access: 22 April 2016), 2008.

- Lee, C., Martin, R.V., van Donkelaar, A., O'Byrne, G., Krotkov, N., Richter, A., Huey, L.G. and Holloway, J.S.: Retrieval of vertical columns of sulfur dioxide from SCIAMACHY and OMI: Air mass factor algorithm development, validation, and error analysis. *Journal of Geophysical Research: Atmospheres*, 114(D22), 2009
- Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q. J., Beekmann, M., and Burrows, J. P.: On the improvement of NO<sub>2</sub> satellite retrievals – aerosol impact on the air mass factors, *Atmos. Meas. Tech.*, 3, 475-493, doi:10.5194/amt-3-475-2010, 2010.
- Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., van den Oord, G. H. J., Bhartia, P. K., Tamminen, J., de Haan, J. F., and Veefkind, J. P.: Science objectives of the Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote Sens.*, 44, 1199–1208, 2006a.
- 10 Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote Sens.*, 44, 1093–1101, 2006b.
- Li X., Brauers, T. and Hofzumahaus, A.: MAX-DOAS measurements of NO<sub>2</sub>, HCHO and CHOCHO at a rural site in Southern China. *Atmos. Chem. Phys. Discuss.*, 12, 3983-4029, 2012.
- Li X., Brauers, T., Shao, M., Garland, R. M., Wagner, T., Deutschmann, T., and Wahner, A.: MAX-DOAS measurements in 15 southern China: retrieval of aerosol extinctions and validation using ground-based in-situ data. *Atmos. Chem. Phys.*, 10, 2079–2089, 2010.
- Li, C., Joiner, J., Krotkov, N.A. and Bhartia, P.K.: A fast and sensitive new satellite SO<sub>2</sub> retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument. *Geophysical Research Letters*, 40(23), 6314-6318, 2013.
- 20 Lin, J.-T., McElroy, M. B., and Boersma, K. F.: Constraint of anthropogenic NO<sub>x</sub> emissions in China from different sectors: a new methodology using multiple satellite retrievals, *Atmos. Chem. Phys.*, 10, 63–78, doi:10.5194/acp-10-63-2010, 2010.
- Lin, J.-T., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Van Roozendaal, M., Clémer, K., and Irie, H.: Retrieving tropospheric nitrogen dioxide from the Ozone Monitoring Instrument: effects of aerosols, surface reflectance anisotropy, and vertical profile of nitrogen dioxide, *Atmos. Chem. Phys.*, 14, 1441-1461, doi:10.5194/acp-14-1441-2014, 2014.
- 25 Lin, J.-T., Liu, M.-Y., Xin, J.-Y., Boersma, K. F., Spurr, R., Martin, R., and Zhang, Q.: Influence of aerosols and surface reflectance on satellite NO<sub>2</sub> retrieval: seasonal and spatial characteristics and implications for NO<sub>x</sub> emission constraints, *Atmos. Chem. Phys.*, 15, 11217-11241, doi:10.5194/acp-15-11217-2015, 2015.
- ~~Lin, J. T., van Donkelaar, A., Xin, J., Che, H., and Wang, Y.: Clear sky aerosol optical depth over East China estimated from visibility measurements and chemical transport modeling, *Atmospheric Environment*, 95, 257-267, doi:10.1016/j.atmosenv.2014.06.044, 2014~~
- 30 Lohmann, U., and Feichter, J.: Global indirect aerosol effects: a review, *Atmos. Chem. Phys.*, 5, 715-737, 10.5194/acp-5-715-2005, 2005.

- Loyola, D., W. Thomas, Y. Livschitz, T. Ruppert, P. Albert, and R. Hollmann: Cloud properties derived from GOME/ERS-2 backscatter data for trace gas retrieval, *IEEE Trans. Geosci. Remote Sens.*, 45(9), 2747–2758, doi:10.1109/TGRS.2007.901043, 2007.
- Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.: Tropospheric NO<sub>2</sub> vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008–2011) and satellite validation, *Atmos. Chem. Phys.*, 13, 1547–1567, doi:10.5194/acp-13-1547-2013, 2013.
- Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO<sub>2</sub> columns, *J. Geophys. Res.*, 108, 4537, doi:10.1029/2003jd003453, 2003.
- McCormic, R. A. and Ludwig, J. H.: Climate modifications by atmospheric aerosols, *Science*, 156, 1358–1359, 1967.
- Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K. V., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, *J. Geophys. Res.*, 113, 1–18, doi:10.1029/2007JD008950, 2008.
- Müller, J.F. and Brasseur, G.: IMAGES: A three-dimensional chemical transport model of the global troposphere. *Journal of Geophysical Research: Atmospheres*, 100(D8), pp.16445–16490, 1995.
- Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Perez Albinana, A.: GOME-2 on MetOp: From in-orbit verification to routine operations, in: *Proceedings of EUMETSAT Meteorological Satellite Conference*, Helsinki, Finland, 12–16 June 2006.
- Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, *Atmos. Meas. Tech.*, 9, 1279–1301, doi:10.5194/amt-9-1279-2016, 2016.
- Nowlan, C.R., Liu, X., Chance, K., Cai, Z., Kurosu, T.P., Lee, C. and Martin, R.V., 2011. Retrievals of sulfur dioxide from the Global Ozone Monitoring Experiment 2 (GOME-2) using an optimal estimation approach: Algorithm and initial validation. *Journal of Geophysical Research: Atmospheres*, 116(D18), 2011
- Ozone Monitoring Instrument (OMI), *IEEE Trans. Geosci. Remote*
- Palmer, P. I., Jacob, D. J., Chance, K. V., Martin, R. V. D. R. J., Kurosu, T. P., Bey, I., Yantosca, R. M., and Fiore, A. M.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, *J. Geophys. Res.*, 106, 14539–14550, doi:10.1029/2000JD900772, 2001.
- Pinardi, G., Yu, H., Hendrick, F., Tack, F., Granville, J., Lambert, J.-C. and Van Roozendaal, M.: End-to-end Validation of Total and Tropospheric NO<sub>2</sub> Columns from Atmospheric Composition Satellite Sensors, *Atmospheric Composition Validation and Evolution*, ESRIN, Italy, March 2013.
- Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy*. Springer-Verlag Heidelberg, Berlin, 229–375, 2008.

- Richter, A., Burrows, J.P., Nüß H., Granier, C. and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space. *Nature*, 437(7055), pp.129-132, 2005
- Richter, A., Wittrock, F. and Burrows, J.P.: SO<sub>2</sub> measurements with SCIAMACHY. In *Proc. Atmospheric Science Conference* (pp. 8-12). May, 2006
- 5 Richter, A., Wittrock, F., Schönhardt, A. and Burrows, J.P.: Quantifying volcanic SO<sub>2</sub> emissions using GOME2 measurements. *Geophys. Res. Abstr.*, EGU2009-7679, EGU General Assembly, 2009.
- Rix, M., Valks, P., Hao, N., Loyola, D. G., Schlager, H., Huntrieser, H. H., Flemming, J., Koehler, U., Schumann, U., and Inness, A.: Volcanic SO<sub>2</sub>, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010, *J. Geophys. Res.*, 117, D00U19, doi:10.1029/2011JD016718, 2012.
- 10 Rozanov, A., Rozanov, V., Buchwitz, M., Kokhanovsky, A., and Burrows, J.: SCIATRAN 2.0 - A new radiative transfer model for geophysical applications in the 175-2400 nm spectral region, in: *Atmospheric Remote Sensing: Earth's Surface, Troposphere, Stratosphere and Mesosphere - I*, edited by Burrows, J. and Eichmann, K., vol. 36 of *ADVANCES IN SPACE RESEARCH*, 1015-1019, 2005.
- Seinfeld, J. H. and Pandis, S. N.: *Atmospheric Chemistry and Physics—From Air Pollution to Climate Change*. John Wiley, New York, 1998.
- 15 Sens., 44, 1259–1266.
- Shaiganfar, R., Beirle, S., Sharma, M., Chauhan, A., Singh, R. P., and Wagner, T.: Estimation of Nox emissions from Delhi using Car MAX-DOAS observations and comparison with OMI satellite data, *Atmos. Chem. Phys.*, 11, 10871–10887, doi:10.5194/acp-11-10871-2011, 2011.
- 20 Spurr, R.J.D., Kurosu, T.P. and Chance, K.V.: A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 68(6), pp.689-735, 2001
- Spurr, R.: LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems, in *Light Scattering Reviews 3*, edited by A. Kokhanovsky, pp. 229 – 275, Springer, Berlin, doi:10.1007/978-3-540-48546-9\_7, 2008.
- 25 Stammes, P.: Errors in UV reflectivity and albedo calculations due to neglecting polarisation, *SPIE* 2311, 227-235, 1994.
- Stammes, P., Sneep, M., de Haan, J. F., Veefkind, J. P., Wang, P. and Levelt, P. F.: Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation, *J. Geo-phys. Res.*, 113, D16S38, doi:10.1029/2007JD008820, 2008.
- Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T., and Zhang, Q.: Key chemical NOx sink uncertainties and how they influence top-down emissions of nitrogen oxides, *Atmos. Chem. Phys.*, 13, 9057-9082, doi:10.5194/acp-13-9057-2013, 2013.
- 30 Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P.-F., and Guenther, A.: How consistent are top-down hydrocarbon emissions based on

- formaldehyde observations from GOME-2 and OMI?, *Atmos. Chem. Phys.*, 15, 11861-11884, doi:10.5194/acp-15-11861-2015, 2015.
- 5 Theys, N., De Smedt, I., Gent, J., Danckaert, T., Wang, T., Hendrick, F., Stavrakou, T., Bauduin, S., Clarisse, L., Li, C., Krotkov, N., Yu, H., Brenot, H., and Van Roozendael, M.: Sulfur dioxide vertical column DOAS retrievals from the Ozone Monitoring Instrument: Global observations and comparison to ground - based and satellite data. *Journal of Geophysical Research: Atmospheres*, 120(6), 2470-2491, 2015.
- Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO<sub>2</sub> column retrieval for GOME-2, *Atmos. Meas. Tech.*, 4, 1491–1514,doi:10.5194/amt-4-1491-2011, 2011.
- 10 van der A, R. J., Eskes, H. J., Boersma, K. F., Van Noije, T. P. C., Van Roozendael, M., De Smedt, I., Peters, D.H.M.U. and Meijer, E.W., : Trends, seasonal variability and dominant NO<sub>x</sub> source derived from a ten year record of NO<sub>2</sub> measured from space. *Journal of Geophysical Research: Atmospheres*, 113(D4), 2008
- Van Roozendael, M., Fayt, C., Post, P., Hermans, C., Lambert, J.-C.: Retrieval of BrO and NO<sub>2</sub> from UV-Visible Observations, in: *Sounding the troposphere from space: a new Era for Atmospheric Chemistry. The TROPOSAT Final Report*. Peter Borrell, Patricia M. Borrell, John P. Burrows and Ulrich Platt (editors), Springer Verlag, 155-166, 2003
- 15 Veefkind, J.P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H.J., De Haan, J.F., Kleipool, Q. and Van Weele, M.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120, pp.70-83, 2012.
- 20 Vlemmix, T., Piders, A. J. M., Stammes, P., Wang, P., and Levelt, P. F.: Retrieval of tropospheric NO<sub>2</sub> using the MAX-DOAS method combined with relative intensity measurements for aerosol correction, *Atmos. Meas. Tech.*, 3, 1287–1305, 2010.
- Vlemmix, T., Piders, A. J. M., Berkhout, A. J. C., Gast, L. F. L., Wang, P., and Levelt, P. F.: Ability of the MAX-DOAS method to derive profile information for NO<sub>2</sub>: can the boundary layer and free troposphere be separated?, *Atmos. Meas. Tech.*, 4, 2659–2684, 2011.
- 25 Vlemmix, T., Eskes, H. J., Piders, A. J. M., Schaap, M., Sauter, F. J., Kelder, H., & Levelt, P. F.: MAX-DOAS tropospheric nitrogen dioxide column measurements compared with the Lotos-Euros air quality model, *Atmospheric Chemistry and Physics*, 15(3), 1313-1330, 2015a
- Vlemmix, T., Hendrick, F., Pinardi, G., De Smedt, I., Fayt, C., Hermans, C., Piders, A., Wang, P., Levelt, P., and Van Roozendael, M.: MAX-DOAS observations of aerosols, formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches, *Atmos. Meas. Tech.*, 8, 941-963, doi:10.5194/amt-8-941-2015, 2015b.
- 30 Wagner, T., B. Dix, C. von Friedeburg, Frieß U., Sanghavi, S., Sinreich, R. and Platt, U.: MAX-DOAS O<sub>4</sub> measurements: A new technique to derive information on atmospheric aerosols—Principles and information content, *J. Geophys. Res.*, 109, D22205, 2004.

- Wagner, T., Beirle, S., Brauers, T., Deutschmann, T., Frieß, U., Hak, C., Halla, J. D., Heue, K. P., Junkermann, W., Li, X., Platt, U. and Pundt-Gruber, I.: Inversion of tropospheric profiles of aerosol extinction and HCHO and NO<sub>2</sub> mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets, *Atmos. Meas. Tech. Discuss.*, 4, 3891–3964, 2011.
- 5 Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J., and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, *Atmos. Meas. Tech.*, 7, 1289–1320, doi:10.5194/amt-7-1289-2014, 2014.
- Wang, P., Stammes, P., van der A, R., Pinardi, G. and van Roozendael, M.: FRESCO+: an improved O<sub>2</sub> A-band cloud retrieval algorithm for tropospheric trace gas retrievals, *Atmos. Chem. Phys.*, 8, 6565–6576, 2008.
- 10 Wang, Y., Li, A., Xie, P. H., Chen, H., Xu, J., Wu, F. C., G., L. J., and Q., L. W.: Retrieving vertical profile of aerosol extinction by multi-axis differential optical absorption spectroscopy, *Acta Phys. Sin.*, 16, doi: 10.7498/aps.62.180705, <http://wulixb.iphy.ac.cn/EN/abstract/abstract54703.shtml#>, 2013a.
- Wang, Y., Li, A., Xie, P. H., Chen, H., Mou, F. S., Xu, J., Wu, F. C., Zeng, Y., Liu, J. G. and Liu, W. Q.: Measuring tropospheric vertical distribution and vertical column density of NO<sub>2</sub> by multi-axis differential optical absorption spectroscopy, *Acta Phys. Sin.*, 16, DOI:10.7498/aps.62.200705, <http://wulixb.iphy.ac.cn/EN/abstract/abstract56201.shtml>, 2013b.
- 15 Wang, Y., Penning de Vries, M., Xie, P. H., Beirle, S., Dörner, S., Remmers, J., Li, A., and Wagner, T.: Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, *Atmos. Meas. Tech.*, 8, 5133-5156, doi:10.5194/amt-8-5133-2015, 2015.
- 20 Wang, Y., Lampel, J., Xie, P., Beirle, S., Li, A., Wu, D., and Wagner, T.: Ground-based MAX-DOAS observations of tropospheric aerosols, NO<sub>2</sub>, SO<sub>2</sub> and HCHO in Wuxi, China, from 2011 to 2014, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-282, in review, 2016.
- Williams, J. E., Scheele, M. P., van Velthoven, P. F. J., Cammas, J.-P., Thouret, V., Galy-Lacaux, C., and Volz-Thomas, A.: The influence of biogenic emissions from Africa on tropical tropospheric ozone during 2006: a global modelling study, *Atmos. Chem. Phys.*, 9, 5729–5749, doi:10.5194/acp-9-5729-2009, 2009
- 25 Wittrock, F., Oetjen, H., Richter, A., Fietkau, S., Medeke, T., Rozanov, A., and Burrows, J. P.: MAX-DOAS measurements of atmospheric trace gases in Ny-Alesund – Radiative transfer studies and their application, *Atmos. Chem. Phys.*, 4, 955–966, 2004.
- Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, *Geophys. Res. Lett.*, 33, L16804, doi:10.1029/2006GL026310, 2006a.
- 30 Wittrock, F.: The Retrieval of Oxygenated Volatile Organic Compounds by Remote Sensing Techniques. Ph.D., University of Bremen, Bremen, Germany, 2006b, Available at: [http://www.doas-bremen.de/paper/diss\\_wittrock\\_06.pdf](http://www.doas-bremen.de/paper/diss_wittrock_06.pdf) (last accessed Dec 2015).

- Xue, L., Ding, A., Gao, J., Wang, T., Wang, W., Wang, X., Lei, H., Jin, D., and Qi, Y.: Aircraft measurements of the vertical distribution of sulfur dioxide and aerosol scattering coefficient in China. *Atmospheric Environment*, 44(2), 278-282, 2010.
- Yang, K., Krotkov, N.A., Krueger, A.J., Carn, S.A., Bhartia, P.K. and Levelt, P.F.: Retrieval of large volcanic SO<sub>2</sub> columns from the Aura Ozone Monitoring Instrument: Comparison and limitations. *Journal of Geophysical Research: Atmospheres*, 112(D24), 2007
- 5 Yilmaz, S.: Retrieval of Atmospheric Aerosol and Trace Gas Vertical Profiles using Multi-Axis Differential Optical Absorption Spectroscopy, Dissertation submitted to the Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany for the degree of Doctor of Natural Sciences, 2012.

# Figures & Tables

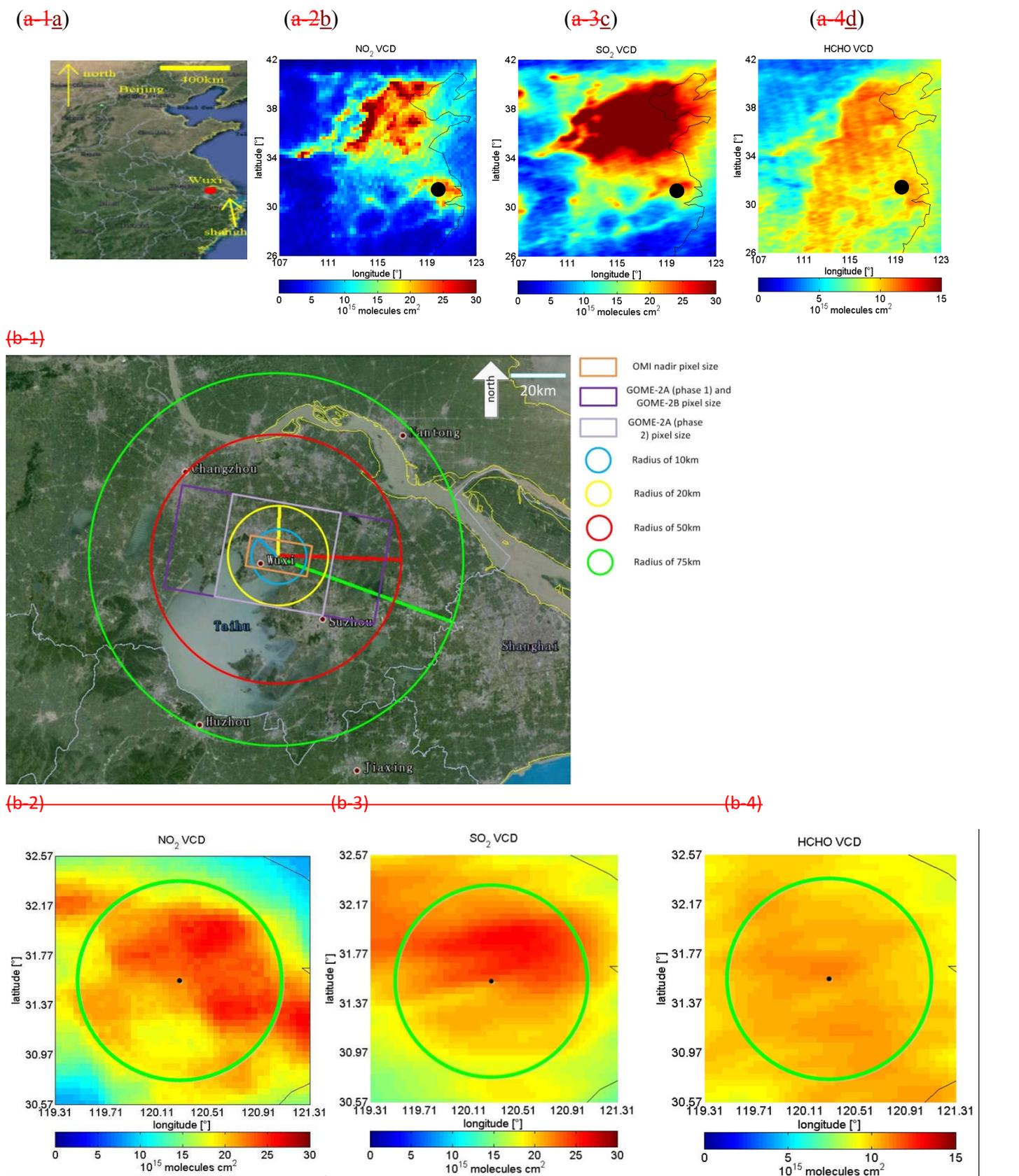
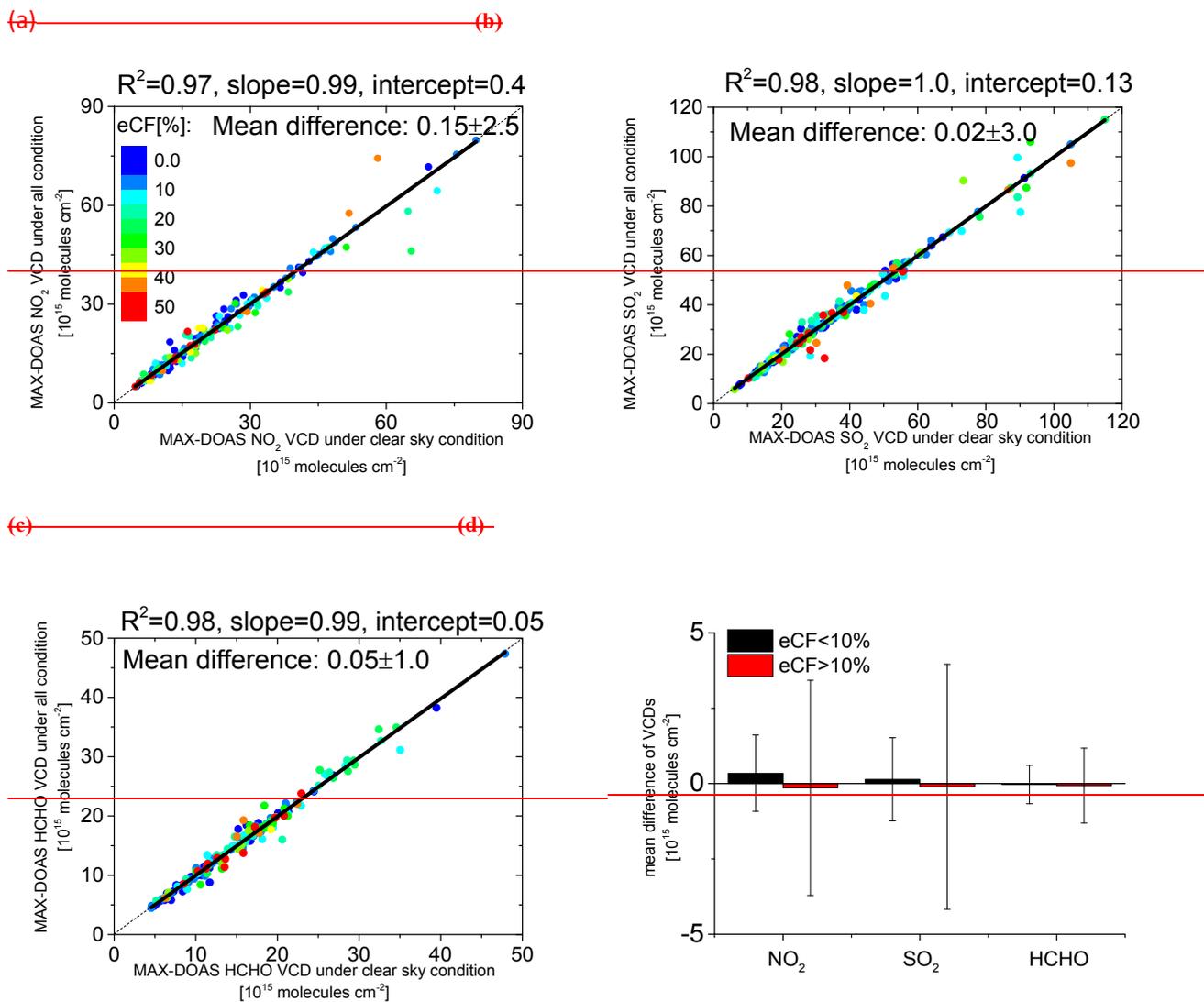
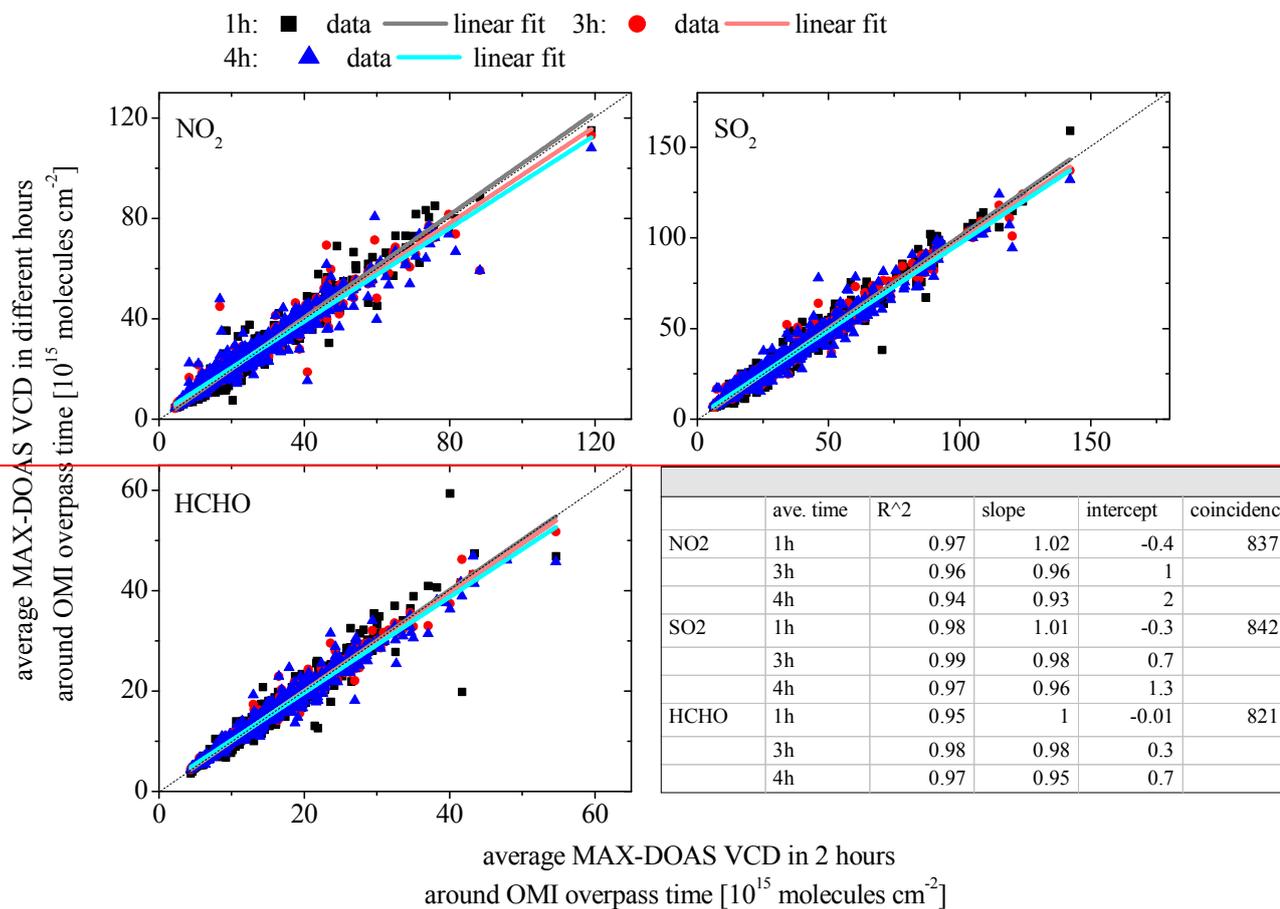


Figure 1: Wuxi city, in which the MAX-DOAS instrument is operated, is marked by the red dot in subfigure (a-1). Subfigures (a-2), (a-3c) and (a-4d) show maps of the averaged tropospheric VCDs of NO<sub>2</sub> from DOMINO 2, SO<sub>2</sub> and HCHO from BIRA derived from OMI observations over eastern China in the period from 2011 to 2014, respectively. The black dots indicate the location of Wuxi. Subfigure (b-1) shows the earth image around Wuxi MAX-DOAS station from google earth service; the rectangles indicate the ground pixel sizes of the

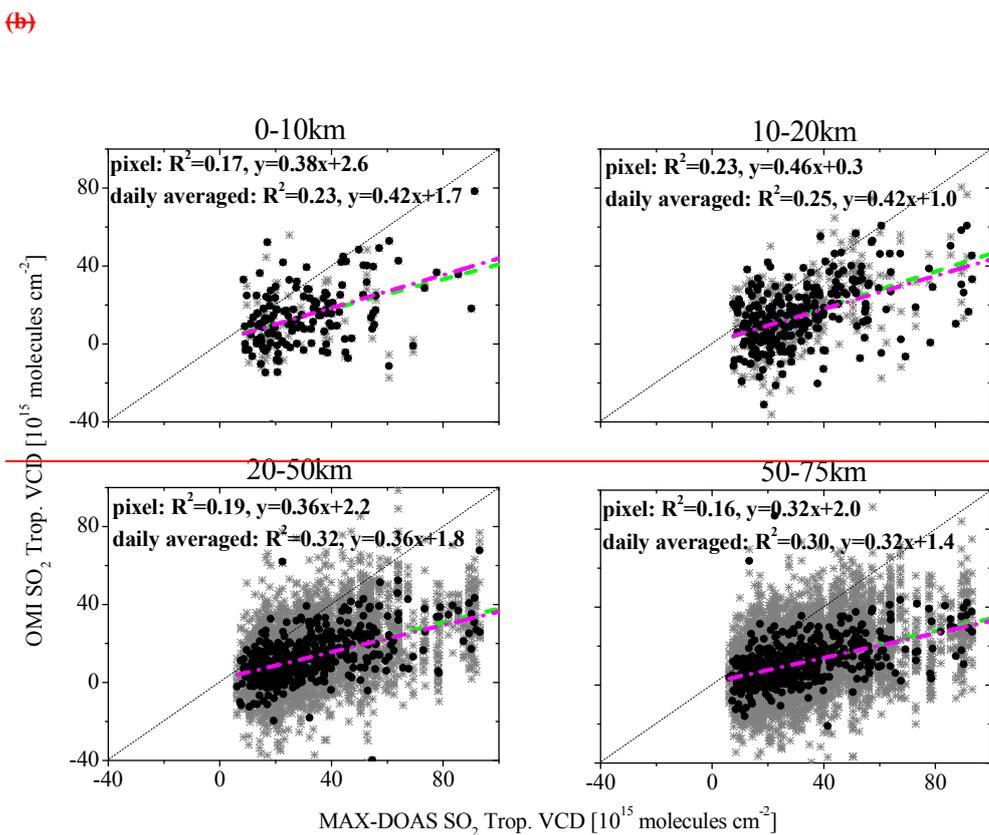
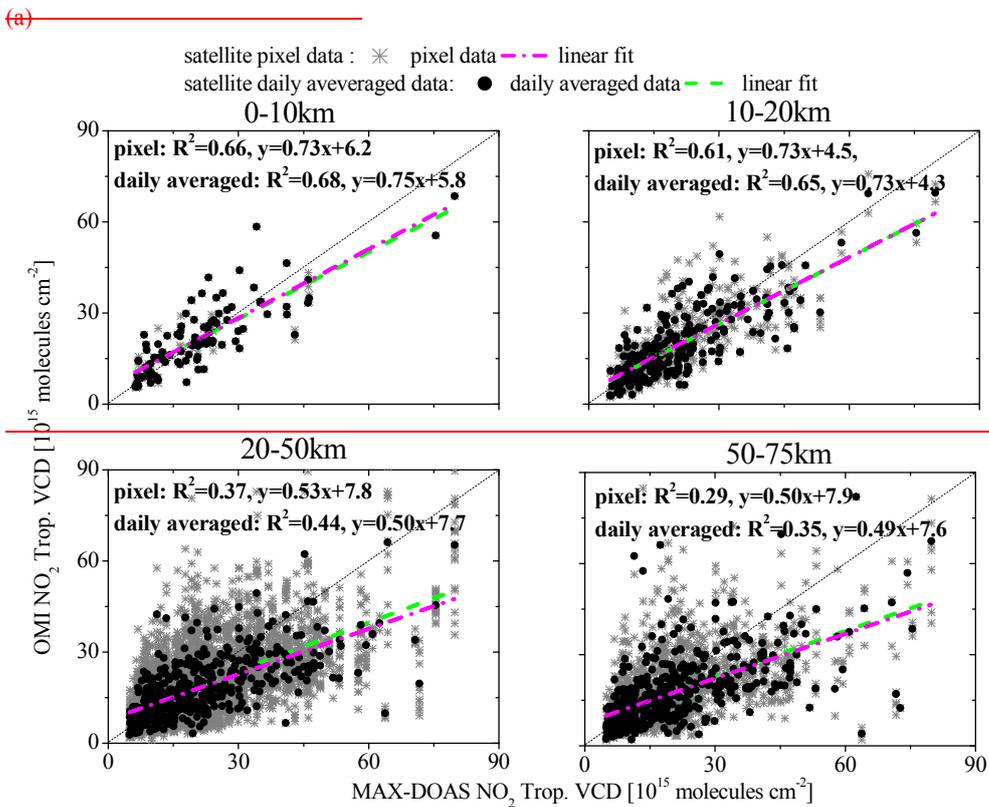
different satellite instruments used in this study. (GOME2-A phase 1 and phase 2 corresponding to the periods before and after 15 July 2013); the circles indicate areas with different radii around Wuxi. The subfigures of (b-2), (b-3) and (b-4) show averaged VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO for the same area as shown in (b-1); the black dots indicate the location of Wuxi and the green circles have a radius of 75km.



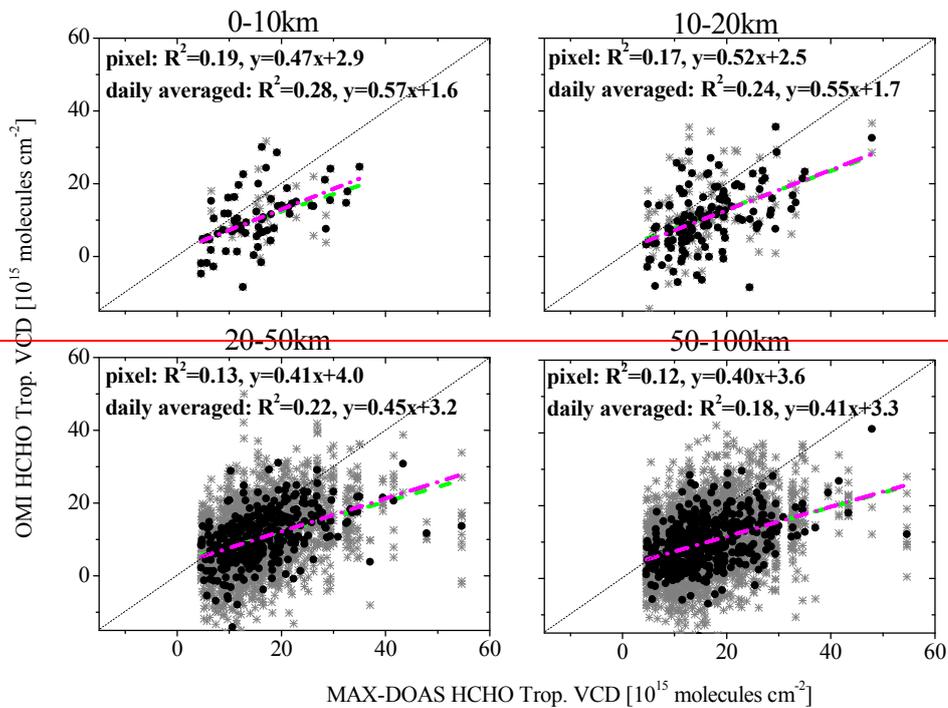
**Figure 2: Daily averaged (during two hours around the OMI overpass time) NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) tropospheric VCDs derived from MAX-DOAS observations under all sky conditions plotted against those under clear sky conditions. The colours indicate the eCF. The correlation coefficients, slopes, intercepts and mean differences  $\pm$  standard deviation are displayed in each subfigure. The mean differences for eCF < 10% and > 10% are plotted in subfigure (d) with the error bars denoting the respective standard deviations.**



**Figure 3: Averaged NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) tropospheric VCDs derived from MAX-DOAS observations in time periods of 1 hour (black dots), 3 hours (red dots) and 4 hours (blue dots) around the OMI overpass time plotted against those in the time period of 2 hours around the OMI overpass time. The linear regression lines for each time period and each species are plotted in each subfigure. The corresponding parameters are listed in the table.**



(c)



**Figure 4: Tropospheric VCDs of NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) derived from OMI observations for pixels within the distance bins of 0-10km, 10-20km, 20-50km and 50-75km away from the Wuxi MAX-DOAS station plotted against the coincident MAX-DOAS results. Only OMI data for the eCF<30% are included. For HCHO, only the data for a fit error < 7×10<sup>15</sup> molecules cm<sup>-2</sup> are included. The grey crosses and black dots show the data for individual satellite pixel and daily averaged data (averaged during two hours around the OMI overpass time), respectively. The linear regression lines and the parameters are shown in each subfigure for the pixel data (green dash-dot lines) and daily averaged data (magenta dash-dot lines), respectively.**

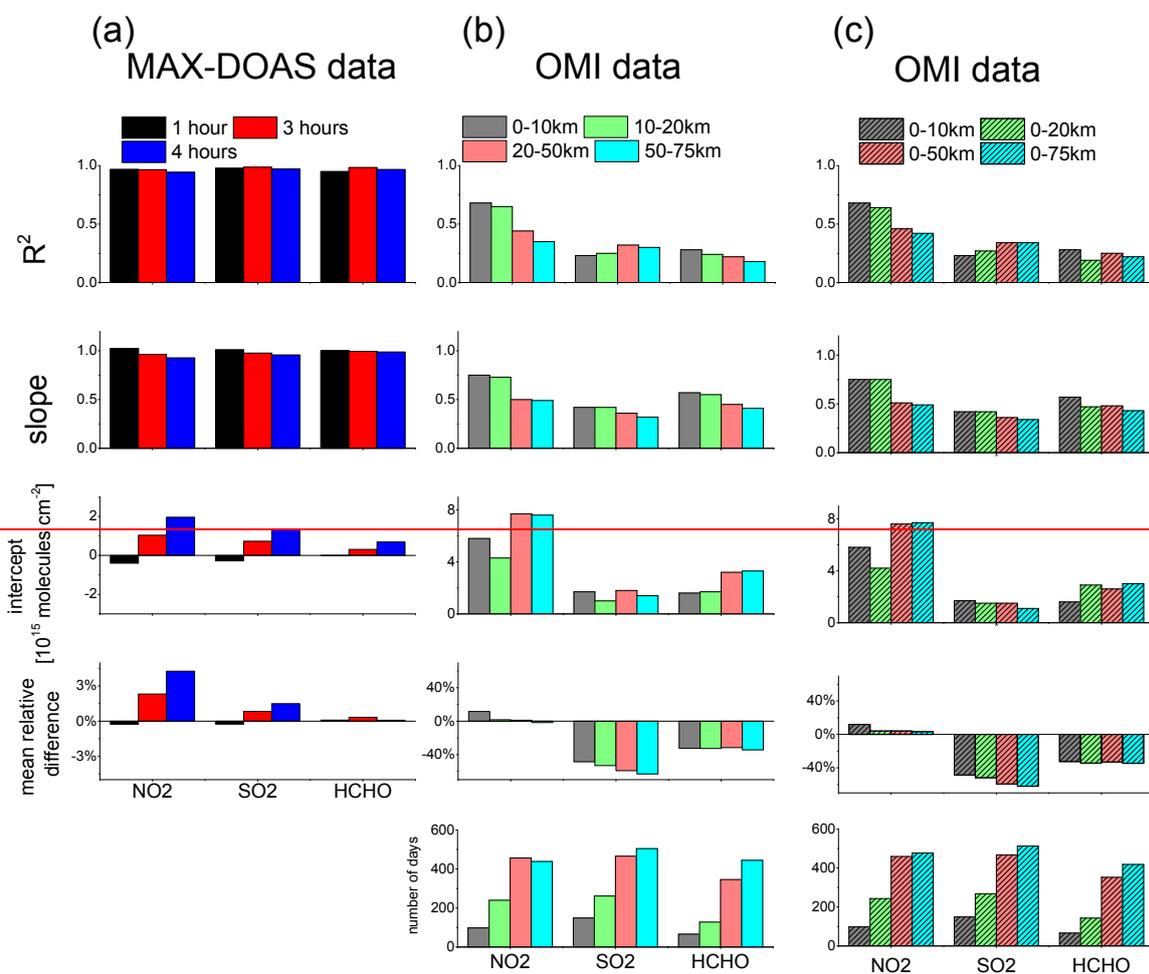


Figure 5: (a)  $R^2$ , slope and intercept of the linear regressions as well as the mean relative differences of the averaged MAX-DOAS tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO in the time periods of 1 hour, 3 hours and 4 hours around the OMI overpass time compared to those in the time period of 2 hours. (b)  $R^2$ , slope and intercept of the linear regressions as well as the mean relative differences of the averaged OMI tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO for the pixels within the distance bins of 0-10km, 10-20km, 20-50km and 50-75km compared to the coincident MAX-DOAS results. At the bottom also the numbers of the days for each comparison are shown. (c) Similar with (b), but for distance bins of 0-10km, 0-20km, 0-50km and 0-75km.

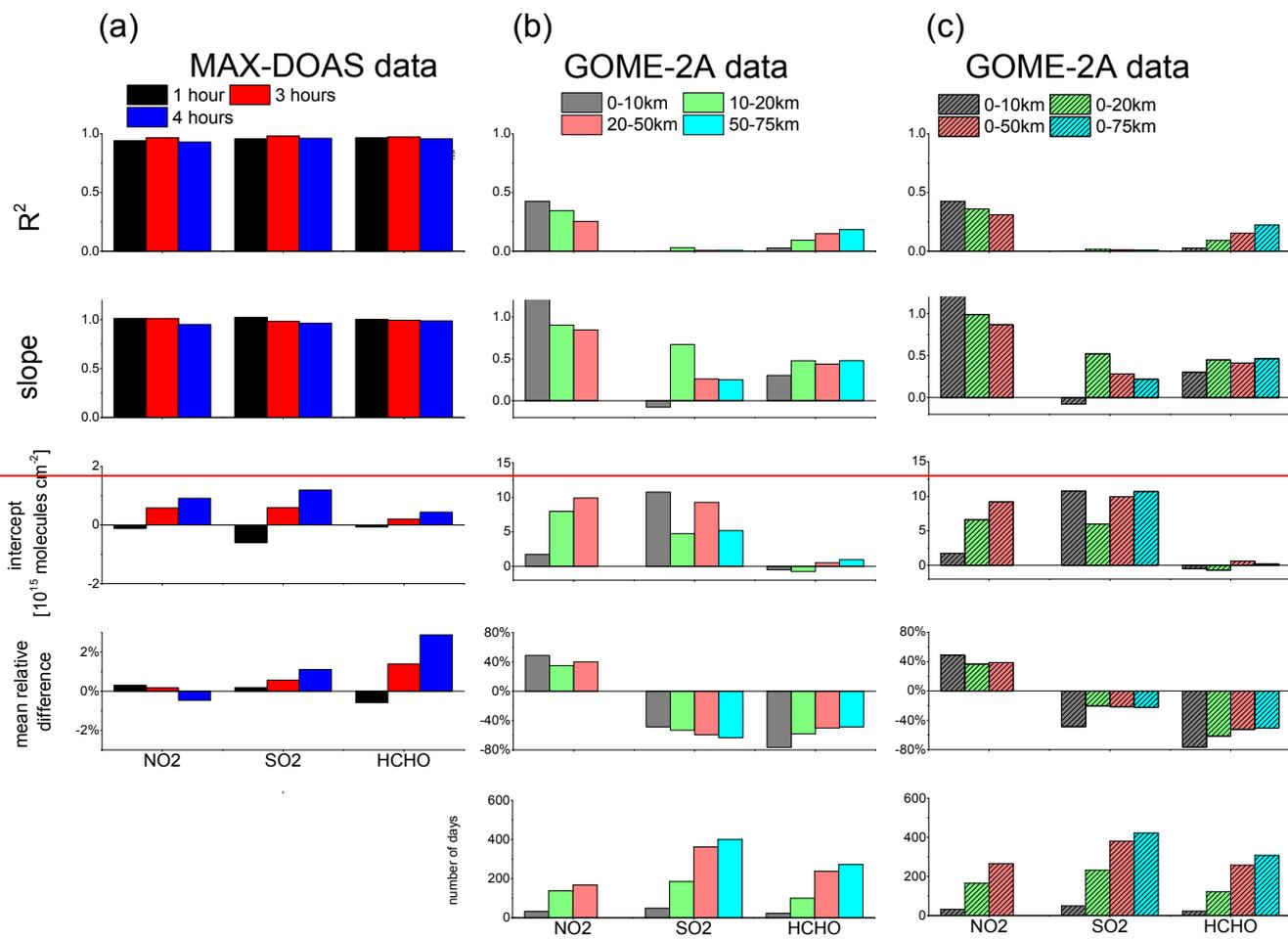
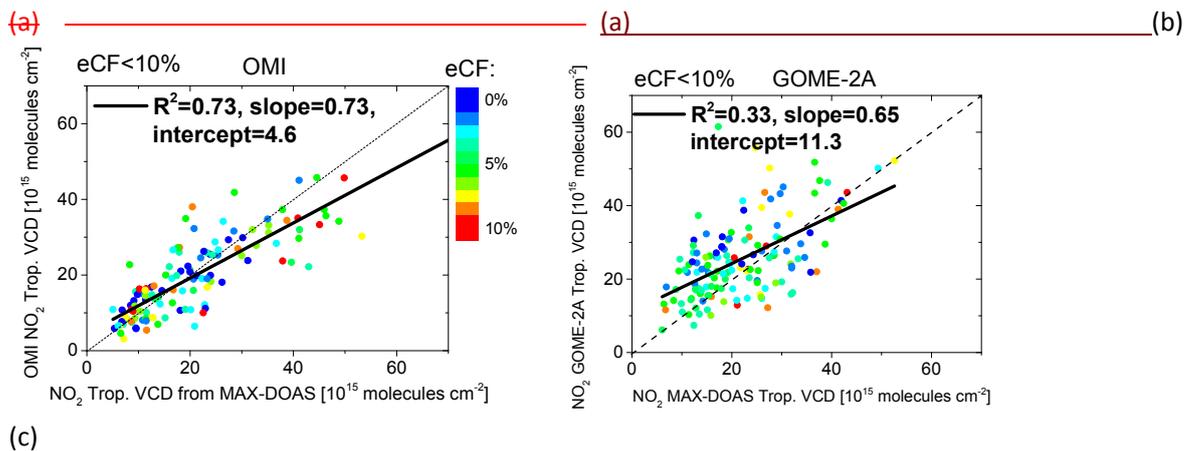
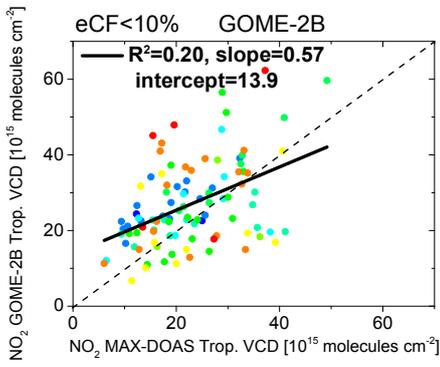
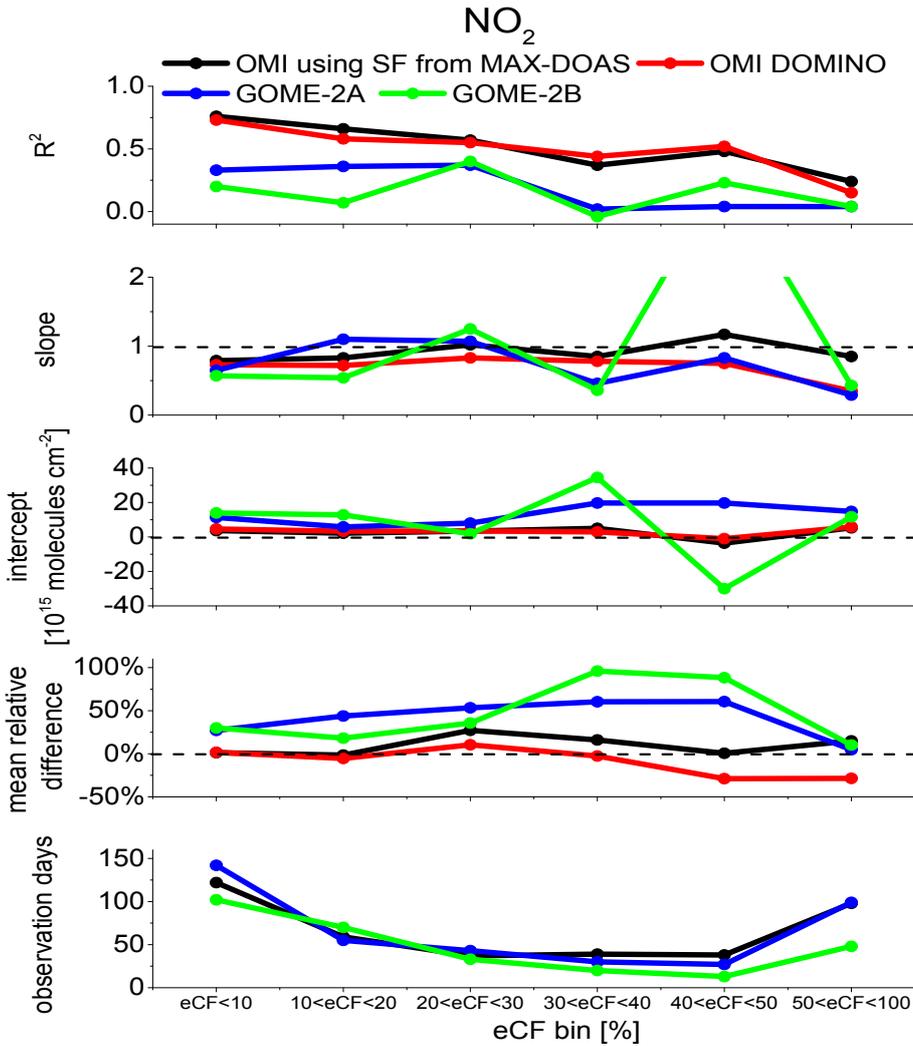


Figure 6: same as Fig. 5, but for GOME-2A data.





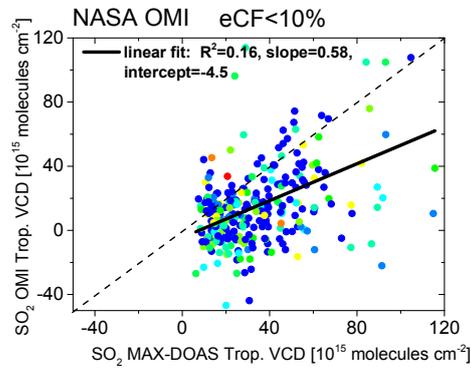
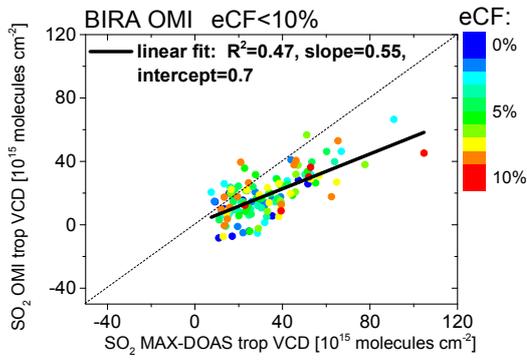
**Figure 72:** Daily average NO<sub>2</sub> tropospheric VCDs derived from OMI (a), GOME-2A (b) and GOME-2B (c) compared with the corresponding MAX-DOAS data for eCF<10%. The colors indicate the eCF.



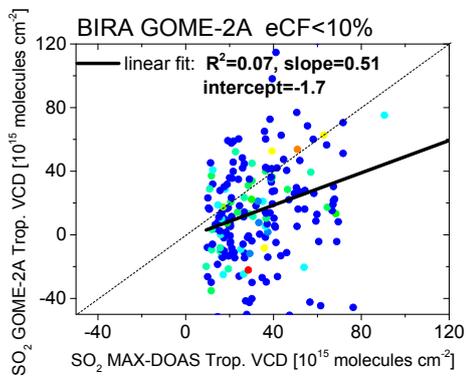
**Figure 83:** R<sup>2</sup>, slopes, intercepts, mean relative differences (and the number of available days) derived from the comparisons of the NO<sub>2</sub> VCDs from different satellite instruments to the MAX-DOAS results for the different eCF bins. Note that the black and red curves represent the improved OMI VCDs with the a-priori shape factors derived from Wuxi MAX-DOAS observations (see section 3.32) and for the DOMINO 2 product, respectively.

(a)

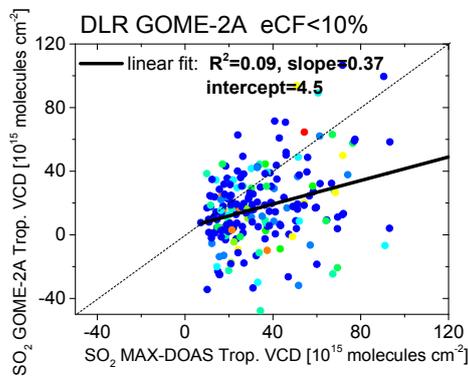
(b)



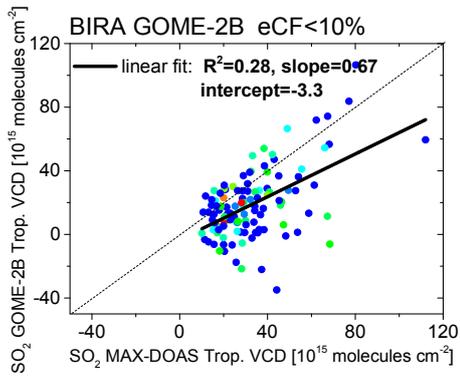
(c)



(d)



(e)



**Figure 94:** Daily averaged OMI SO<sub>2</sub> tropospheric VCDs from BIRA (a) and NASA (b), GOME-2A SO<sub>2</sub> tropospheric VCDs from BIRA (c) and DLR (d) and GOME-2B SO<sub>2</sub> tropospheric VCDs from BIRA (e) for eCF < 10% plotted versus the coincident MAX-DOAS results. The colors indicate the eCF.

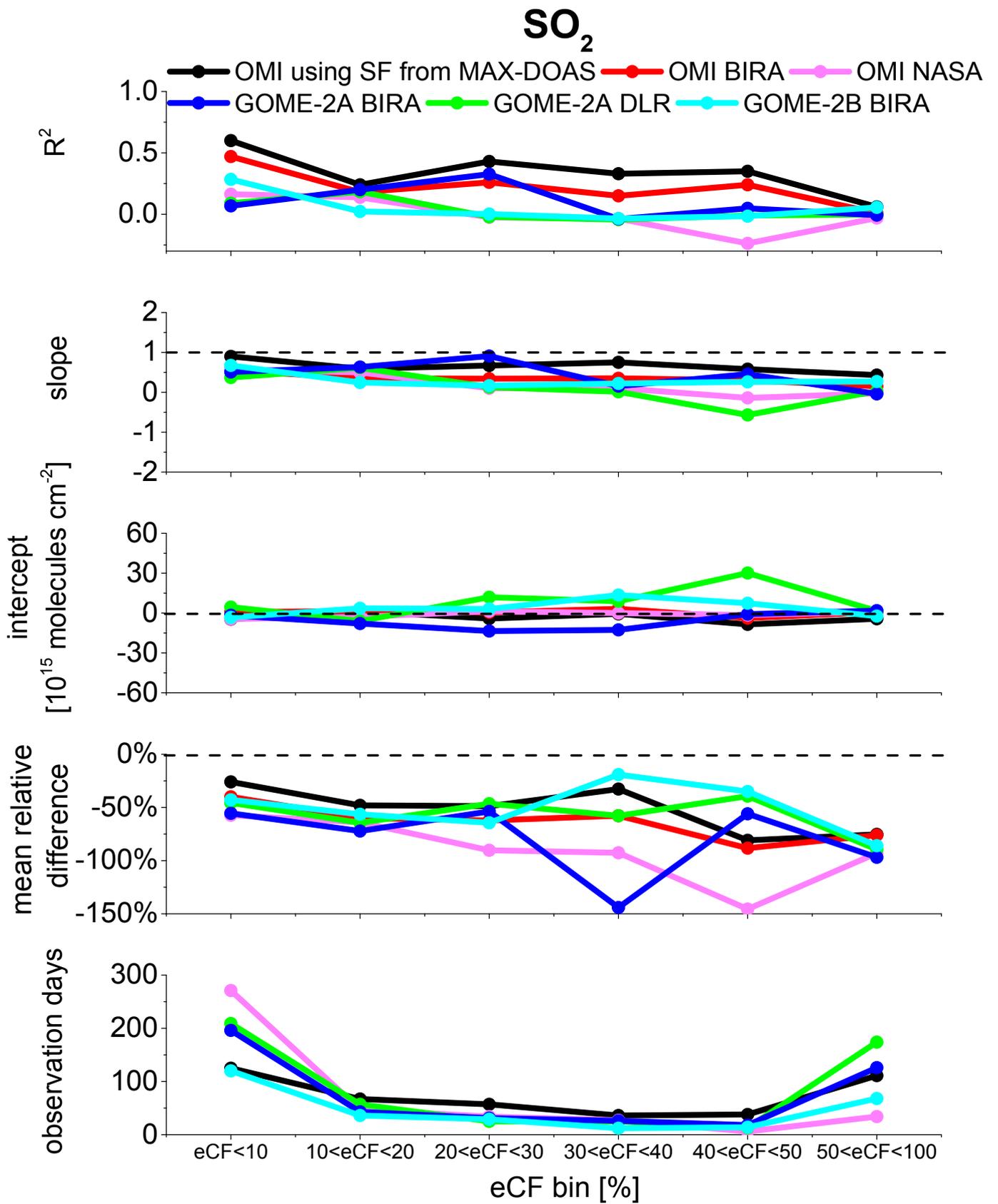
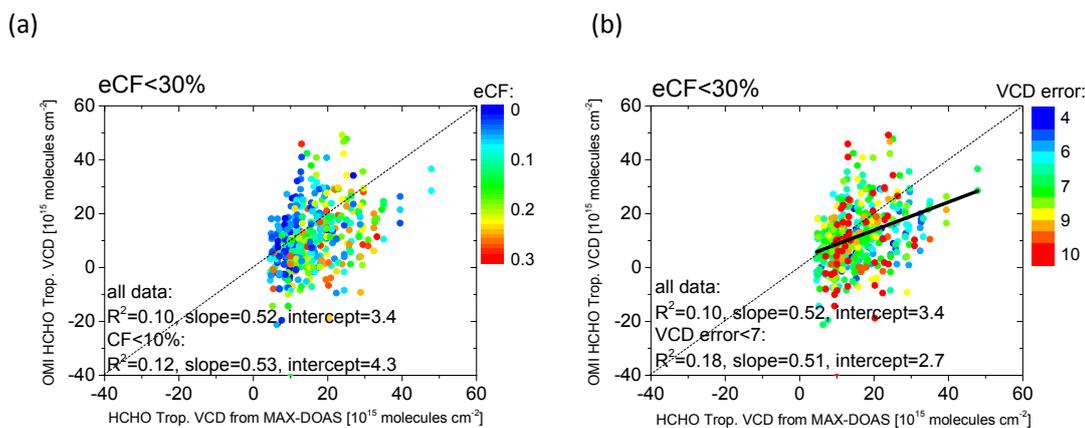
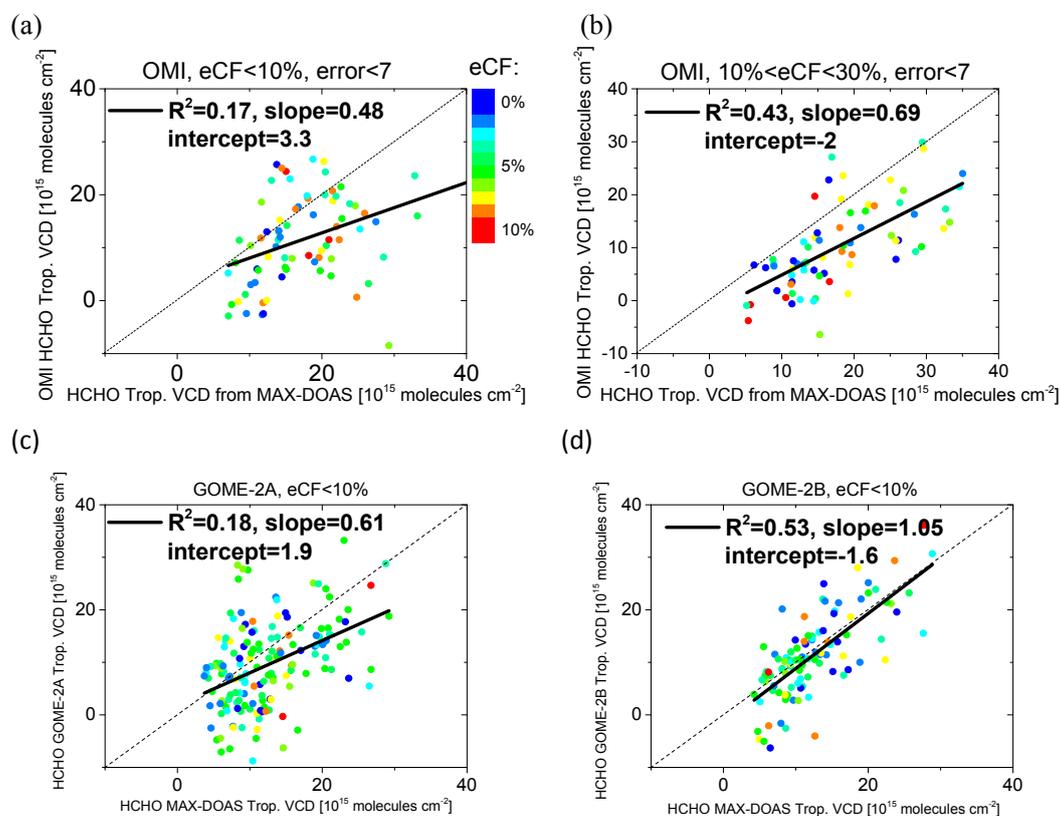


Figure 105: Same as figure Fig. 8-3 but for SO<sub>2</sub>.



**Figure 446:** (a) HCHO tropospheric VCDs for OMI pixels for  $eCF < 30\%$  are plotted against those derived from MAX-DOAS observations with the color map of  $eCF$ ; the linear regression parameters are acquired for  $eCF < 30\%$  and for  $eCF < 10\%$ , respectively. (b) Scattered plots are same as in (a), but with the color map of VCD fit error; linear regression parameters are acquired for all data and for VCD fit error  $< 7 \times 10^{15}$  molecules  $cm^{-2}$ .



**Figure 427:** Same as Fig.-72, but for HCHO.

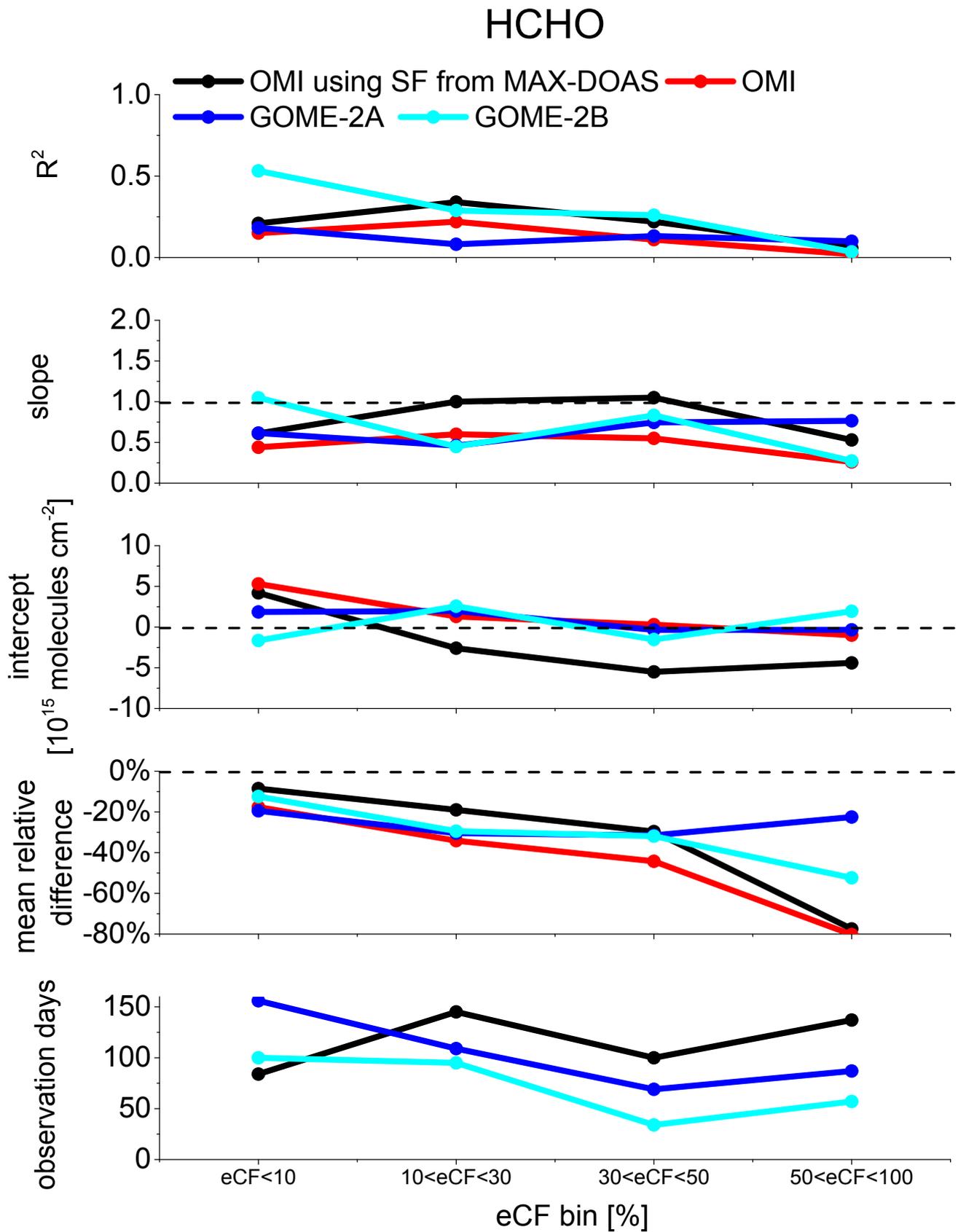
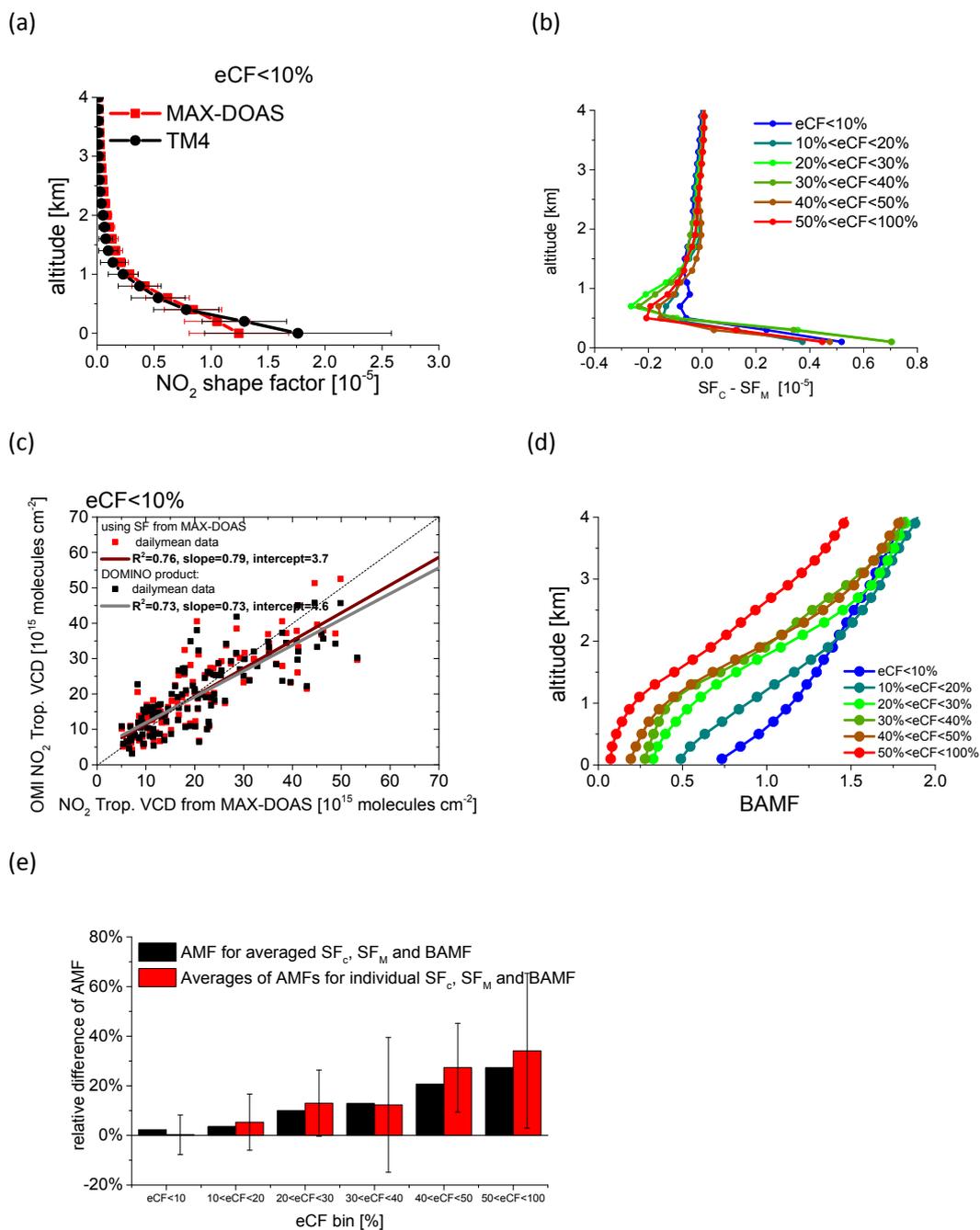


Figure 138: Same as Fig. 83 but for HCHO.



**Figure 149:** (a) Average NO<sub>2</sub> SFs and standard deviations derived from the MAX-DOAS observations and from the TM4 CTM (for the DOMINO product) for eCF  $\leq$  10%. (b) Averaged differences between the NO<sub>2</sub> SFs from CTM (SF<sub>C</sub>) and from MAX-DOAS (SF<sub>M</sub>) for different eCF bins. (c) Daily averages of the original DOMINO NO<sub>2</sub> product and modified NO<sub>2</sub> product (based on MAX-DOAS SF) plotted against those from MAX-DOAS for eCF  $\leq$  10%. (d) Averaged BAMF for satellite observation for different eCF bins. (e) Relative difference (RD) of satellite AMF using SF<sub>C</sub> (AMF<sub>CTM</sub>) or SF<sub>M</sub> (AMF<sub>MAX-DOAS</sub>) for different eCF bins. The error bars indicate the standard deviation of the RDs for each eCF bin. Black columns denote the RDs derived from the averaged SF<sub>C</sub>, SF<sub>M</sub> and BAMF (shown in subfigure (b) and (d)); red columns denote the averaged RDs for individual SF<sub>C</sub>, SF<sub>M</sub> and BAMF of each satellite observation.

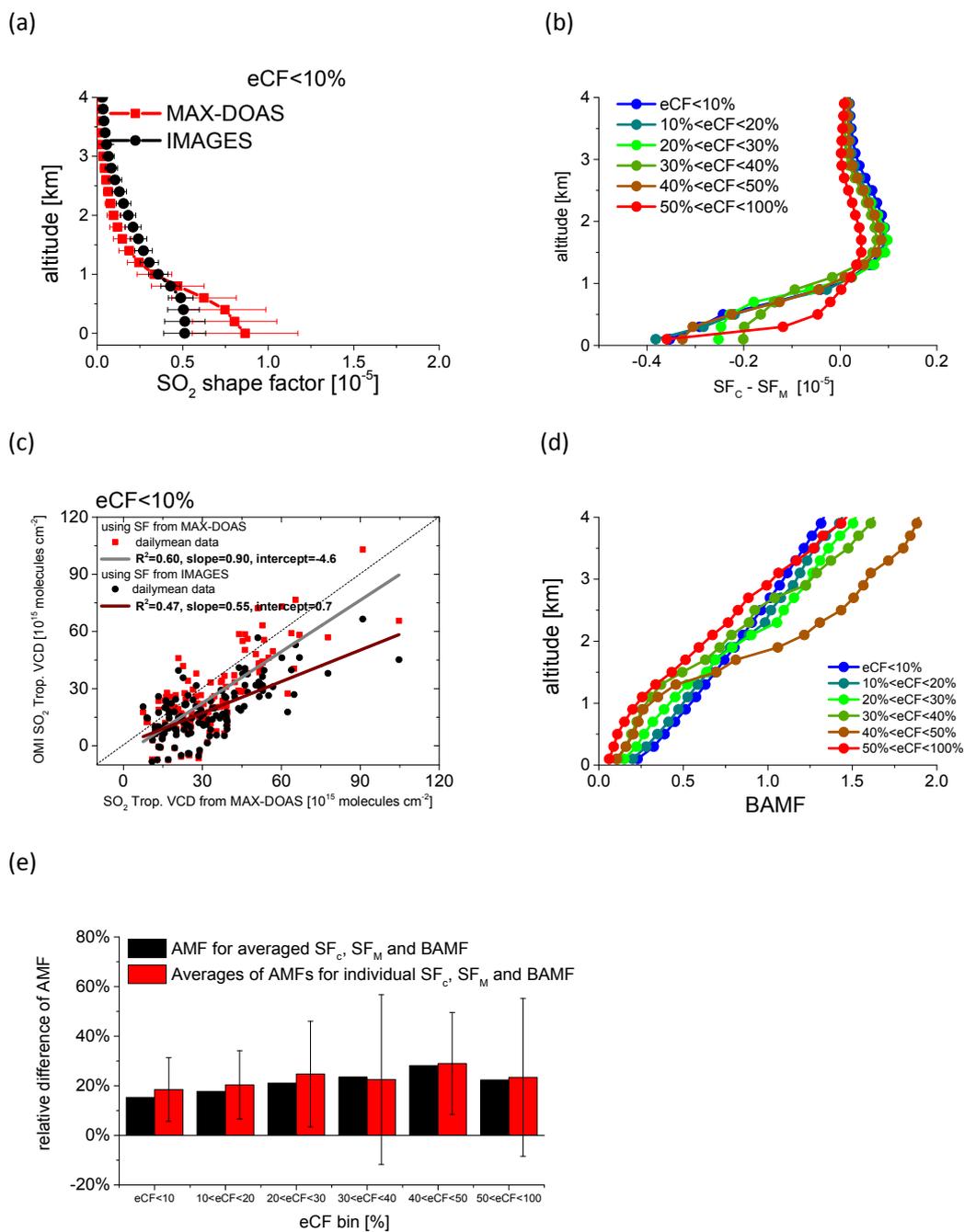
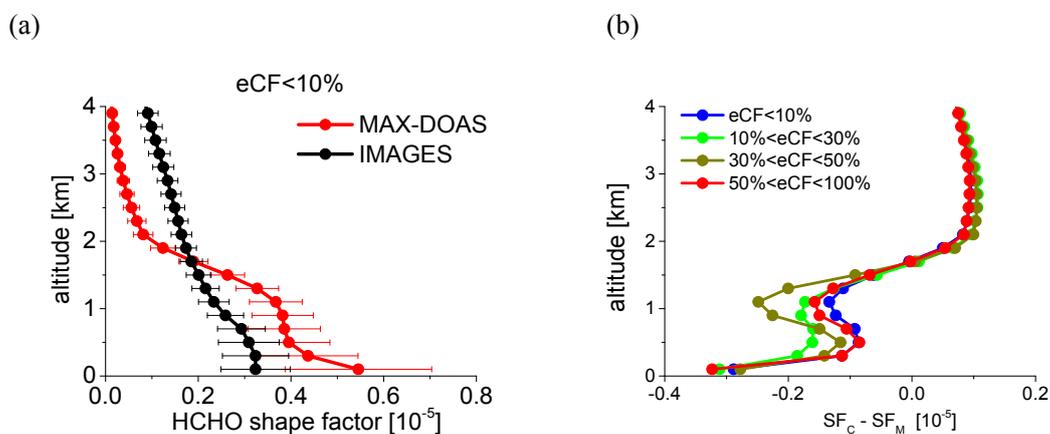
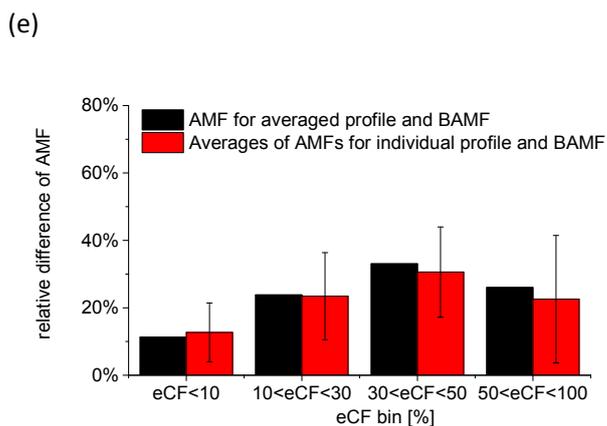
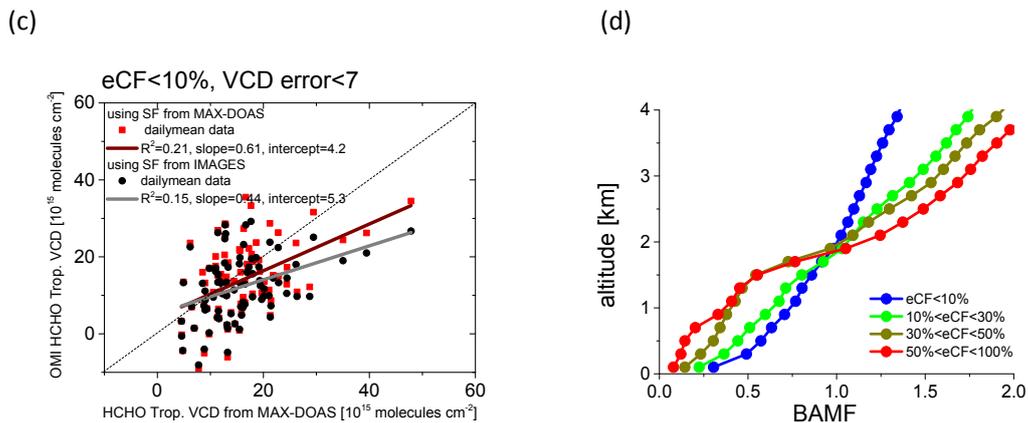
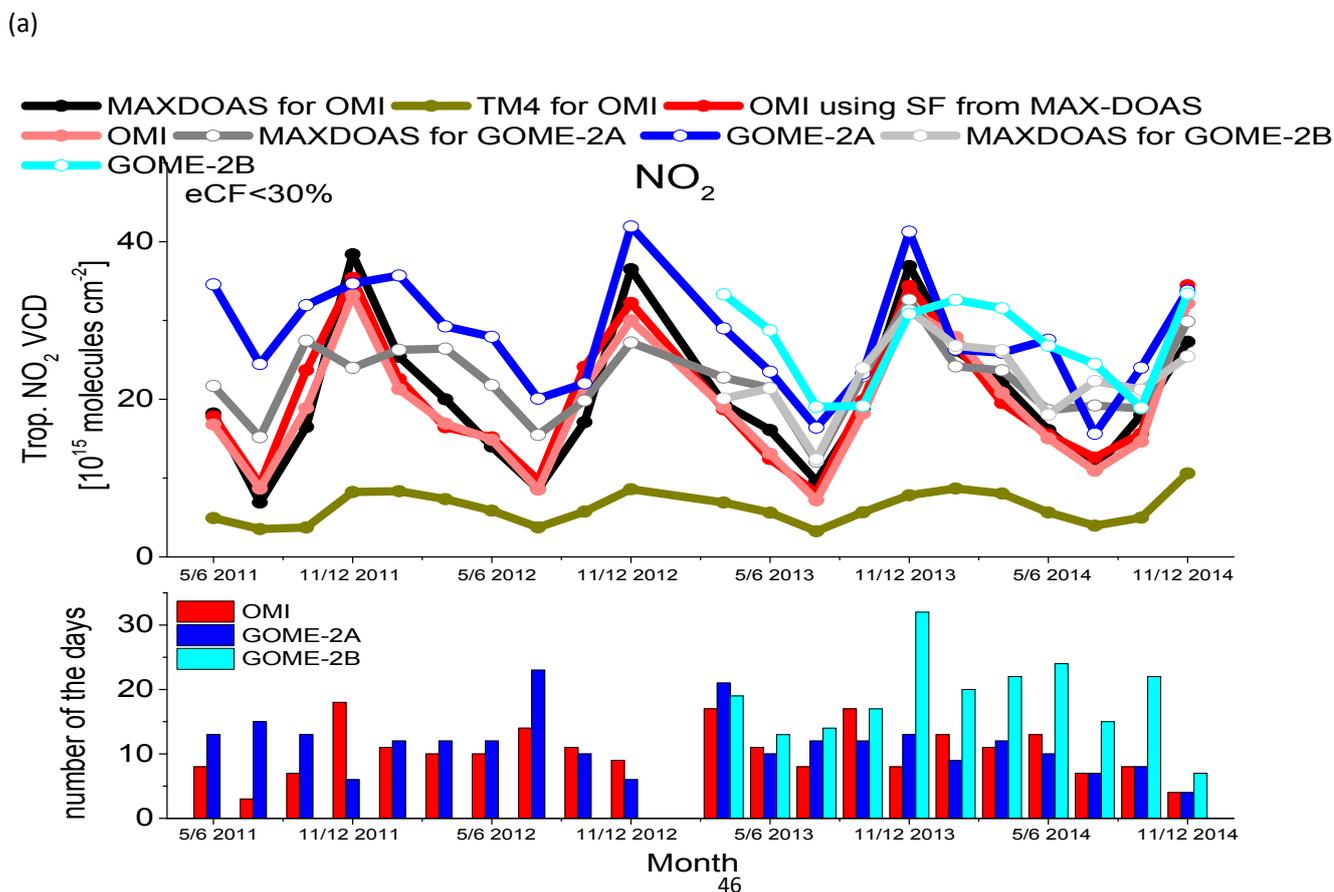


Figure 105: Similar as Fig. 944 but for the OMI BIRA  $SO_2$  product. Note that the SF for the OMI BIRA product is obtained from the IMAGES CTM.

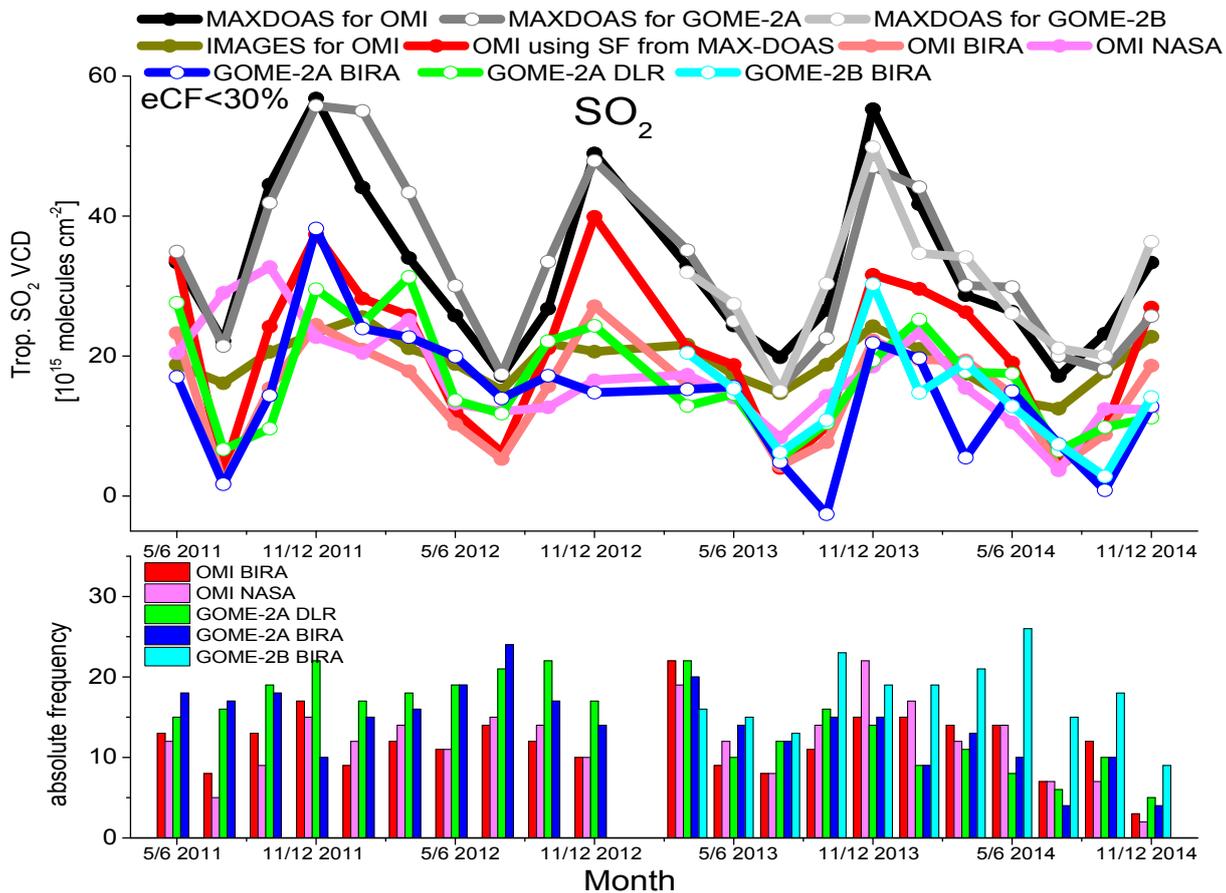




**Figure 4611:** Same as Fig. 449 but for the OMI BIRA HCHO product and eCF bins of 0-10%, 10%-30%, 30% -50% and 50% -100%. Note that the SF for the OMI BIRA product is obtained from the IMAGES CTM.



(b)



(c)

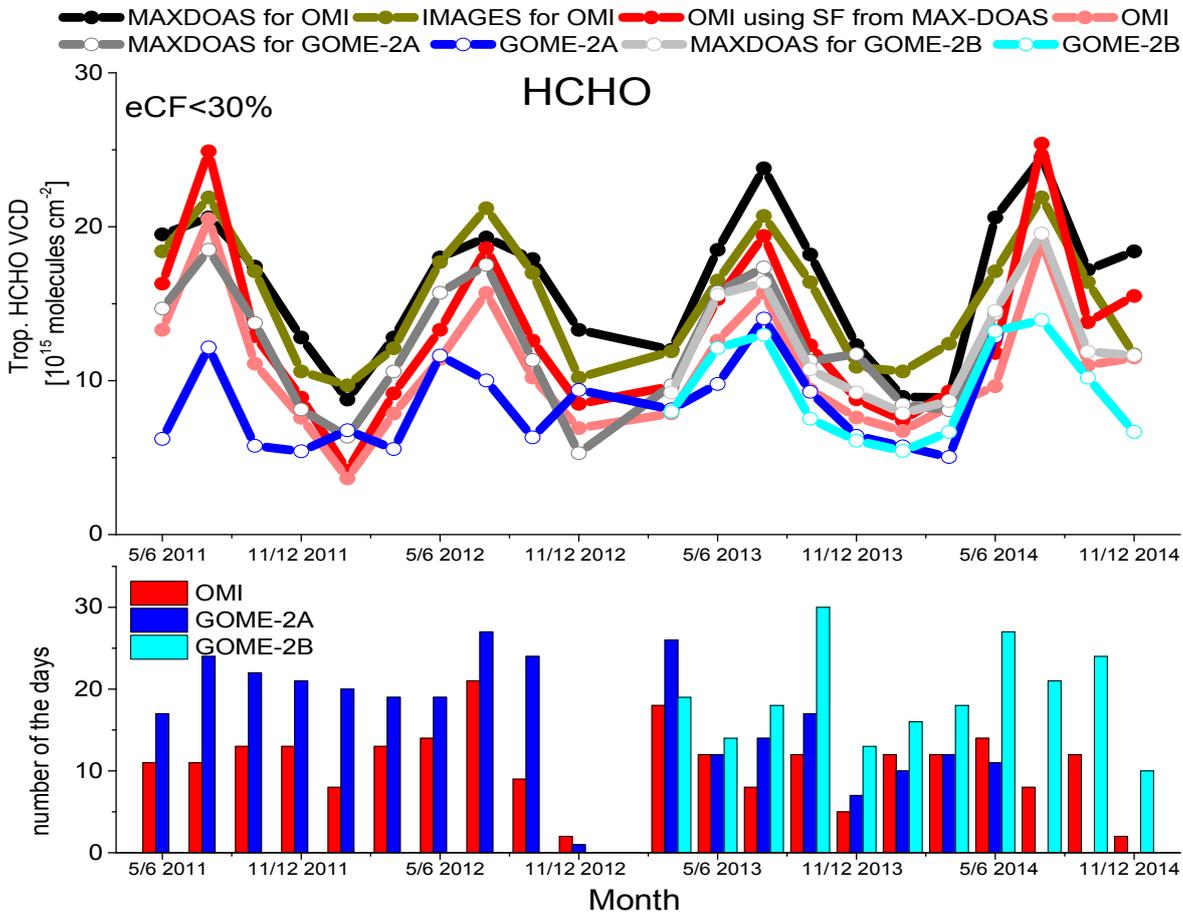


Figure 1712: Bi-monthly averaged tropospheric VCDs of NO<sub>2</sub> (a), SO<sub>2</sub> (b)<sub>2</sub> and HCHO (c) derived from coincident satellite and MAX-DOAS observations for eCF <30%. Also shown are the corresponding CTM results (TM4 for NO<sub>2</sub>, IMAGES for SO<sub>2</sub> and HCHO). In all subfigures the red and light red lines indicate the improved OMI tropospheric VCDs using the SFs from MAX-DOAS and the VCDs from the original OMI products, respectively. The numbers of the available days are shown in the bottom panel of each subfigure.

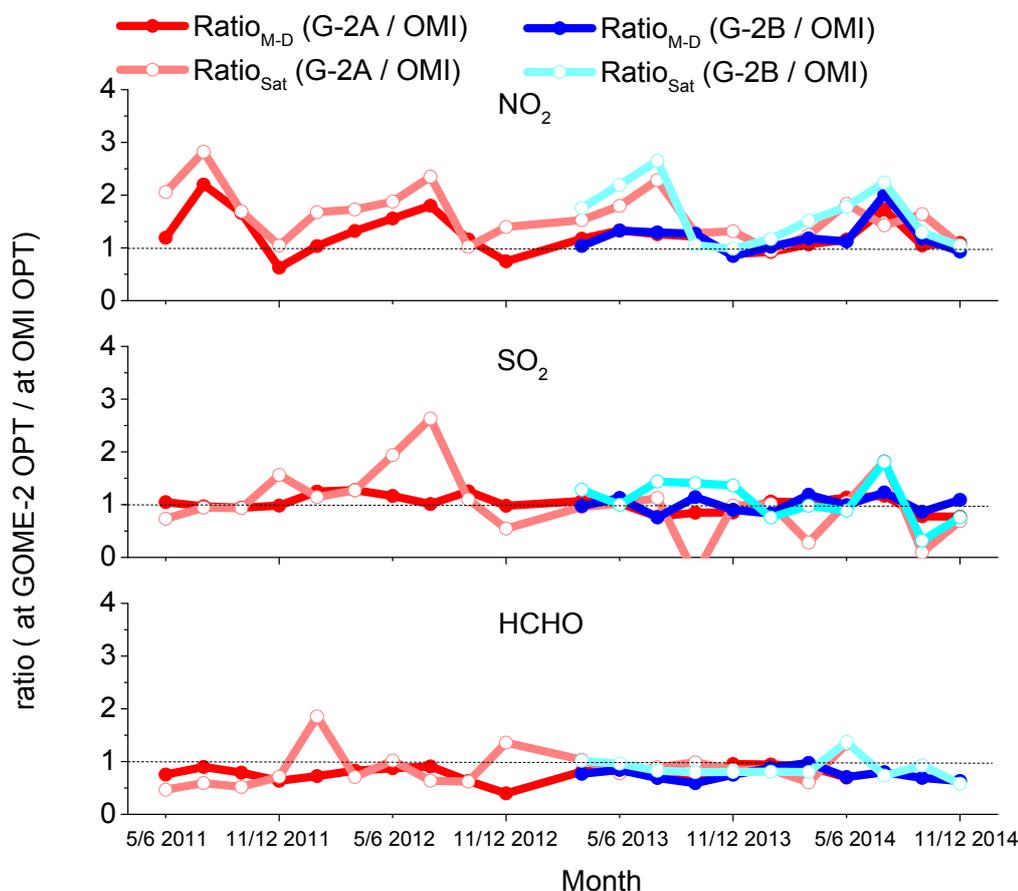


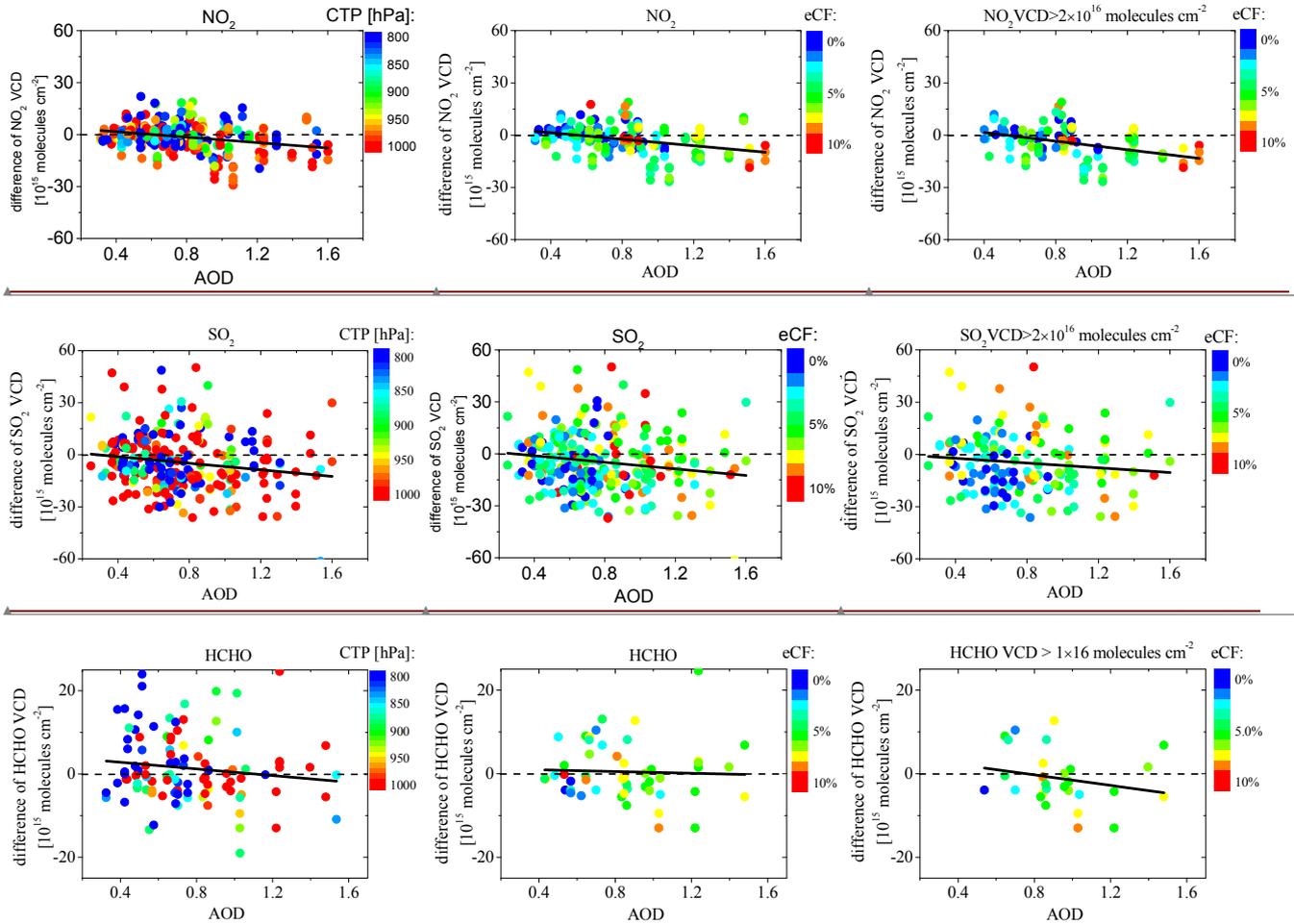
Figure 138: Ratios between the bi-monthly mean tropospheric VCDs from GOME-2A\B and OMI (Ratio<sub>Sat</sub>) as well as the ratios between the corresponding MAX-DOAS observations (Ratio<sub>M-D</sub>) for NO<sub>2</sub> (a), SO<sub>2</sub> (b)<sub>2</sub> and HCHO (c), respectively. The light red (dark red) and light blue (dark blue) curves are corresponding to GOME-2A and GOME-2B results (coincident MAX-DOAS results with GOME-2A and GOME-2B), respectively. Note that for SO<sub>2</sub> the OMI and GOME-2A data from BIRA are used for the ratio calculations. The mean ratios for the shown data sets are presented in Table 1.

(a)

eCF<10%

eCF<10% and CP>900hPa

eCF<10%, CP>900hPa, VCD> threshold

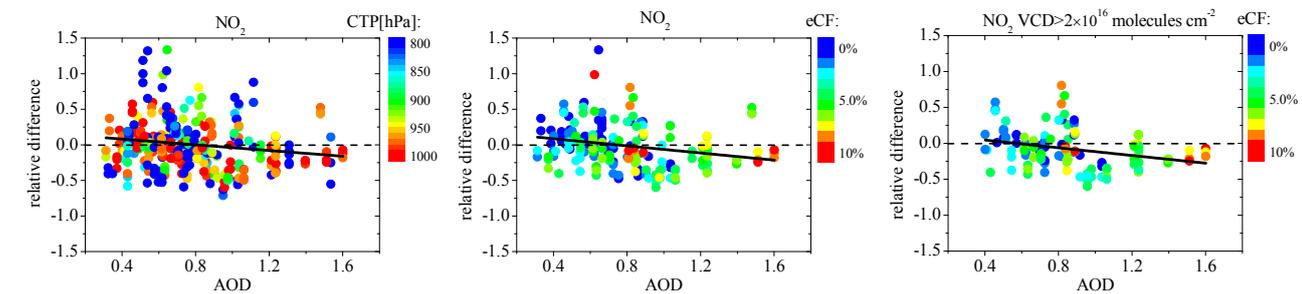


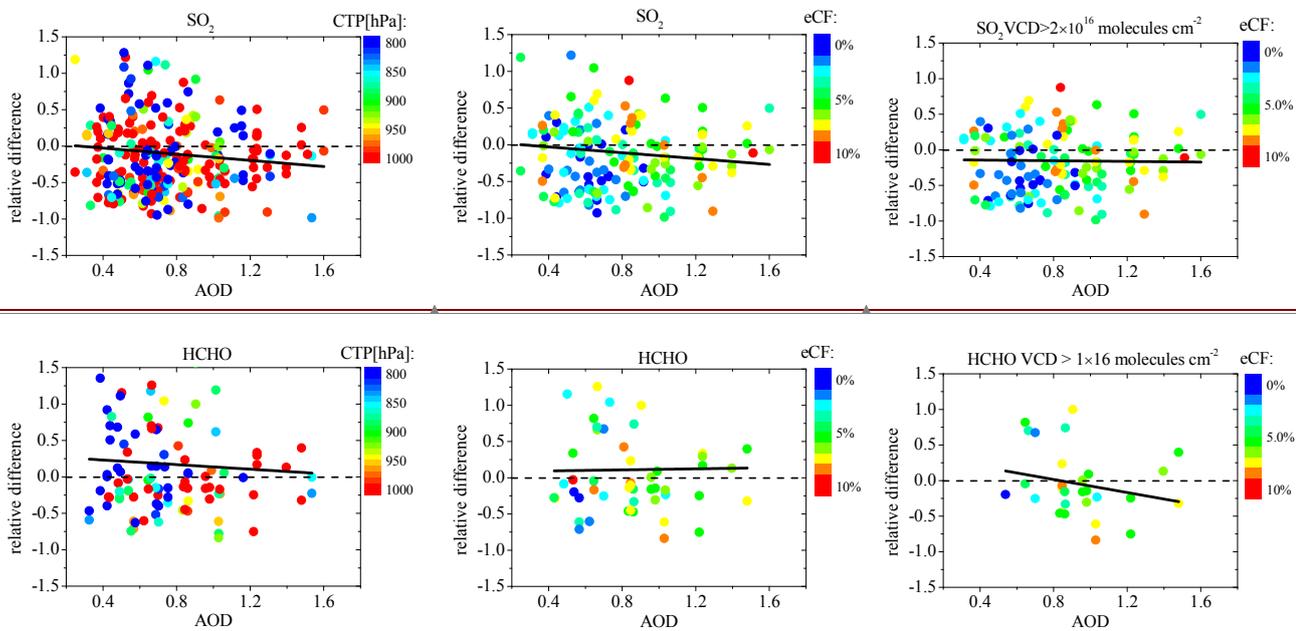
(b)

eCF<10%

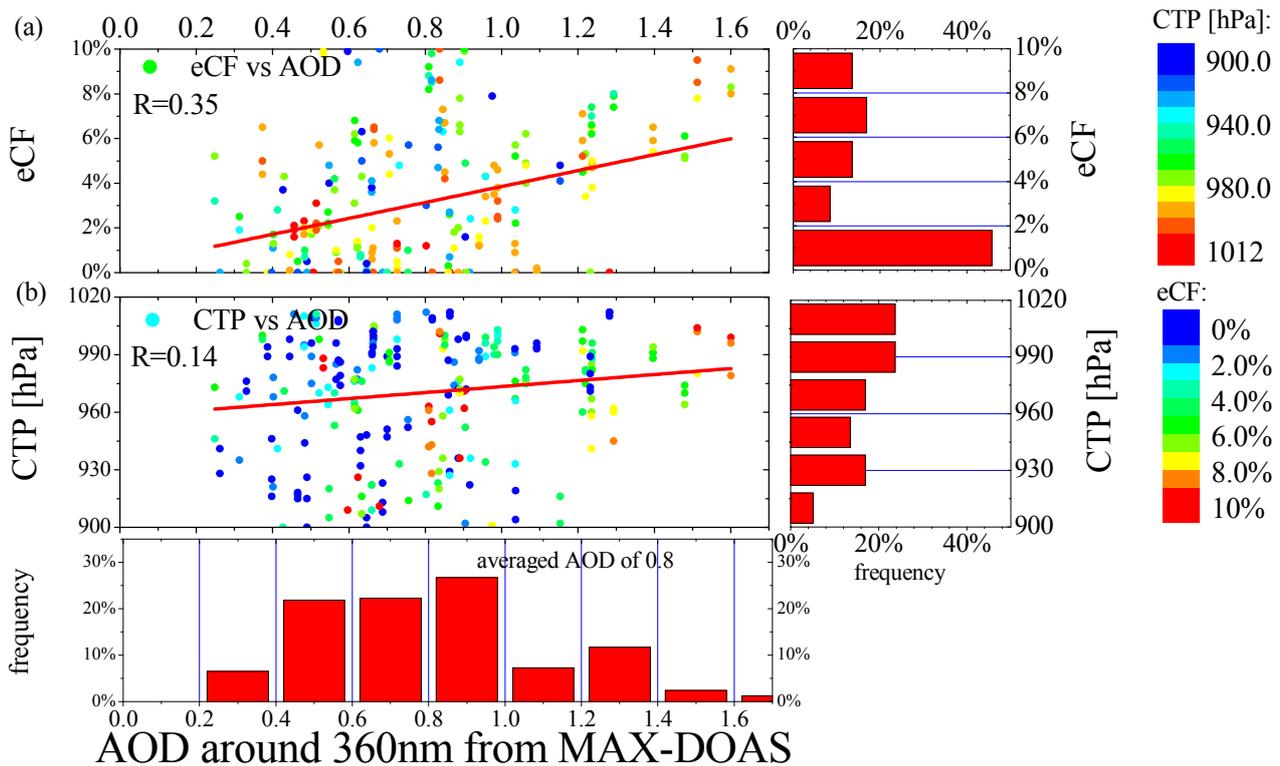
eCF<10% and CP>900hPa

eCF<10%, CP>900hPa, VCD> threshold

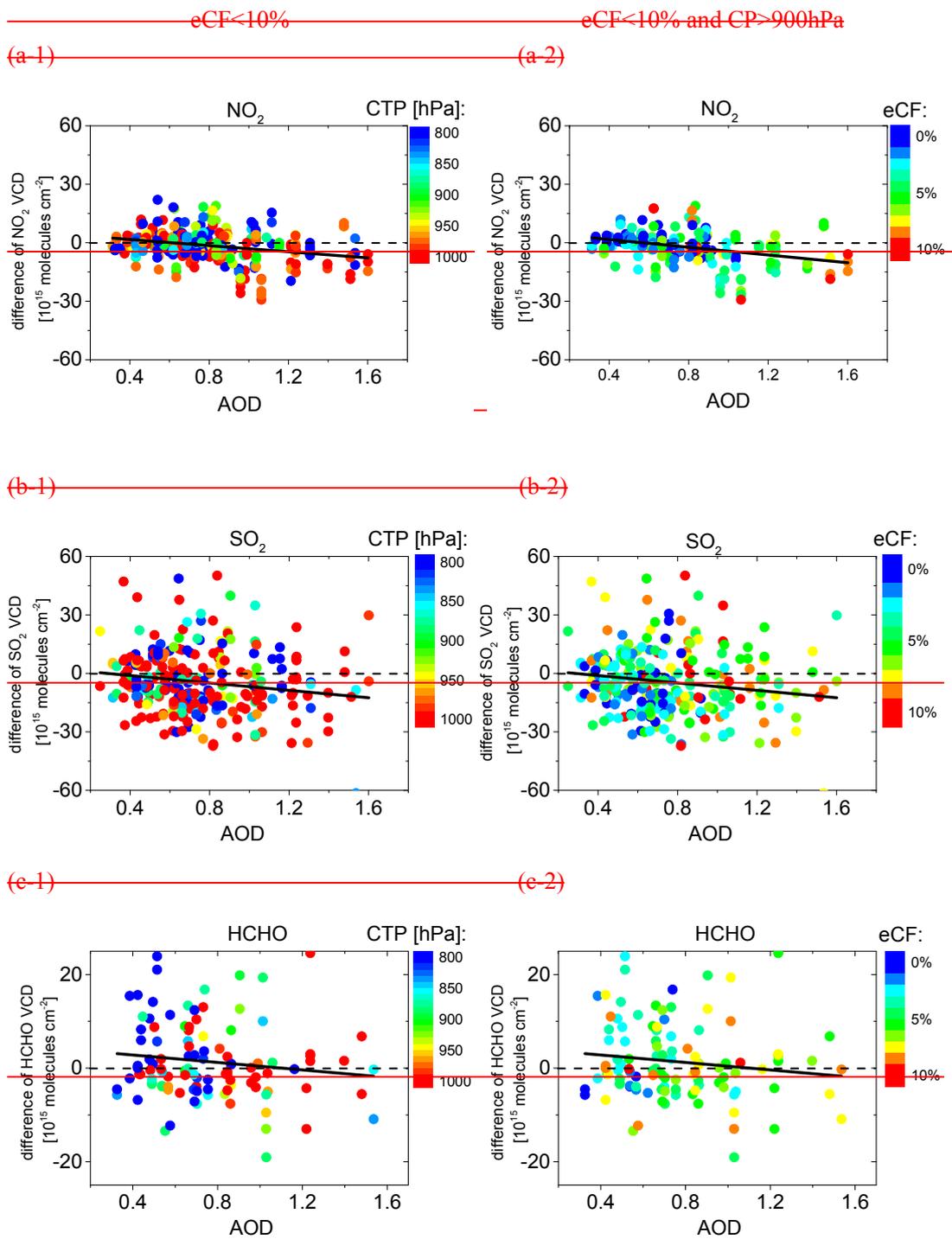




**Figure 14: absolute differences (a) and relative differences (b) of tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO between for individual OMI observations and MAX-DOAS observations plotted against the AODs derived from the MAX-DOAS observations. The data are differently screened in the left, center and right panels: eCF < 10% for the left; eCF < 10% and CTP > 900hPa for the center; and eCF < 10%, CTP > 900hPa, and VCD > a specific threshold for the right (see text). Note that the OMI VCDs are the modified values using SFs derived from MAX-DOAS observations.**

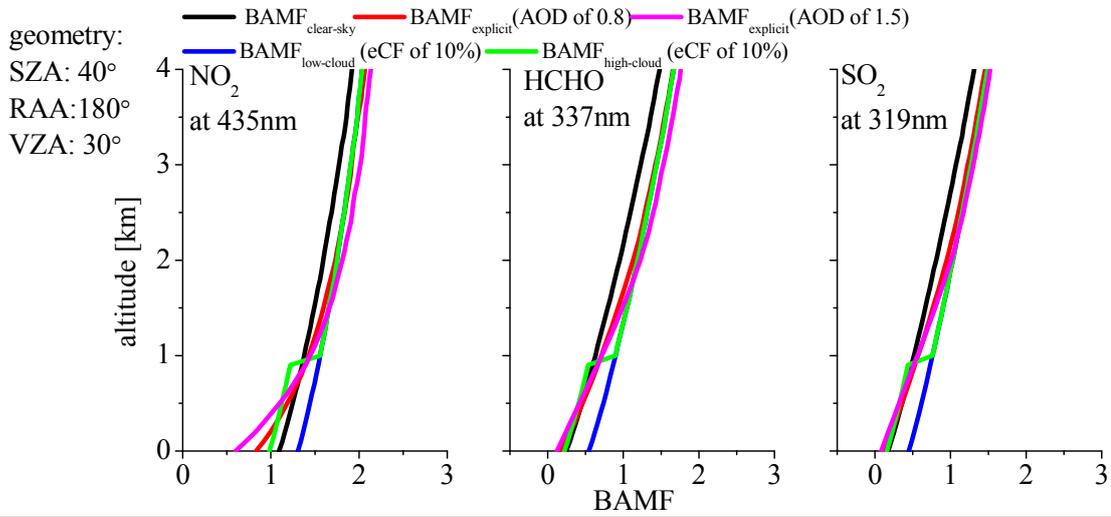


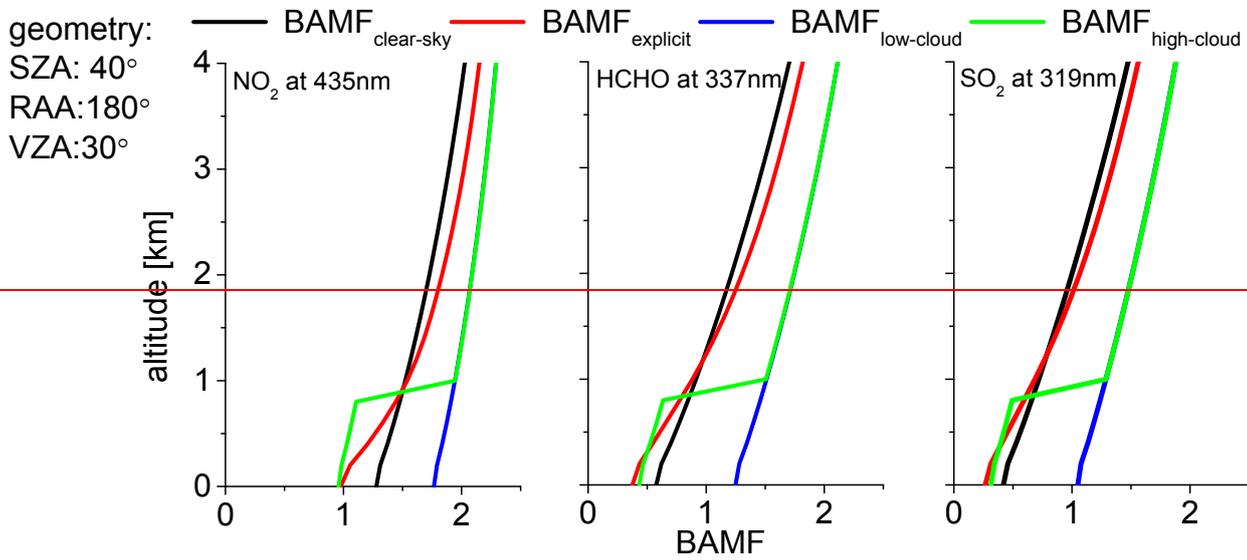
**Figure 15: eCF and CTP from the OMI cloud algorithm for individual OMI observations are plotted against AOD at 360nm derived from MAX-DOAS observation during the whole measurement in the condition of (for eCF < 10% and CTP > 900hPa). The red bars on the right and bottom indicate the frequency of eCF, CTP, and AOD in different value intervals. The red lines are the linear regressions of the scatter plots. The correlation coefficients are shown in the figure. The color of the dots in (a) and (b) indicates CTP and eCF, respectively.**



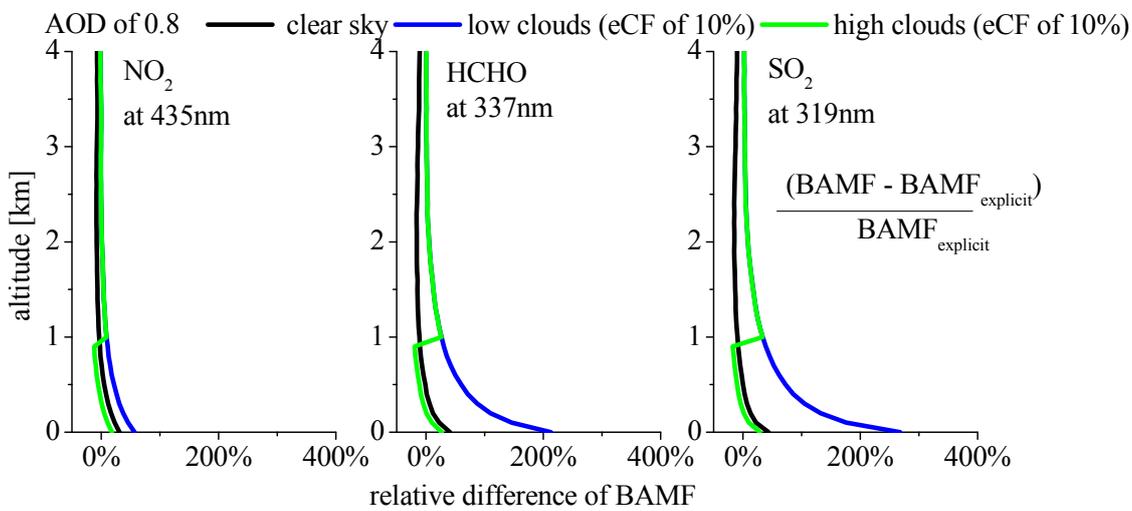
**Figure 19: Left panels: Differences of tropospheric VCDs of  $\text{NO}_2$  (a-1),  $\text{SO}_2$  (b-1) and HCHO (c-1) between for individual OMI observations (for  $eCF < 10\%$ ) and MAX-DOAS observations plotted against the AODs derived from the MAX-DOAS observations.— Right panels: Same data as left, but observations with  $CTP < 900\text{hPa}$  are skipped. The colours indicate the CTP (left) or  $eCF$  (right).—**

-(a)

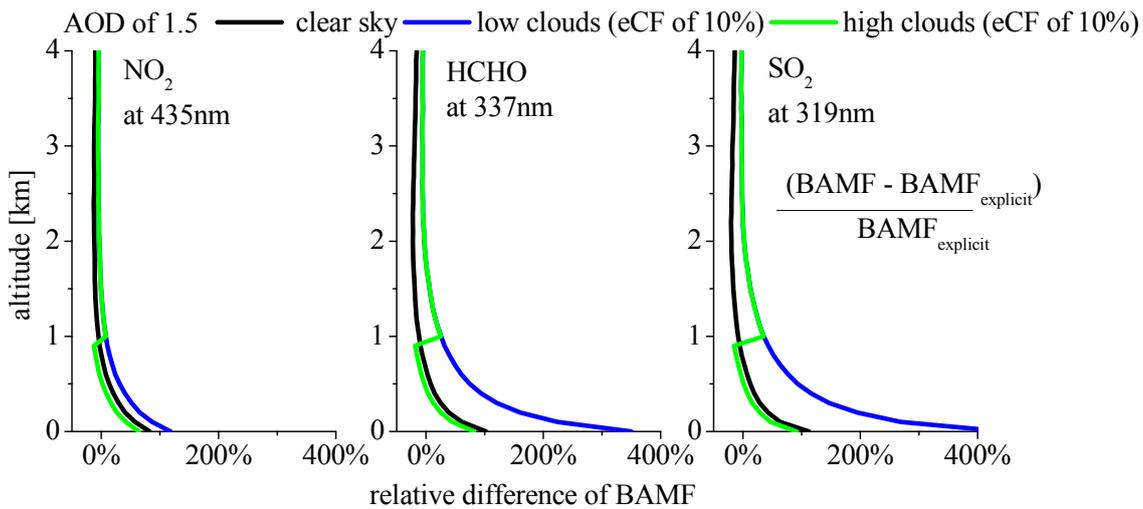




(b) ————— (c)



(c)



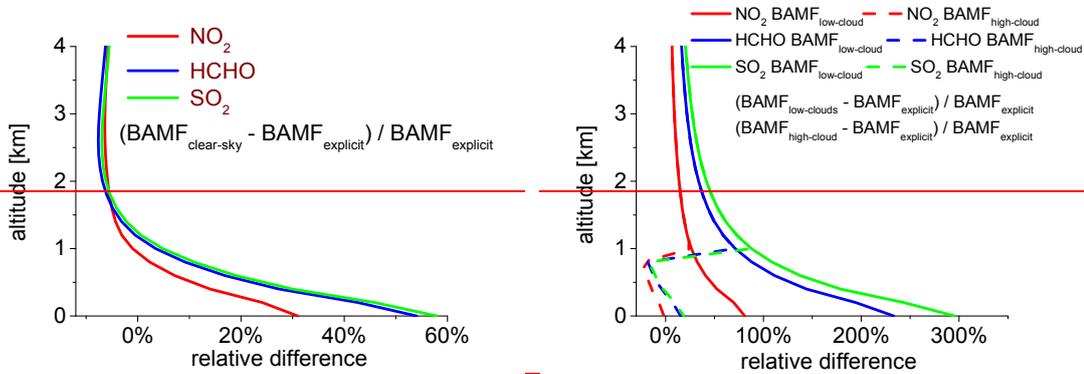


Figure 20: (a) Simulated  $BAMF_{clear-sky}$ ,  $BAMF_{explicit}$  for AOD of 0.8 and 1.5,  $BAMF_{low-cloud}$  (clouds at surface) and  $BAMF_{high-cloud}$  (clouds at 1 km) of  $NO_2$  at 435nm, HCHO at 337nm and  $SO_2$  at 319nm for one typical nadir-satellite observation (SZA of  $40^\circ$ , RAA of  $180^\circ$  and VZA of  $30^\circ$ ). The effective cloud fraction of 10% is used in the calculations. (b) Relative differences between of  $BAMF_{clear-sky}$  and  $BAMF_{explicit}$  (c) Relative differences between  $BAMF_{low-cloud}$  and  $BAMF_{high-cloud}$  and compared to  $BAMF_{explicit}$  for AOD of 0.8. (c) Same with (b) but  $BAMF_{explicit}$  for AOD of 1.5.

(a). Note the different x-axes.

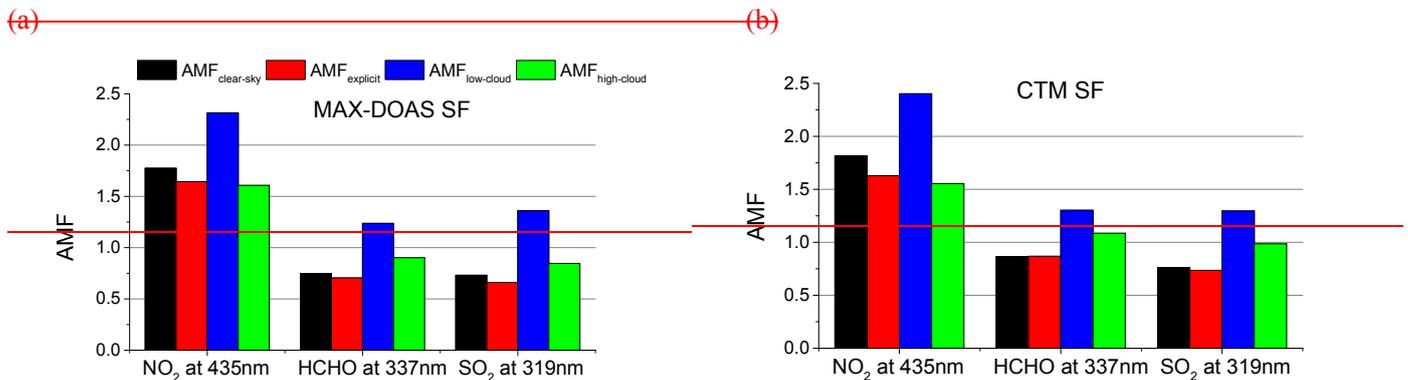
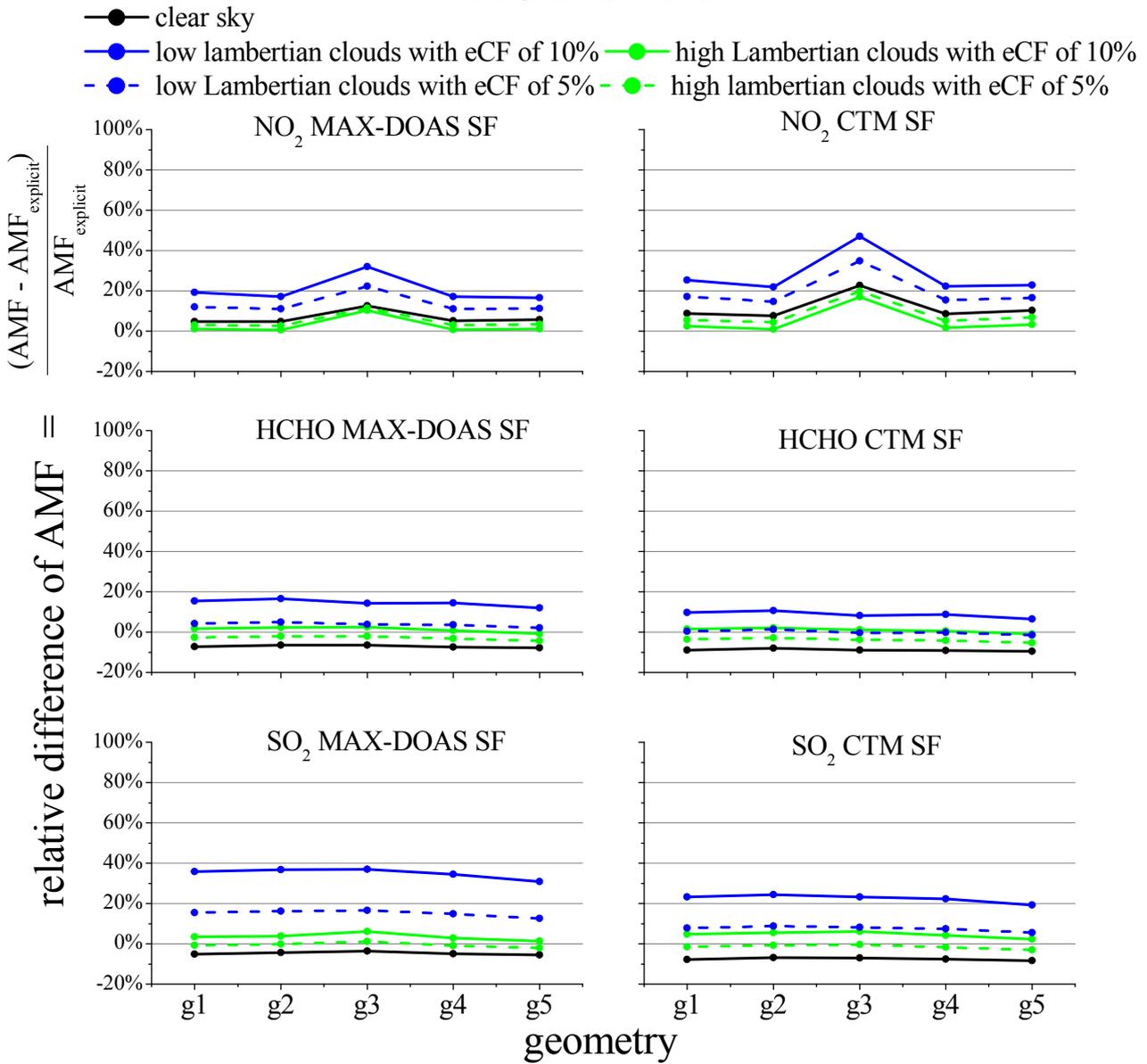


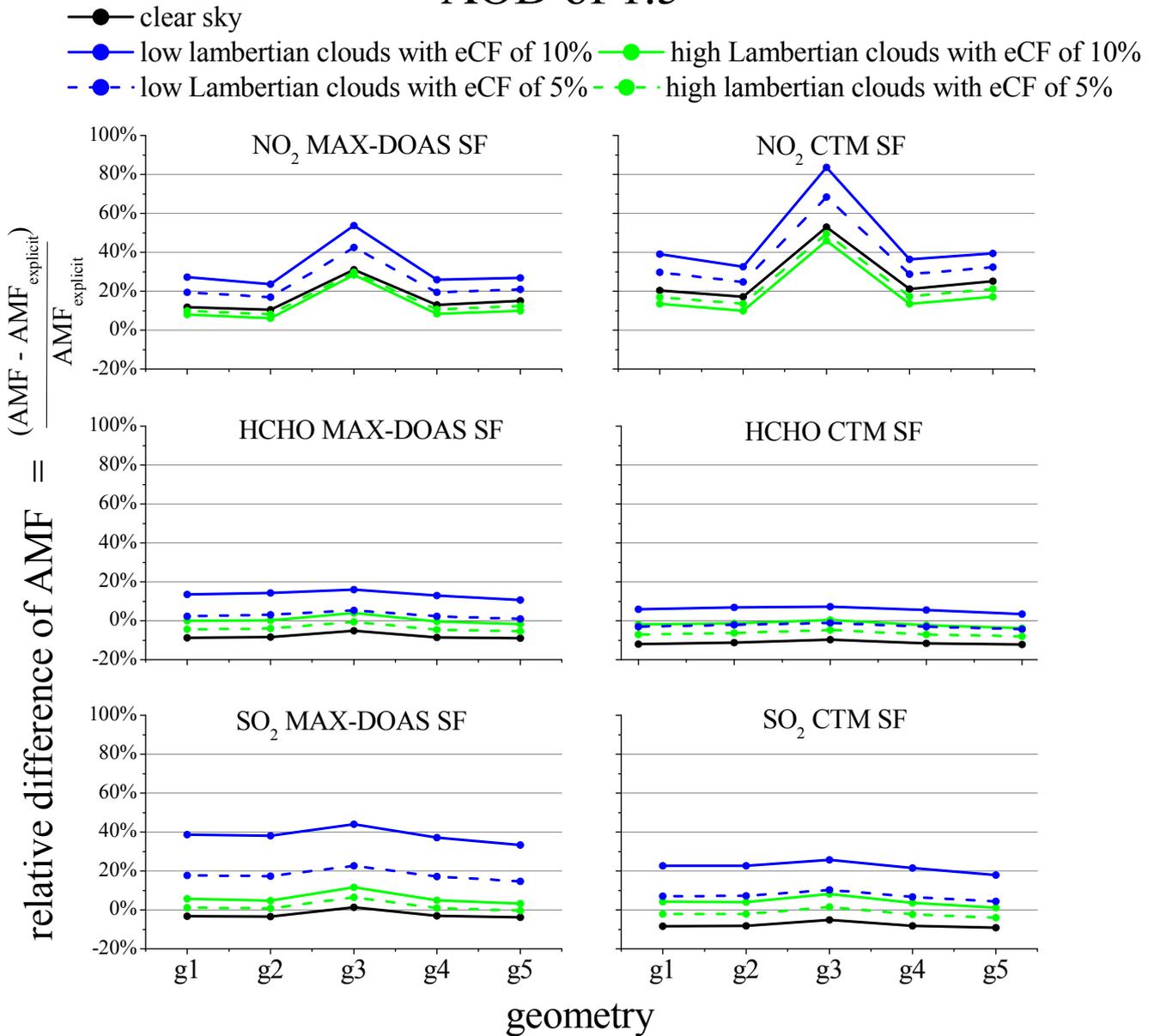
Figure 21: AMFs calculated for different aerosol and cloud assumptions (for details see text) and different trace gases. The TG SFs are obtained from MAX-DOAS (a) or CTM (b), see also Fig. S5 in the Supplement.

# AOD of 0.8



(b)

# AOD of 1.5



**Figure 17: Relative differences between AMFs calculated for different cloud assumptions (for detail see text) and AMFs with for explicit aerosol profiles for three trace gases. The labels at the x-axis indicate for five different observation geometries (see Table 2) for three trace gases. The MAX-DOAS and CTM SFs are used for the calculations shown in the left and right column. Explicit aerosol profiles of AOD of 0.8 and 1.5 are used in subfigure (a) and (b), respectively.**

Table 1 Mean ratios for the data presented in Fig. 4813.

	Ratio <sub>M-D</sub> (G-2A / OMI)	Ratio <sub>Sat</sub> (G-2A / OMI)	Ratio <sub>M-D</sub> (G-2B / OMI)	Ratio <sub>Sat</sub> (G-2B / OMI)
NO <sub>2</sub>	1.25	1.62	1.20	1.61
SO <sub>2</sub>	1.02	1.02	1.01	1.09
HCHO	0.78	0.88	0.76	0.87

Table 2 Daily averaged AODs derived from MAX-DOAS observations, eCFs and CTPs derived from OMI for six cloud free days with strong aerosol pollution.

date	AOD from MAX-DOAS	OMI eCF [%]	OMI CTP [hPa]
Jan 26, 2012	0.56	9	955
Oct 28, 2013	0.64	4	962
Dec 10, 2011	0.69	5	830
Nov 20, 2013	0.75	9	942
Apr 22, 2012	0.85	6	994
Nov 19, 2013	1.66	9	995

Table 2 Observation geometry scenarios for BAMF and AMF calculations with different aerosol and cloud assumptions.

Scenario	Solar zenith angle [°]	View zenith angle [°]	Relative azimuth angle [°]
g1	40	30	180
g2	10	30	180
g3	70	30	180
g4	40	0	180
g5	40	30	0

# **Supplement for “Validation of OMI, GOME-2A, and GOME-2B tropospheric NO<sub>2</sub>, SO<sub>2</sub>, and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products”**

Yang Wang<sup>1</sup>, Steffen Beirle<sup>1</sup>, Johannes Lampel<sup>1,2</sup>, Mariliza Koukouli<sup>3</sup>, Isabelle De Smedt<sup>4</sup>, Nicolas Theys<sup>4</sup>, Ang Li<sup>5</sup>, Dexia Wu<sup>5</sup>, Pinhua Xie<sup>5,6,7</sup>, Cheng Liu<sup>8,6,5</sup>, Michel Van Roozendael<sup>4</sup>, Trissevgeni Stavrakou<sup>4</sup>, Jean-François Müller<sup>4</sup>, and Thomas Wagner<sup>1</sup>

<sup>1</sup> Max Planck Institute for Chemistry, Mainz, Germany

<sup>2</sup> Institute of Environmental Physics, University of Heidelberg, Heidelberg, Germany

<sup>3</sup> Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>4</sup> Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

<sup>5</sup> Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China

<sup>6</sup> CAS Center for Excellence in Urban Atmospheric Environment, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China

<sup>7</sup> School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei, China

<sup>8</sup> School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China

*Correspondence to:* Yang Wang (y.wang@mpic.de), Ang Li (Angli@aiofm.ac.cn), Cheng Liu (chliu81@ustc.edu.cn)

## **1 Effects of variations of the coincidence criteria on the validation**

Because of the large ground pixel size of the satellite observations, MAX-DOAS results are averaged over a time period around the satellite overpass time to (partly) compensate the effect of horizontal gradients of the TG concentrations. In principle the time period is a function of the satellite pixel size, the wind speed and the life time of the trace gases. Although some factors change frequently, here we use one fixed time period for the long-term comparisons for simplicity. In this section, we test the effect on the satellite validation for four time periods including 1 hour, 2 hours, 3 hours and 4 hours around the satellite overpass time. Scatter plots of the average MAX-DOAS data over three time periods (1 hour, 3 hours and 4 hours) against those over 2 hours are shown in Fig. S1. The correlation coefficients are close to unity for all time periods. However, the slopes become systematically smaller for larger time periods (up to -10%) because of temporal smoothing. The results of the linear regressions and mean relative differences from the comparisons are also shown in Fig. S5a and will be discussed below together with the effect of the selected coincidence area of the satellite products.

In principle for the satellite validation the satellite pixel closest to the MAX-DOAS instrument need to be selected. However, in order to minimise the random noise of the satellite data, it is useful to calculate the average of several satellite observations close to the measurement site (see e.g. Irie et al., 2012 and Ma et al., 2013). As selection criterion, a distance between the centre of the satellite pixel and the measurement site can be specified. This optimum distance depends on many factors, such as the satellite ground pixel size, the selected time period over which the MAX-DOAS results are averaged, the expected horizontal gradients of the trace gas and uncertainties of the satellite data. A distance of < 20 km has been used for NO<sub>2</sub> comparisons (e.g. Ma et al., 2013 and Chan et al., 2015), 100 km for HCHO (De Smedt et al., 2015) and SO<sub>2</sub> (Theys et al., 2015). Irie et al. (2012) already found that the correlations and slopes of the linear regressions of the NO<sub>2</sub> tropospheric VCDs from OMI and GOME-2A against those from MAX-DOAS observations depend systematically on the distance to the MAX-DOAS station.

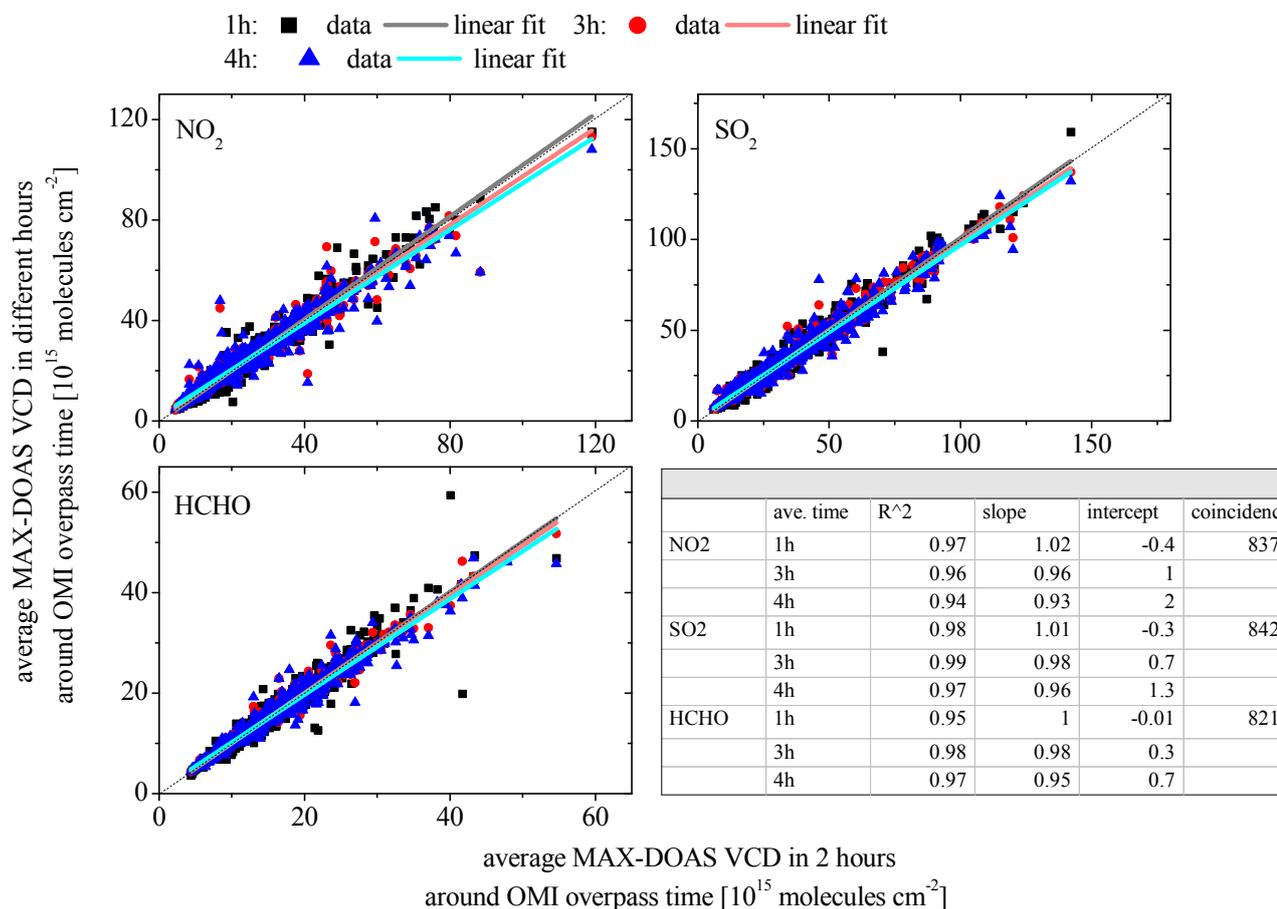
We test the effect of the variation of the distance between 10 km to 75 km on the comparison between the satellite data (OMI and GOME-2) and the MAX-DOAS data for all three TGs. The areas for the four radii (10km, 20km, 50km and 75km) and the pixel sizes

of OMI and GOME-2 are shown in the earth view image downloaded from the Google Earth service in Fig. S2a. For distances larger than 20 km, the cities of Suzhou, Changzhou, Huzhou and Nantong are included in the area. Because of transport of the pollutants between the cities and the different residence times, different horizontal distributions of the NO<sub>2</sub>, SO<sub>2</sub> and HCHO VCDs are found around Wuxi as shown in Figs. S2b, c and d, respectively. HCHO has a smoother distribution than SO<sub>2</sub>, which is smoother than NO<sub>2</sub>. The satellite data for pixels with the distances of 0-10km, 10-20km, 20-50km and 50-75km to the MAX-DOAS station are compared with the MAX-DOAS results.

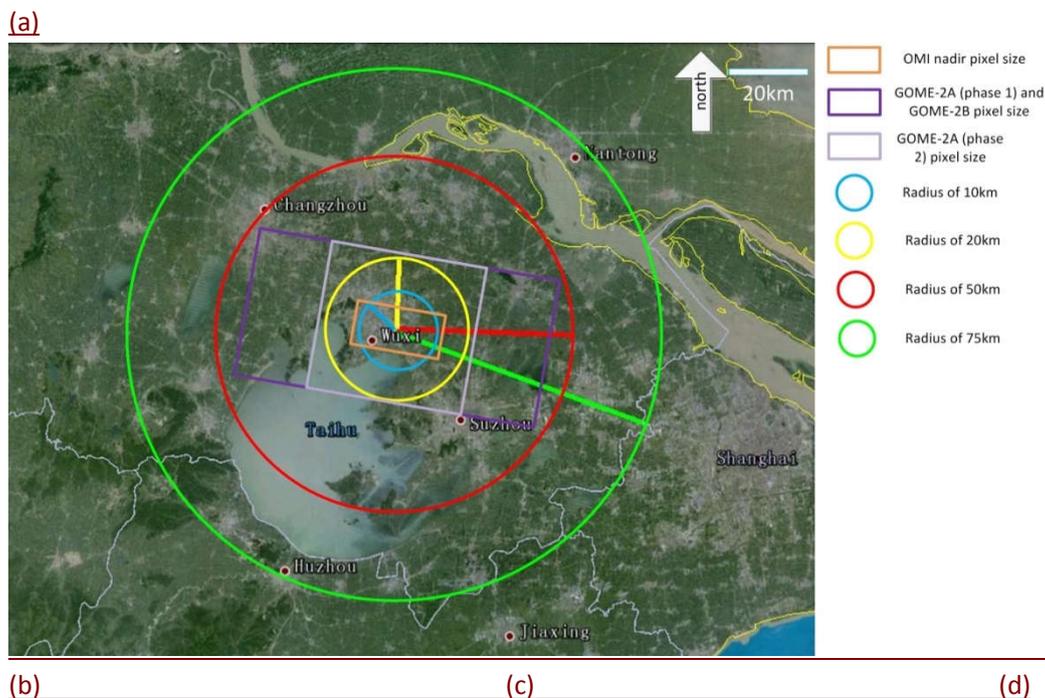
We compare both the results for individual satellite pixels and daily averages for the four radii with the average MAX-DOAS data over 2 hours around the satellite overpass time. The comparisons for OMI NO<sub>2</sub>, SO<sub>2</sub> and HCHO for pixels with distances of 0-10km, 10-20km, 20-50km and 50-75km are shown in Fig. S3a, b and c, respectively (the comparisons for pixels with the distances of <10km, <20km, <50km and <75km are shown in Fig. S4). We use the SO<sub>2</sub> OMI product from BIRA for this study, because it shows in general a higher correlation with the MAX-DOAS data. We found that the linear regressions for the daily averaged data are quite similar to those for the individual pixel data. Only the correlation coefficients are higher. The results of the linear regressions and the mean relative differences for the two distance categories as indicated in Fig. S3 and S4 are shown in Fig. S5 b and c, respectively. The slopes decrease with increasing distance for the three gases. The decrease of the slopes (from 0.75 to 0.49 and R<sup>2</sup> from 0.66 to 0.29) are stronger for NO<sub>2</sub> than for SO<sub>2</sub> and HCHO. This finding is consistent with the typically stronger horizontal inhomogeneity of NO<sub>2</sub>. The mean differences for HCHO show almost no dependence on the distance. This finding can be explained by the more homogenous distribution of HCHO compared to NO<sub>2</sub> and SO<sub>2</sub>. A significant decrease of the slopes from 0.73 to 0.50 and the R<sup>2</sup> from 0.65 to 0.44 is found for NO<sub>2</sub> with increasing distance over 20km. A decrease of the slope is also found for SO<sub>2</sub> for the distances larger than 20km. From these findings we conclude that 20km is a reasonable distance to select OMI NO<sub>2</sub> and SO<sub>2</sub> data for conditions similar to those at Wuxi. In contrast, for HCHO we select a distance of 50 km. Although for such distances the slope is smaller than for shorter distances, we find nearly identical mean differences. Because of this finding and the rather high noise of the HCHO satellite data we select a distance of 50 km, for which the number of available measurements largely increases. The comparison of Fig. S5a and b indicates that the effect of time periods used for averaging the MAX-DOAS results on the validation study is much smaller than the effect of distances for selecting the satellite data. Thus we apply the time period of 2 hours around the satellite overpass time in this study.

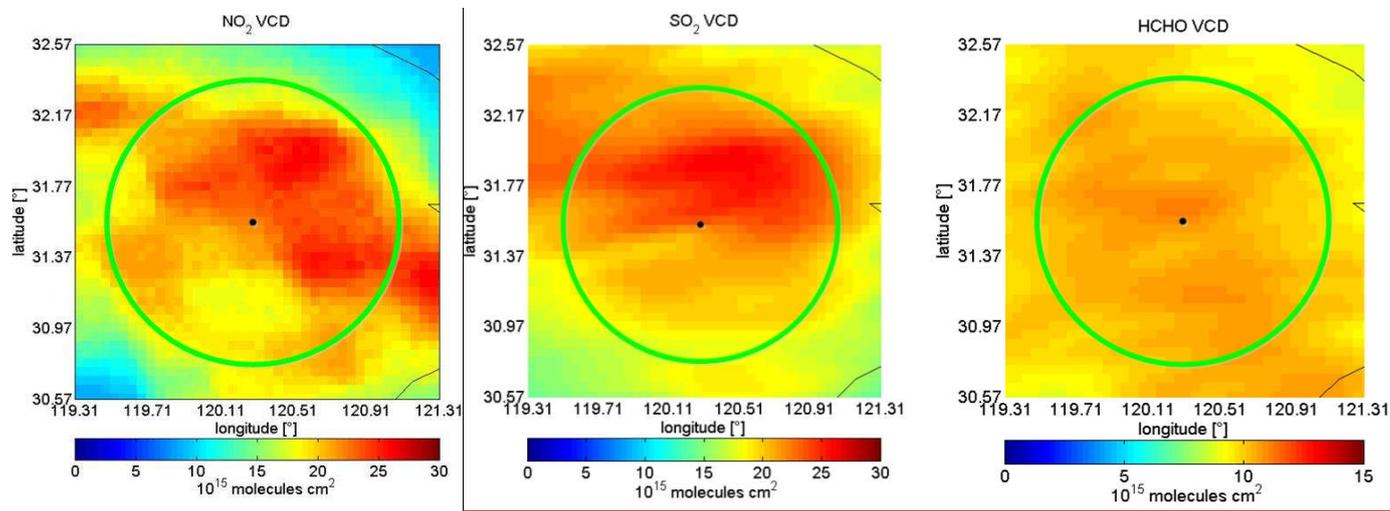
Similar results for GOME-2 data as those for OMI shown in Fig. S5 are shown in Fig. S6. The O3M-SAF GOME-2A SO<sub>2</sub> product from DLR is used for this sensitivity study. Also for the GOME-2 SO<sub>2</sub> data set the effect of the horizontal coincidence criterion is larger than the effect of the time period for the averaging of the MAX-DOAS data is found. Thus also 2 hours around the satellite overpass time will be used for GOME-2 comparisons in this study. The largest changes of the slopes for the three trace gases are found around the distance of 10km, but the results for the selection criterion of 0-10km should be treated with care because of the low number of available measurements. The changes of the slopes for distances larger than 20km are smaller than 0.06 for NO<sub>2</sub> and 0.04 for HCHO, but are larger for SO<sub>2</sub>. However, the results of the linear regressions for SO<sub>2</sub> should again be treated with care because of the rather low correlation coefficients. From these results we select 50km as a reasonable distance for GOME-2 data of NO<sub>2</sub>, SO<sub>2</sub> and HCHO.

In summary, in the validation studies (section 3) of the main manuscript, the MAX-DOAS results are selected within the period from 12:30 LT to 14:30 LT for the comparisons with OMI and from 08:30 LT to 10:30 LT for the comparisons with GOME-2A/B. The OMI NO<sub>2</sub> and SO<sub>2</sub> (HCHO) data are selected for satellite pixels with the distance of <20km (<50km) from the Wuxi station. The GOME-2A/B data of the three species are selected for the distances < 50km.

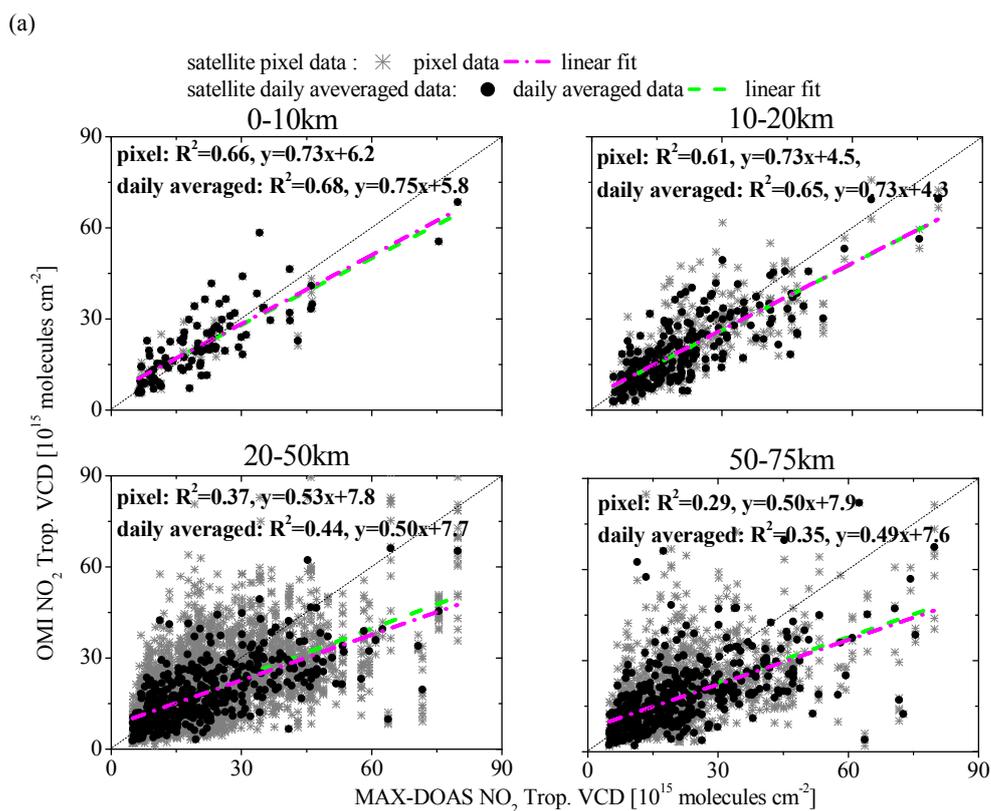


**Figure 3S1:** Averaged NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) tropospheric VCDs derived from MAX-DOAS observations in time periods of 1 hour (black dots), 3 hours (red dots) and 4 hours (blue dots) around the OMI overpass time plotted against those in the time period of 2 hours around the OMI overpass time. The linear regression lines for each time period and each species are plotted in each subfigure. The corresponding parameters are listed in the table.

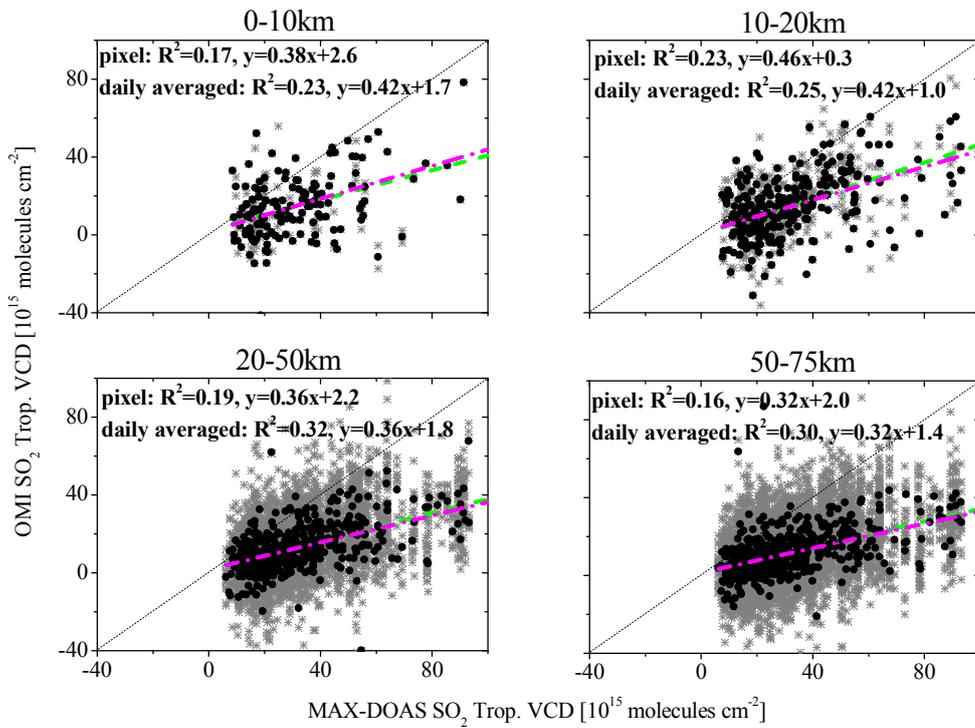




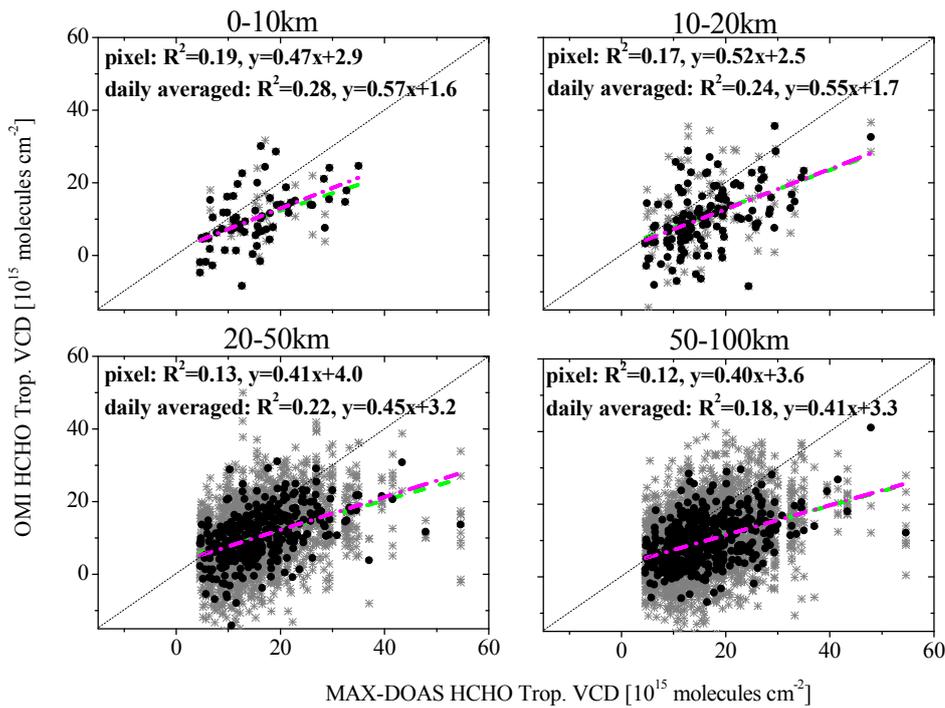
**Figure S2:** Subfigure (a) shows the earth image around Wuxi MAX-DOAS station from google earth service; the rectangles indicate the ground pixel sizes of the different satellite instruments used in this study. (GOME2-A phase 1 and phase 2 corresponding to the periods before and after 15 July 2013); the circles indicate areas with different radii around Wuxi. The subfigures of (b), (c) and (d) show averaged VCDs of  $\text{NO}_2$ ,  $\text{SO}_2$  and  $\text{HCHO}$  for the same area as shown in (b-1); the black dots indicate the location of Wuxi and the green circles have a radius of 75km.



(b)

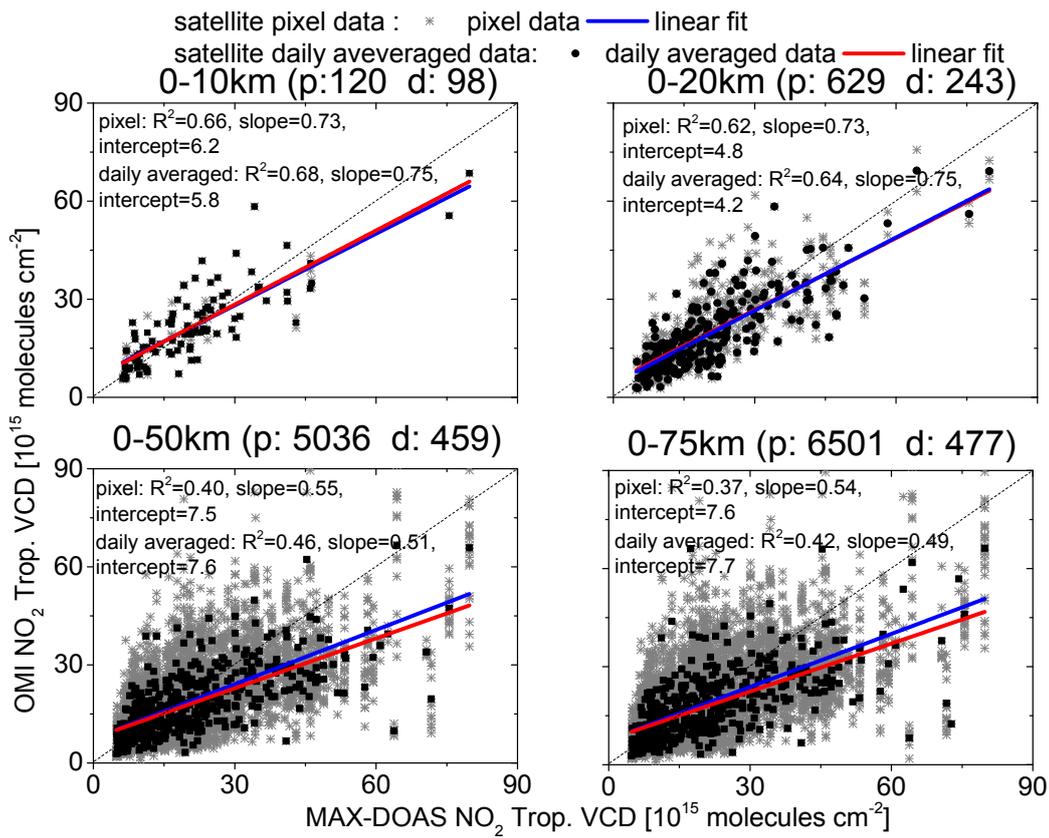


(c)

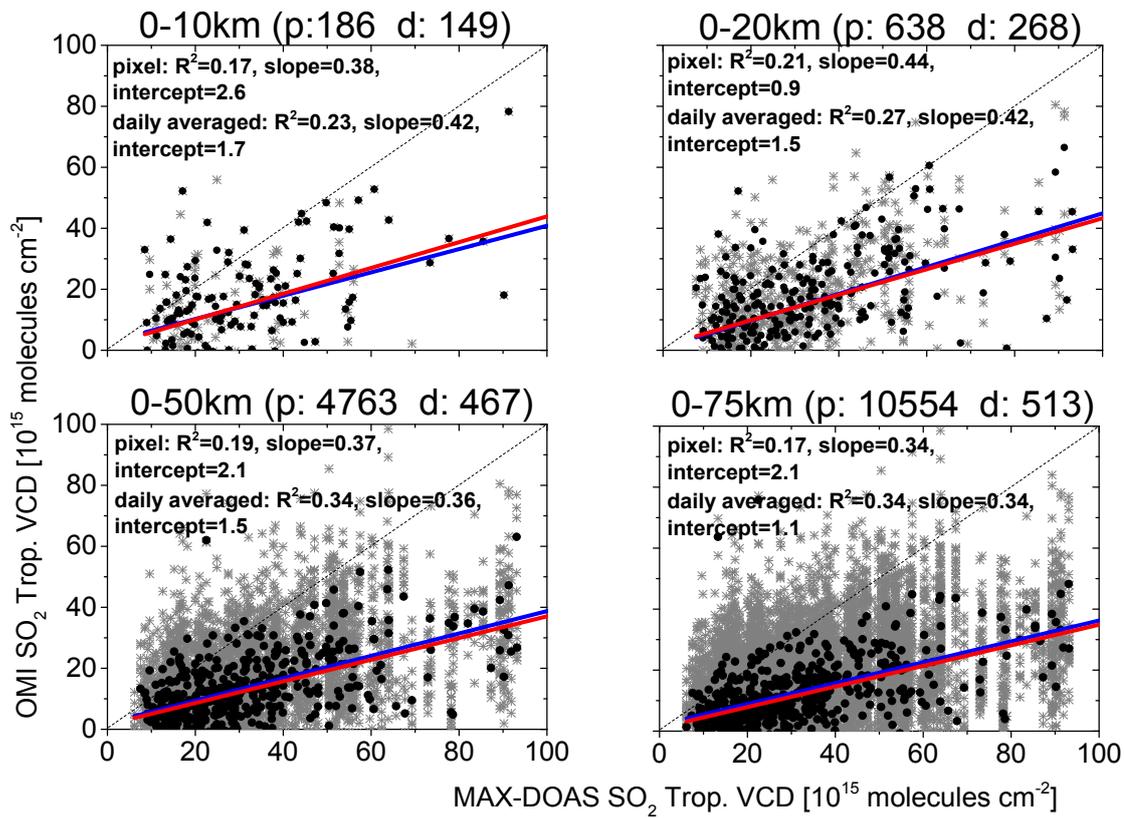


**Figure 4S3:** Tropospheric VCDs of NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) derived from OMI observations for pixels within the distance bins of 0-10km, 10-20km, 20-50km and 50-75km away from the Wuxi MAX-DOAS station plotted against the coincident MAX-DOAS results. Only OMI data for the eCF<30% are included. For HCHO, only the data for a fit error <  $7 \times 10^{15}$  molecules cm<sup>-2</sup> are included. The grey crosses and black dots show the data for individual satellite pixel and daily averaged data (averaged during two hours around the OMI overpass time), respectively. The linear regression lines and the parameters are shown in each subfigure for the pixel data (green dash lines) and daily averaged data (magenta dash-dot lines), respectively.

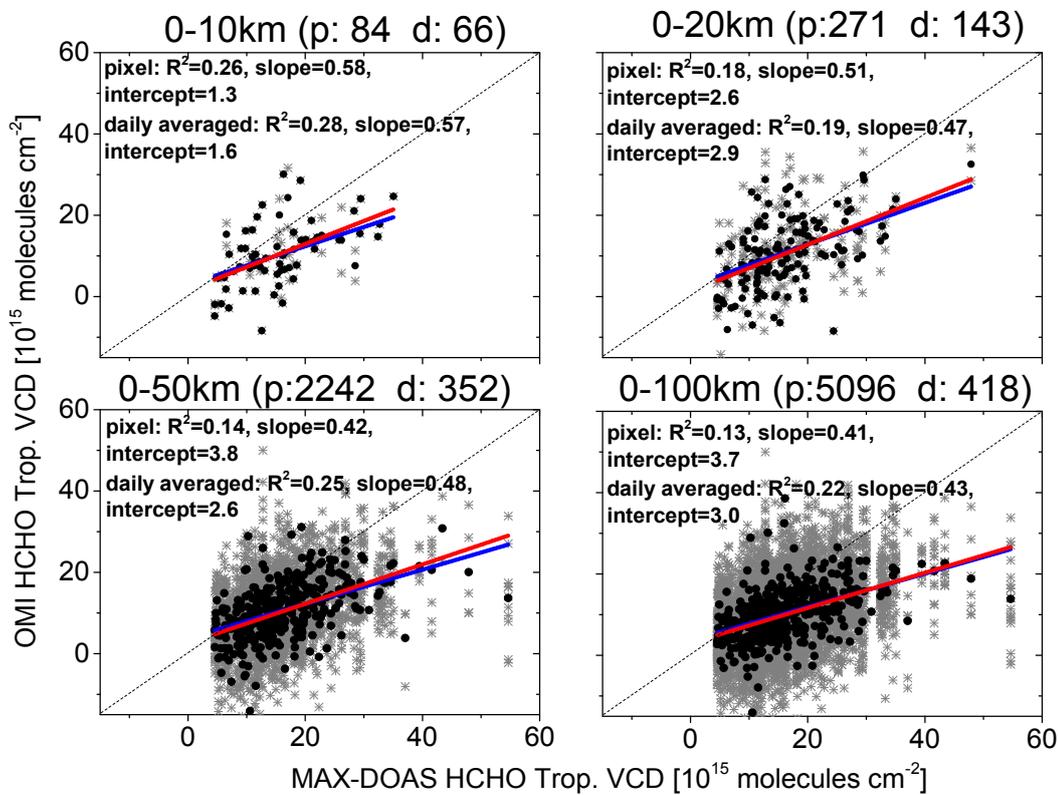
(a)



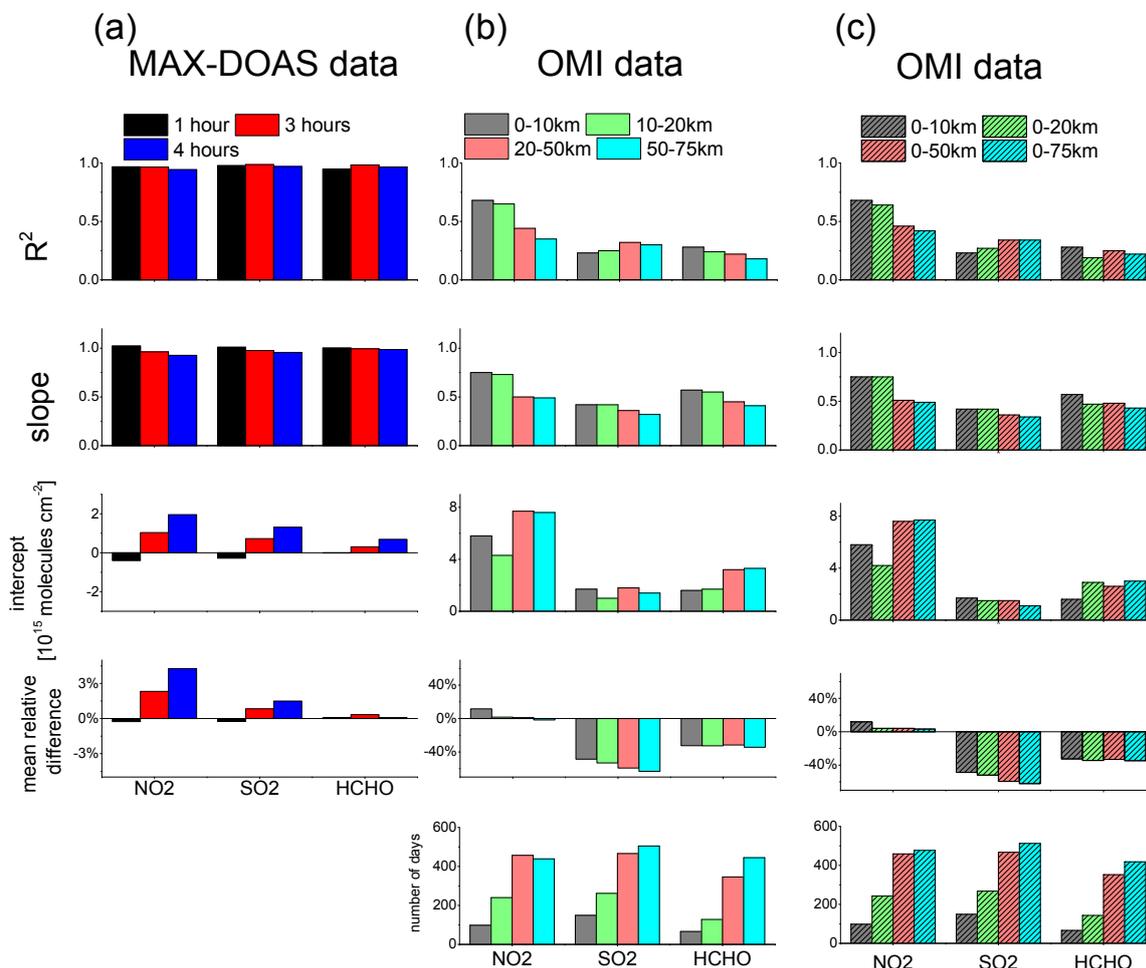
(b)



(c)



**Figure S1S4:** Tropospheric VCDs of NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) derived from OMI observations for the pixel with the distance bins of 0-10km, 0-20km, 0-50km and 0-75km away from the Wuxi MAX-DOAS station are plotted against the coincident MAX-DOAS results. Only the OMI data for the eCF<30% are included. For HCHO, only the data for VCD fit error < 7×10<sup>15</sup> molecules cm<sup>-2</sup> are included. The grey and black dots show the data for each satellite pixel and daily averaged data (averaged during two hours around the OMI overpass time), respectively. The corresponding numbers of the pixels (p) and days (d) are shown in each subfigure. The linear regression lines and the parameters are shown in each subfigure for the pixel data (blue lines) and daily averaged data (red lines), respectively.



**Figure S1S5:** (a)  $R^2$ , slope and intercept of the linear regressions as well as the mean relative differences of the averaged MAX-DOAS tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO in the time periods of 1 hour, 3 hours and 4 hours around the OMI overpass time compared to those in the time period of 2 hours. (b)  $R^2$ , slope and intercept of the linear regressions as well as the mean relative differences of the averaged OMI tropospheric VCDs of NO<sub>2</sub>, SO<sub>2</sub> and HCHO for the pixels within the distance bins of 0-10km, 10-20km, 20-50km and 50-75km compared to the coincident MAX-DOAS results. At the bottom also the numbers of the days for each comparison are shown. (c) Similar with (b), but for distance bins of 0-10km, 0-20km, 0-50km and 0-75km.

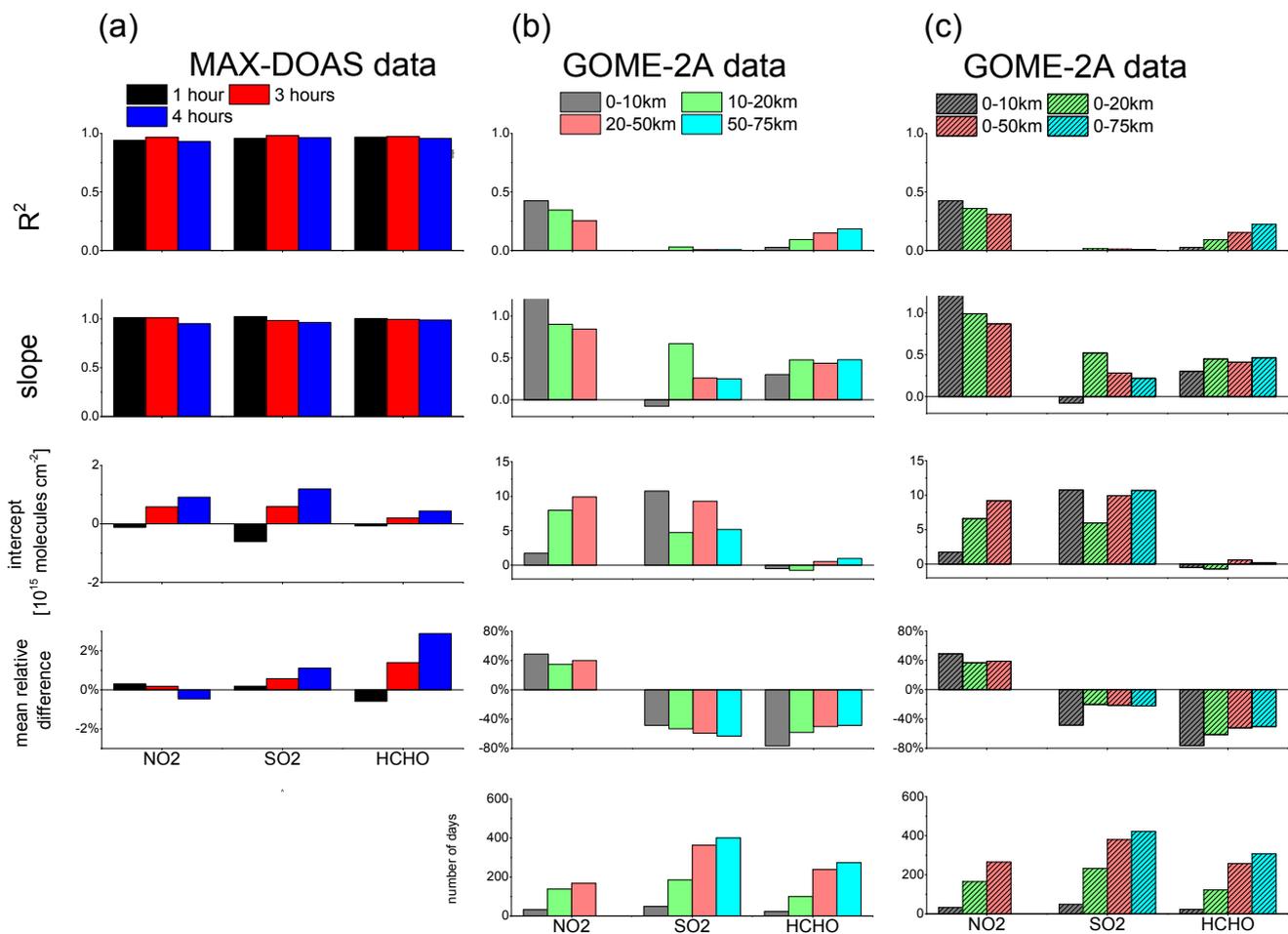
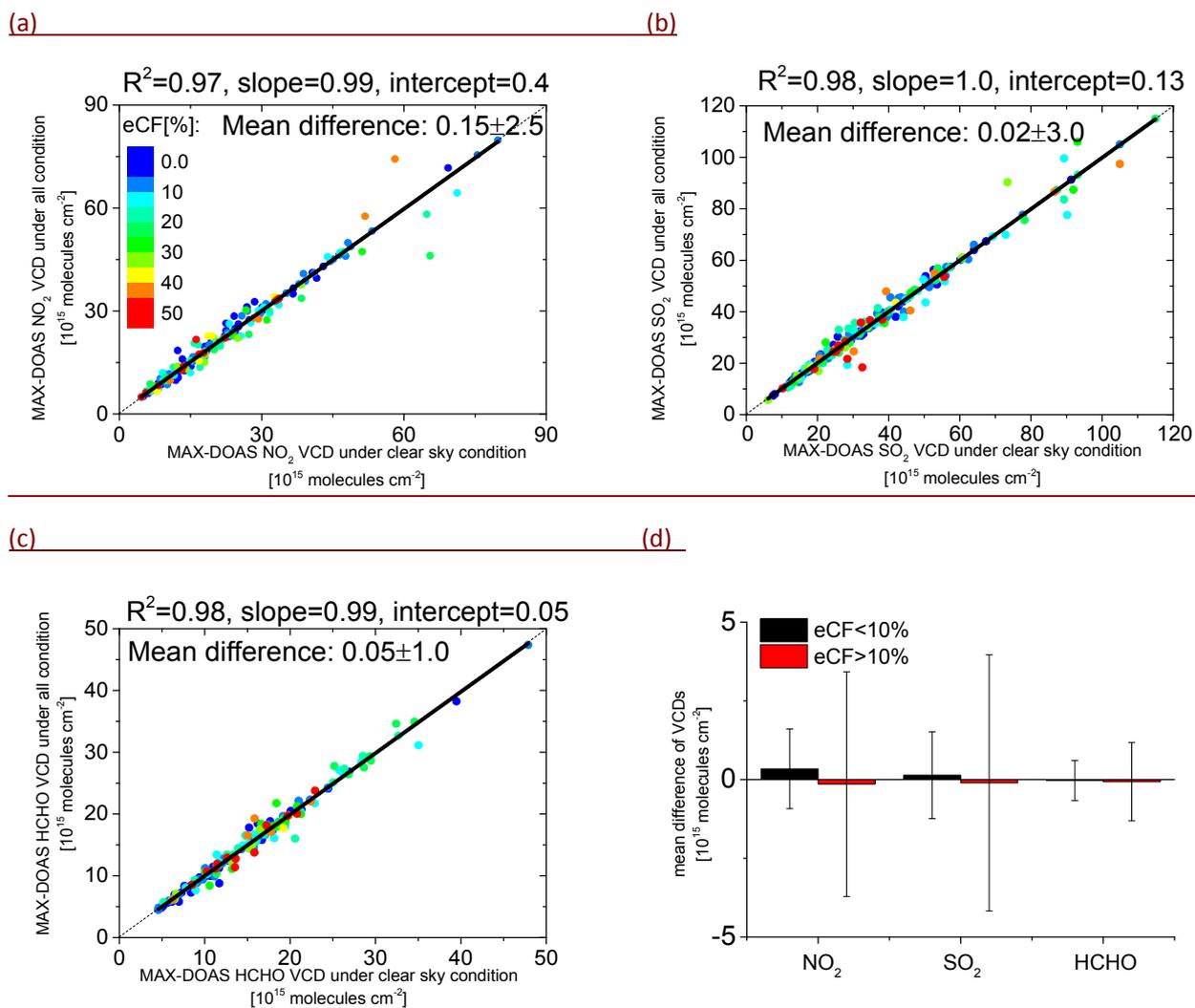


Figure 6S6: same as Fig. 5S5, but for GOME-2A data.

## 2 Cloud effect on MAX-DOAS tropospheric VCDs around the satellite overpass time

In the validation procedure the MAX-DOAS VCDs are averaged over a time period of  $\pm$ one hour around the satellite overpass time. Typically about ten MAX-DOAS elevation sequences are recorded during that period, during which the cloud conditions can change. This effect is probably most important for the presence of broken cloud cover. Thus in order to evaluate the cloud effect on MAX-DOAS results, we compare the average MAX-DOAS VCDs derived from all measurements in  $\pm$ 1 hour around the satellite overpass time with those from the measurements under clear sky conditions only. Sky conditions are derived from MAX-DOAS measurements (Wang et al., 2015). The OMI overpass time of 13:30 local time (LT) is selected for the investigation of this effect, and similar features are expected for observations around the GOME-2 overpass time. Fig. S7a, b and c show scatter plots and linear regressions of the average MAX-DOAS VCDs from all the measurements in  $\pm$ 1 hour around the satellite overpass time against those under clear sky conditions for  $\text{NO}_2$ ,  $\text{SO}_2$  and HCHO, respectively. Almost 1:1 linear regression lines and correlation coefficients ( $R^2$ ) (the Pearson's product moment correlation coefficient is applied in this paper) close to unity are found for all three species. To quantify the systematic differences of the TG VCDs, the corresponding mean differences (and standard deviations) are displayed in Fig. S7d for  $\text{eCF} < 10\%$  and  $\text{eCF} > 10\%$ , respectively. In general larger standard deviations are found for all three species for  $\text{eCF} > 10\%$ , indicating that larger deviations are related to larger eCF. Mean differences of  $0.15 \times 10^{15}$  molecules  $\text{cm}^{-2}$ ,  $0.02 \times 10^{15}$  molecules  $\text{cm}^{-2}$  and  $0.05 \times 10^{15}$  molecules  $\text{cm}^{-2}$  (corresponding to 0.8%, 0.05% and 0.4% of the average VCDs) are found for  $\text{NO}_2$ ,  $\text{SO}_2$  and HCHO, respectively, indicating that the cloud effect on MAX-DOAS results is probably negligible for the satellite validations. Here it should be noted that the shown comparison results represent only situations, for which clear and cloudy conditions occur during the two-hour period around the satellite overpass time. Thus we cannot rule out that the errors for measurements under continuous cloud cover are larger. However

situations of continuous cloud cover are not relevant for this validation study, because for such conditions no meaningful satellite results can be obtained.



**Figure S7: Daily averaged (during two hours around the OMI overpass time) NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) tropospheric VCDs derived from MAX-DOAS observations under all sky conditions plotted against those under clear sky conditions. The colours indicate the eCF. The correlation coefficients, slopes, intercepts and mean differences  $\pm$  standard deviation are displayed in each subfigure. The mean differences for eCF < 10% and > 10% are plotted in subfigure (d) with the error bars denoting the respective standard deviations.**

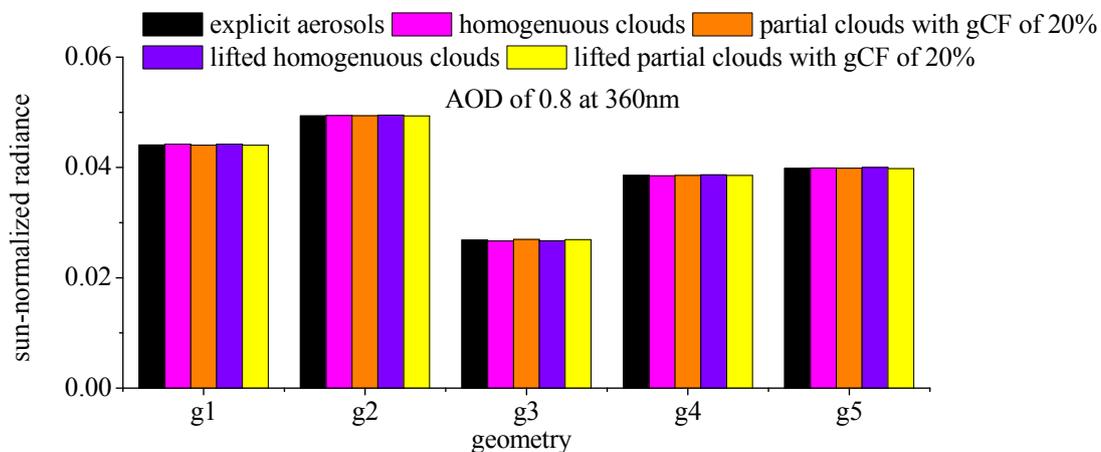
### 3. Comparisons of AMFs between for aerosols and for low clouds

Because some low clouds with a small eCF could interfere with the evaluation of aerosol effects in Fig. 14. In this section we performed McArtim RTM simulations to estimate the effect of low level clouds by comparing the TG AMFs between compared to TG AMFs with for aerosols and with the low clouds. In the simulation the aerosol properties are assumed same with those to be the same as in section 3.5 of the manuscript (two scenarios with either AOD of 0.8 or 1.5). The cloud properties are chosen to obtain the same radiance and O<sub>4</sub> SCDs (at 477nm) as for the aerosol scenarios, but the And the SSA is set to of 1 and the asymmetry parameter g of to 0.85 are assumed for clouds. The simulations are performed for Firstly the radiances and O<sub>4</sub> SCDs at 477nm are simulated for the two aerosol profiles with AOD of 0.8 and 1.5 for the five satellite observation geometries as listed in Table 2 in the main manuscript. Secondly the cloud optical depths can be derived based on the radiance. We derived chose two different cloud types: one is the a) a homogeneous clouds with low optical depth which homogeneously covering the entire a satellite pixel, and b); another an optically thick cloud covering only 20% of the is the partial clouds which partially cover a satellite pixel with geometrical cloud fraction (gCF) of 20%. It should be noted that the

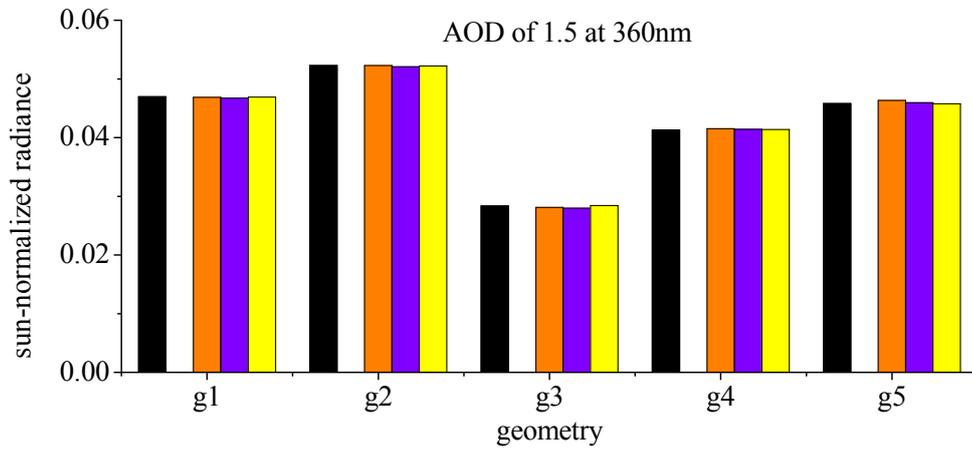
cloud extinction profiles can not unambiguously be determined based on the radiance and  $O_4$  SCDs only. Thus we assumed two different types of clouds, which represent the most extreme cases: Afterwards, the cloud height can be derived from the  $O_4$  SCDs. However only effective cloud height can be derived, the real cloud top and bottom height can not be derived. Therefore we define the cloud extinction profile using two methods: one case is the 'near-surface clouds' with a constant extinction starting from surface. For this cloud type the cloud top height can be derived based on  $O_4$  SCDs; another cloud type is the 'lifted clouds' with the fixed vertical extension of 400 meters (constant extinction in the cloud layer). For this cloud type the height of the cloud middle center can be derived from the  $O_4$  SCDs. Thus, Overall in total all simulations are performed for four types of clouds, which are referred to as 'near-surface homogeneous clouds', 'near-surface partial clouds', 'lifted homogeneous clouds', and 'lifted partial clouds', are derived from the radiance and  $O_4$  SCDs corresponding to the each aerosol profile and observation geometry. The comparisons of the corresponding sun-normalized radiances and  $O_4$  SCDs between with the aerosol profiles and with the derived cloud profiles are shown in Fig. S8 and S9, respectively. The consistent radiances and  $O_4$  SCDs between the aerosols and clouds can be seen. The derived cloud extinction profiles are shown in Fig. S10. Here one aspect needs to be clarified it should be noted that the simulations for the near-surface homogeneous clouds can not match the  $O_4$  SCD with respect to the derived for the aerosols with AOD of 1.5. Finally the AMFs with for the derived clouds are compared with those with for the corresponding aerosols for the different observation geometries. The corresponding relative differences of AMFs are shown in Fig. S11. In general the differences are smaller than 10% for  $NO_2$ , and 5% for  $SO_2$  and HCHO, except for the g3 geometry, for which the  $NO_2$  AMF difference amounts up to 20%. Thus we conclude that in general the influence of residual clouds on the satellite TG retrievals is of importance. Here it should be noted that especially over polluted regions situations with  $eCF < 10\%$  and  $CTP > 900$  hPa typically represent cases with aerosol pollution. If in addition residual clouds are present, they typically co-exist with high aerosol amounts. Thus our simulation results (based on pure cloudy cases without aerosols) represent not typical but rather extreme cases.

Therefore the effects of the low clouds in Fig. 14 of the main manuscript are not significant.

-(a)

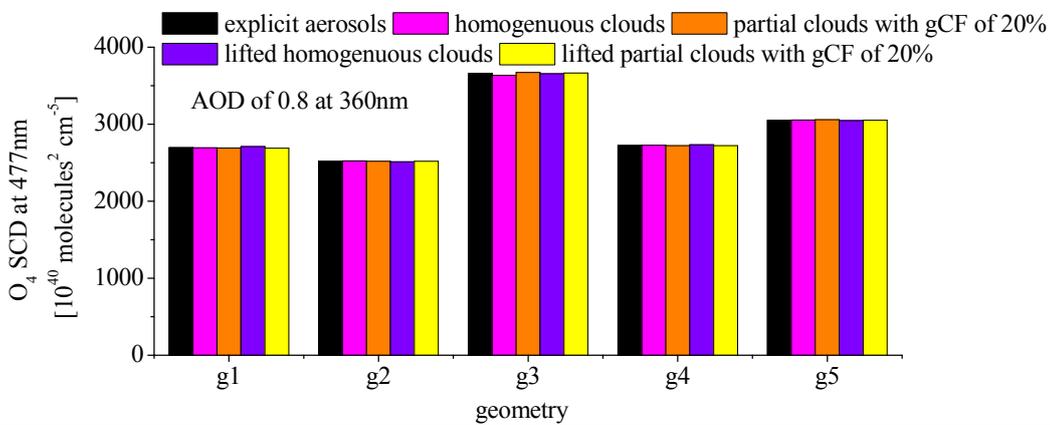


(b)

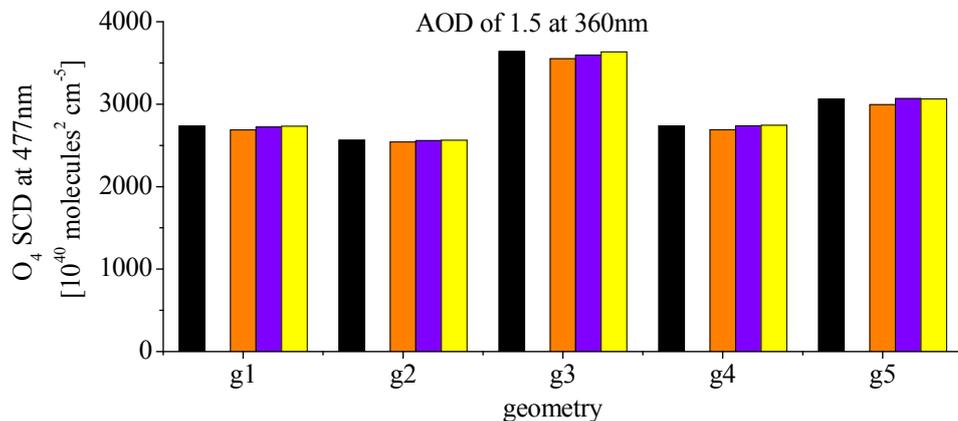


**Figure S8: sun-normalized radiances simulated by RTM at 477nm with for the two aerosol cases and the different clouds scenarios shown in Fig. S10. The different labels at the x-axis indicate for different observation geometries (see Table 2 in the main text). The radiances for the explicit aerosols of AOD of 0.8 and corresponding clouds are shown in (a) and those for AOD of 1.5 are shown in (b).**

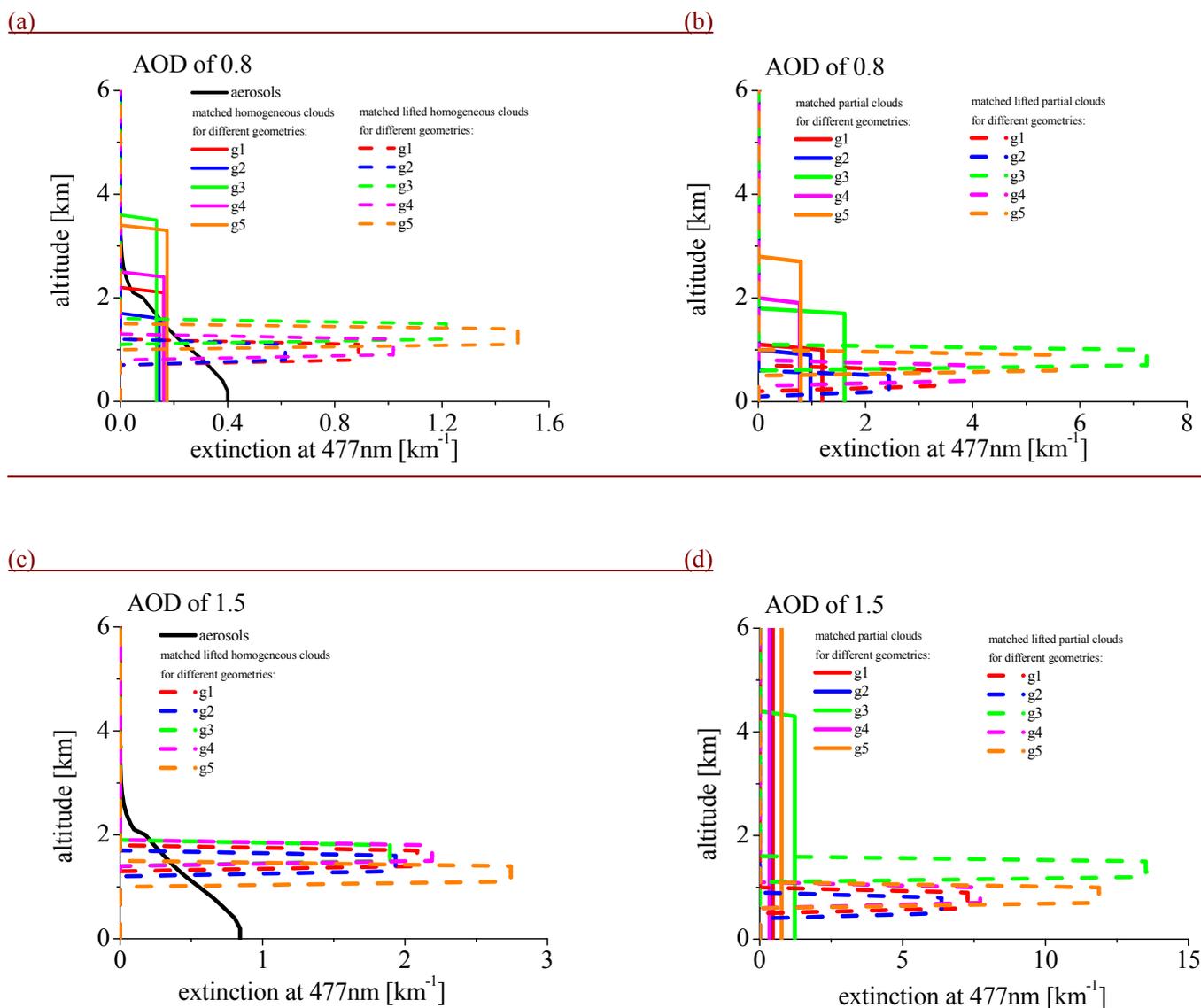
(a)



(b)

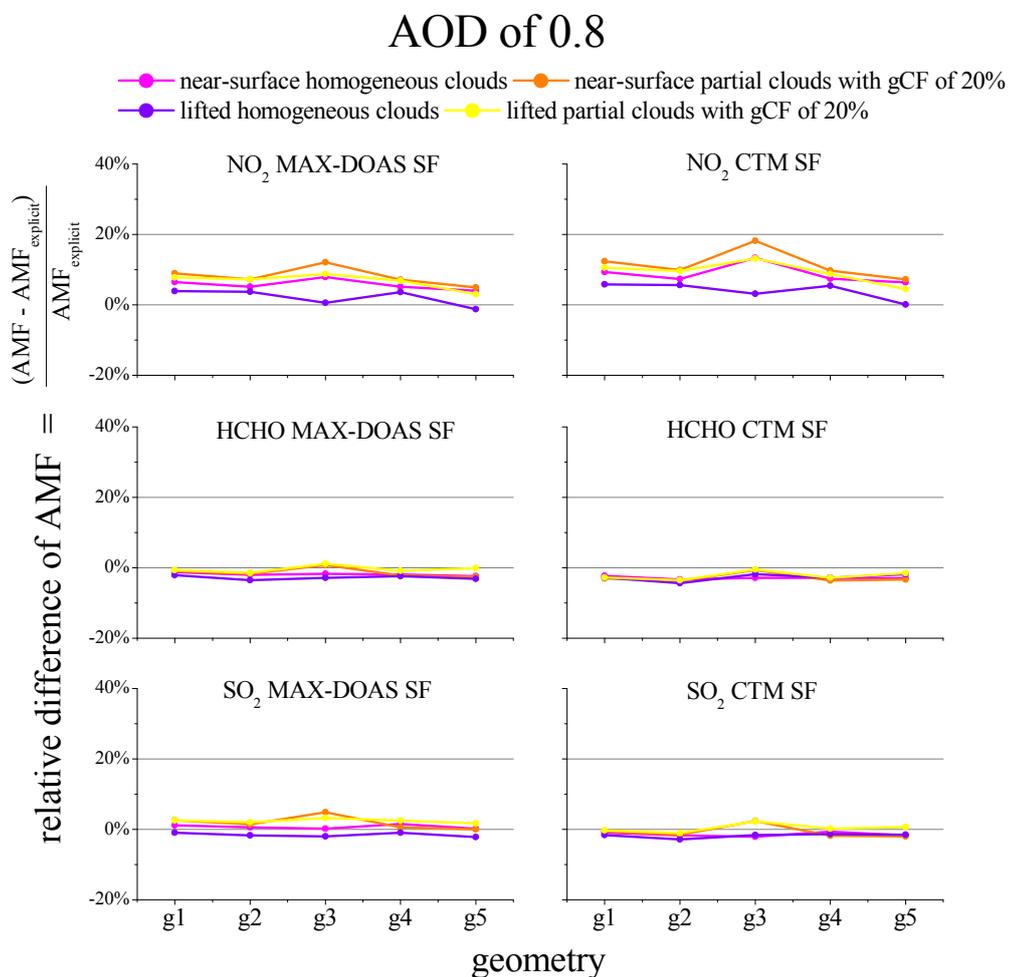


**Figure S9: O<sub>4</sub> SCDs simulated by RFM at 477nm for the two aerosol cases and the different clouds scenarios shown in Fig. S107 with the aerosols and clouds shown in Fig. S17. The different labels at the x-axis indicate different observation geometries for different observation geometries (see Table 2 in the main text). The radiances for the explicit aerosols of AOD of 0.8 and corresponding clouds are shown in (a) and those for AOD of 1.5 are shown in (b).**



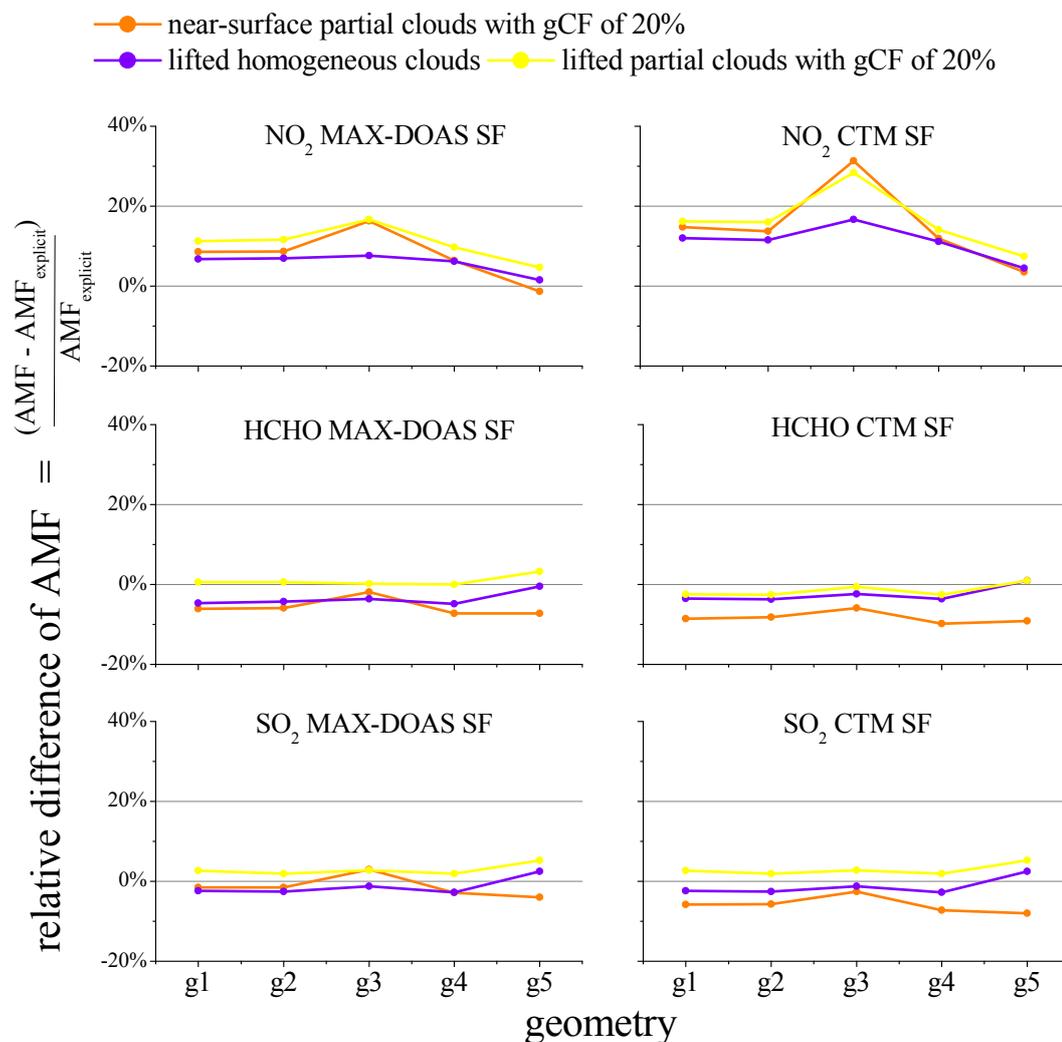
**Figure S10: Derived cloud extinction profiles for the four cloud types and viewing geometries (with SSA of 1 and  $g$  of 0.85) profiles at 477nm with which RTM simulates radiances and  $O_4$  SCDs matching those with the explicit aerosol profiles (The black curves indicate the aerosol extinction profiles, the coloured lines the cloud extinction profiles.) are denoted by the colourful curves. The homogeneous clouds and partial clouds with geometrical cloud fraction (gCF) of 20% (see text) are shown in (a) and (b). The solid and dashed curves indicate the near-surface and lifted clouds, respectively. (a) and (b) are for the aerosol of AOD of 0.8; (c) and (d) are for AOD of 1.5. The different colours denote different observation geometries (see Table 2 in the main text).**

(a)



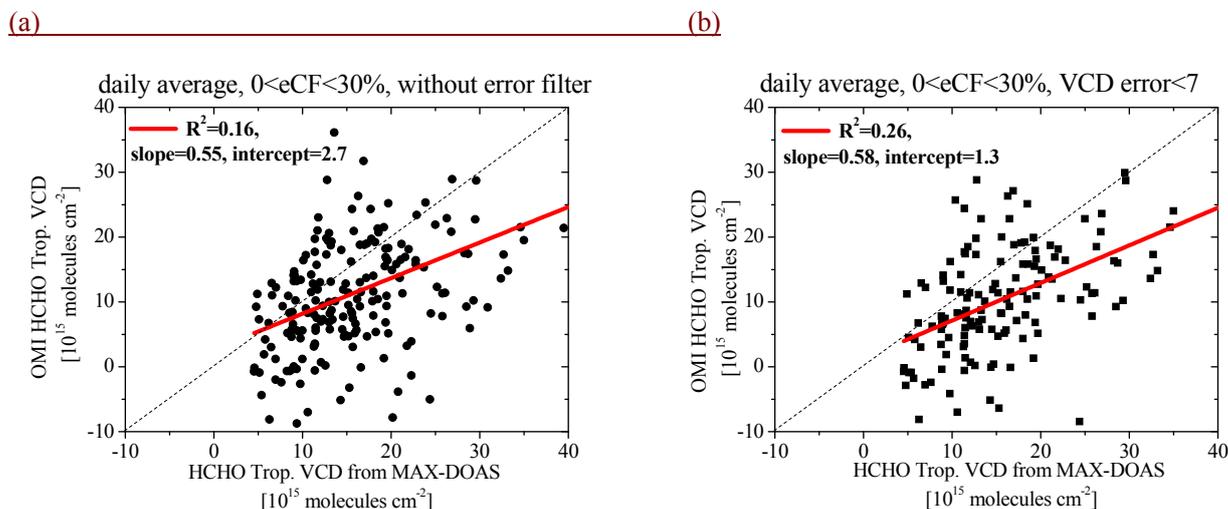
(b)

# AOD of 1.5



**Figure S11: Relative differences of the three TG AMFs between for derived the four cloud types of low clouds and for the corresponding aerosol profiles. The different labels at the x-axis indicate different observation geometries for five observation geometries (see Table 2). for three trace gases. The MAX-DOAS and CTM SFs are used for the calculations shown in the left and right columns. The aerosol profiles of AOD of 0.8 and 1.5 are used in subfigure (a) and (b), respectively.**

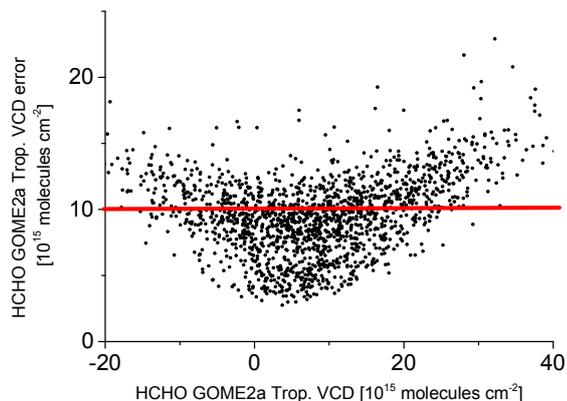
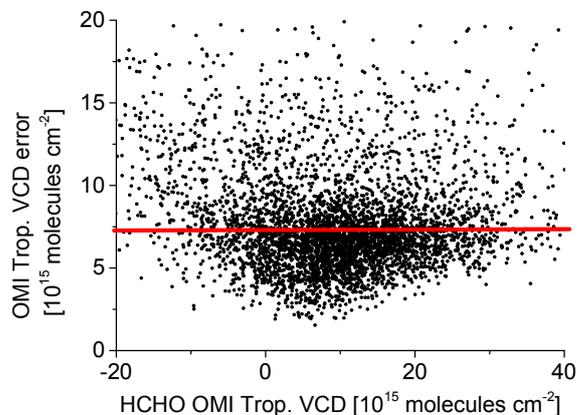
## 4 Other figures



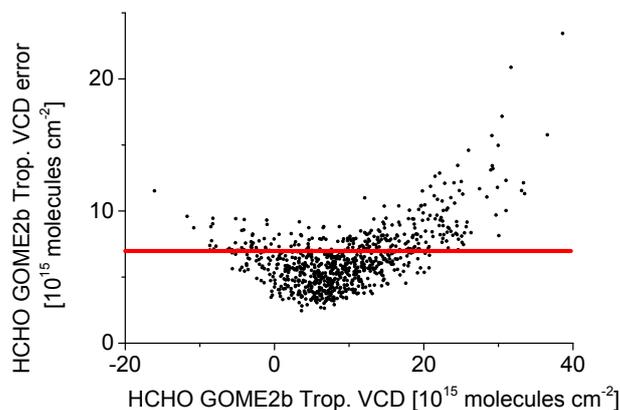
**Figure S12: daily averaged HCHO tropospheric VCD derived from OMI observations are plotted against those derived from MAX-DOAS observations for eCF<30%. And linear regressions are also shown. The OMI data before and after the filter of VCD fit error <math>7 \times 10^{15}</math> molecules  $\text{cm}^{-2}</math> are plotted in subplot (a) and (b), respectively.$**

**(a) OMI**

**(b) GOME-2A**



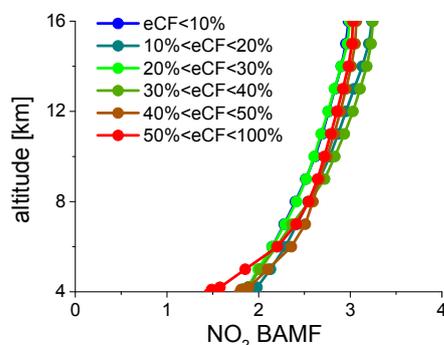
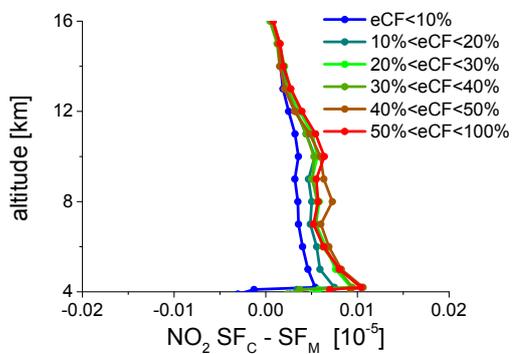
**(c) GOME-2B**



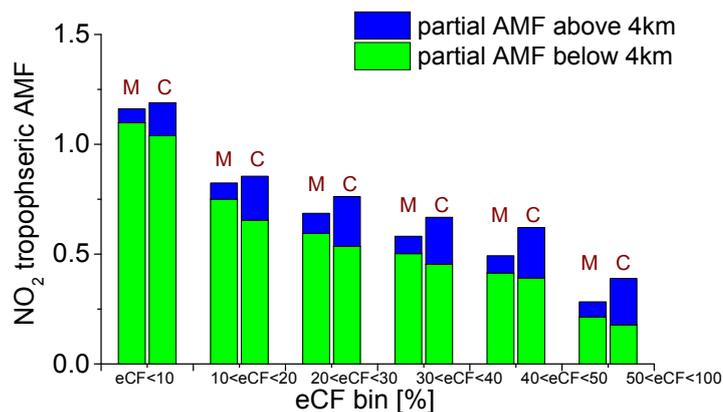
**Figure S13: HCHO tropospheric VCDs for OMI pixels for eCF<30% are plotted against those derived from MAX-DOAS observations with the color map of eCF; the linear regression parameters are acquired for eCF<30% and for eCF<10%, respectively. (b) Scattered plots are same as in (a)**

**(a)**

**(b)**



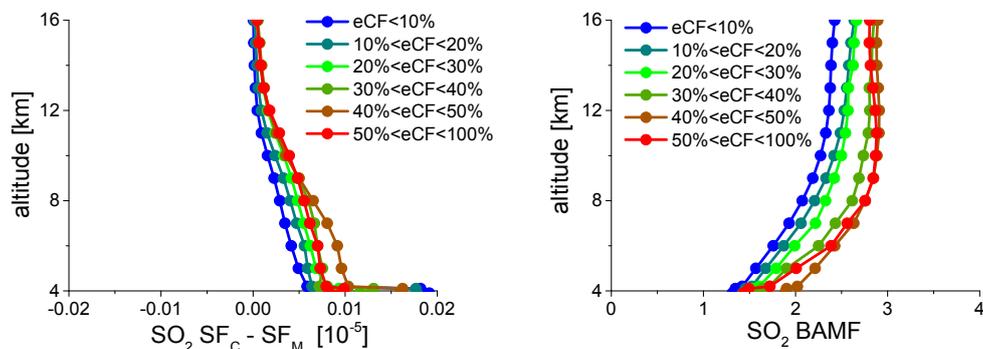
(c)



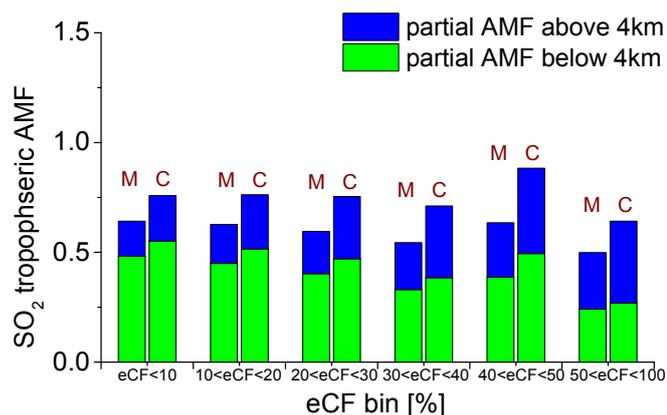
**Figure S14: (a) Averaged difference between the  $\text{NO}_2$  SF from CTM ( $\text{SF}_C$ ) and from MAX-DOAS ( $\text{SF}_M$ ) for different eCF bins in the altitude range of 4km to 16km. (b) Averaged  $\text{NO}_2$  BAMF for satellite observation for different eCF bins in the altitude range of 4km to 16km. (c)  $\text{NO}_2$  tropospheric AMFs calculated with averaged  $\text{SF}_M$  (marked by “M”) and  $\text{SF}_C$  (marked by “C”), respectively; the partial AMFs below and above 4km are marked by green and blue columns, respectively.**

(a)

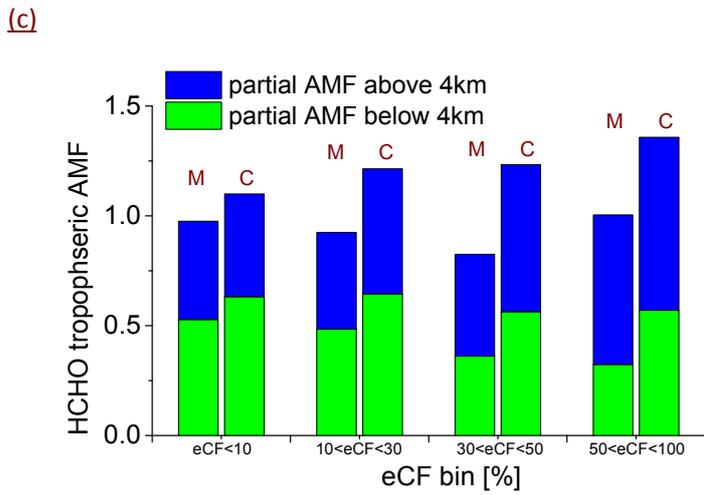
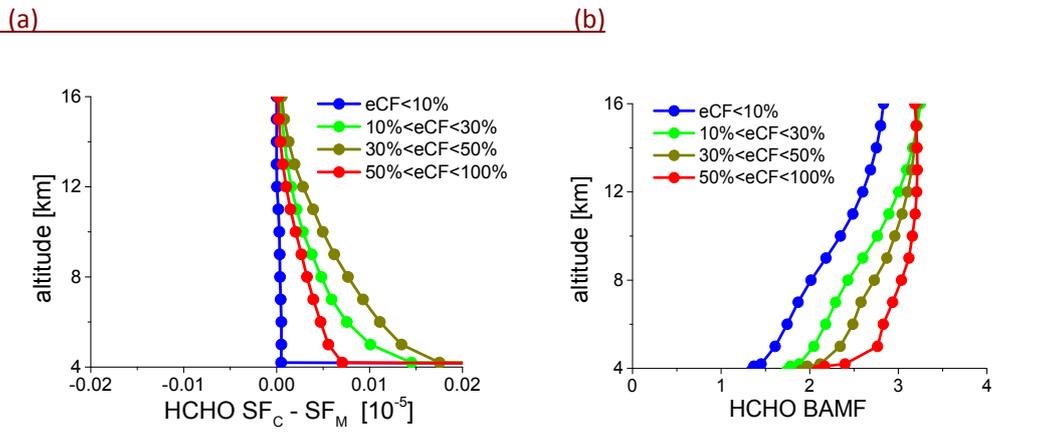
(b)



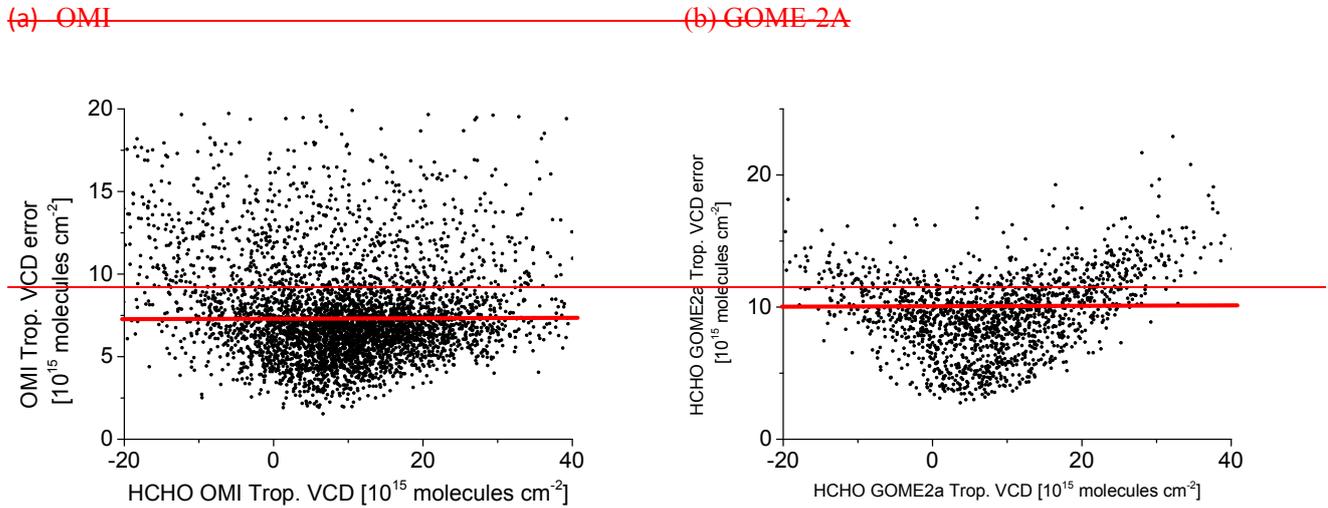
(c)



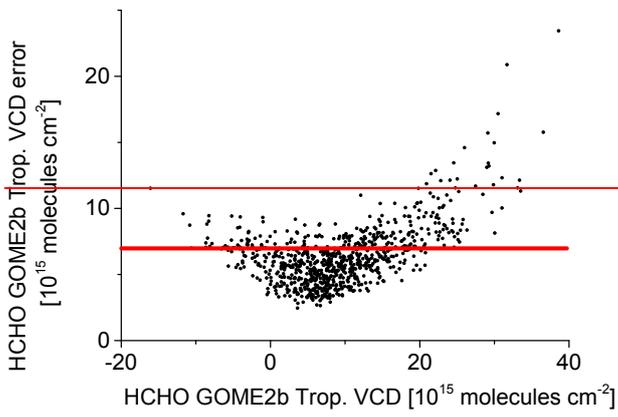
**Figure S15: Same as Fig. S14, but for  $\text{SO}_2$ .**



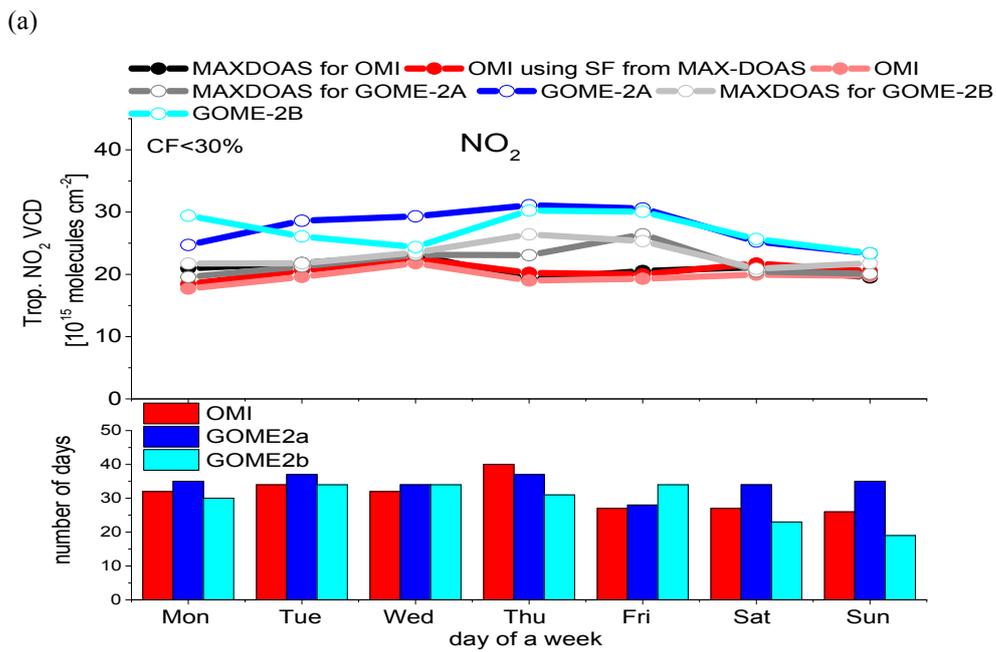
**Figure S16: Same as Fig. S14, but for HCHO.**



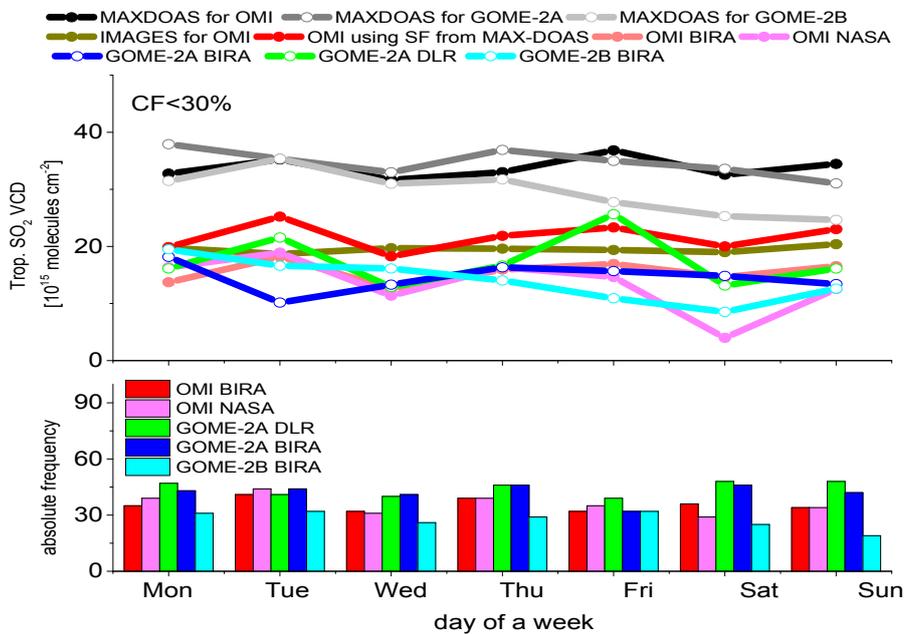
(e) GOME-2B



**Figure S2: HCHO tropospheric VCDs for OMI pixels for eCF<30% are plotted against those derived from MAX-DOAS observations with the color map of eCF; the linear regression parameters are acquired for eCF<30% and for eCF<10%, respectively. (b) Scattered plots are same as in (a)**



(b)



(c)

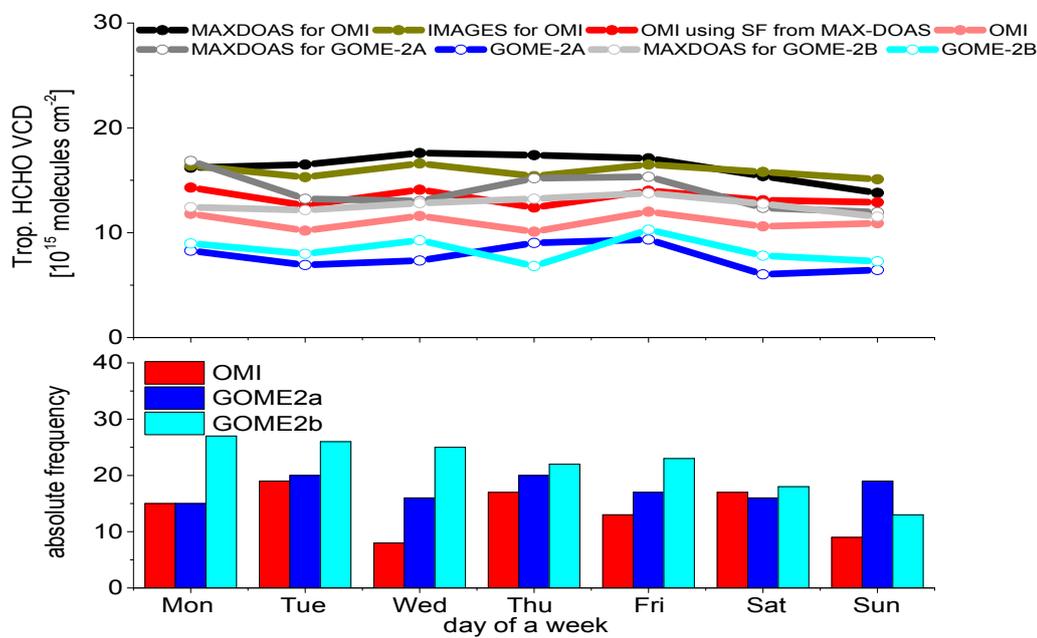
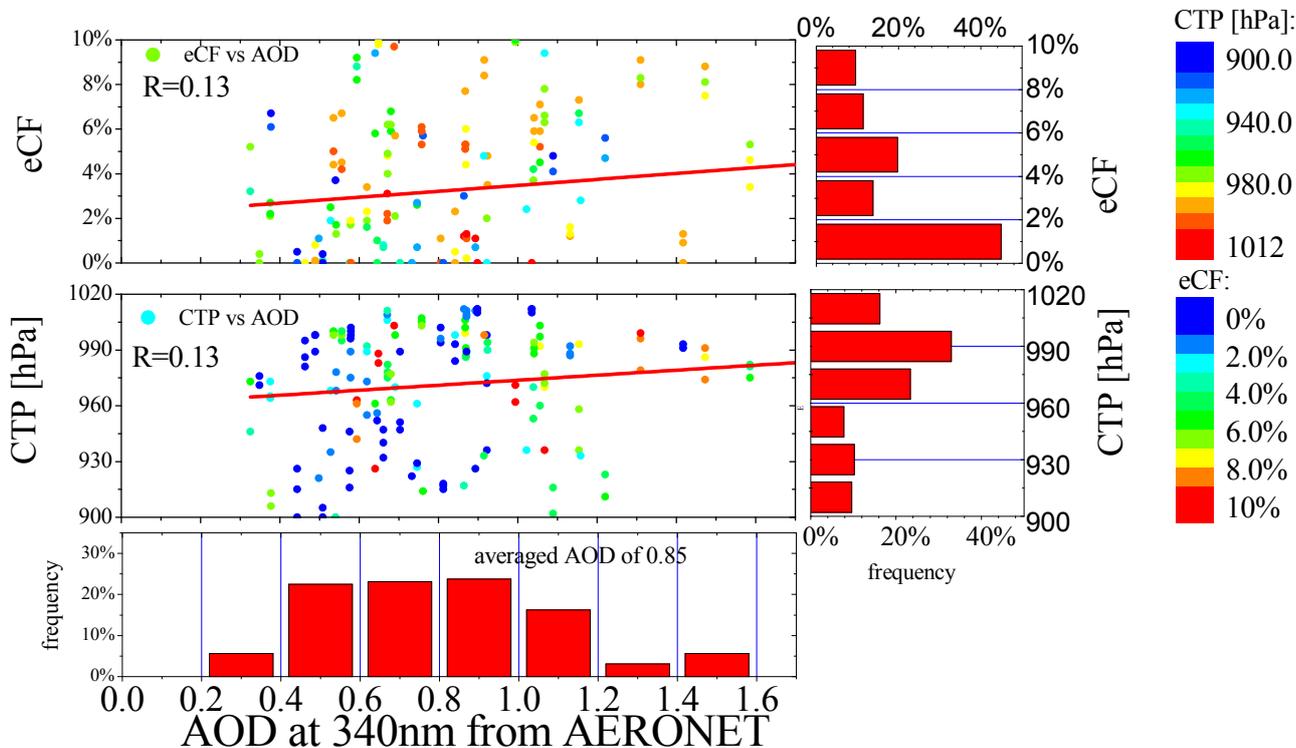
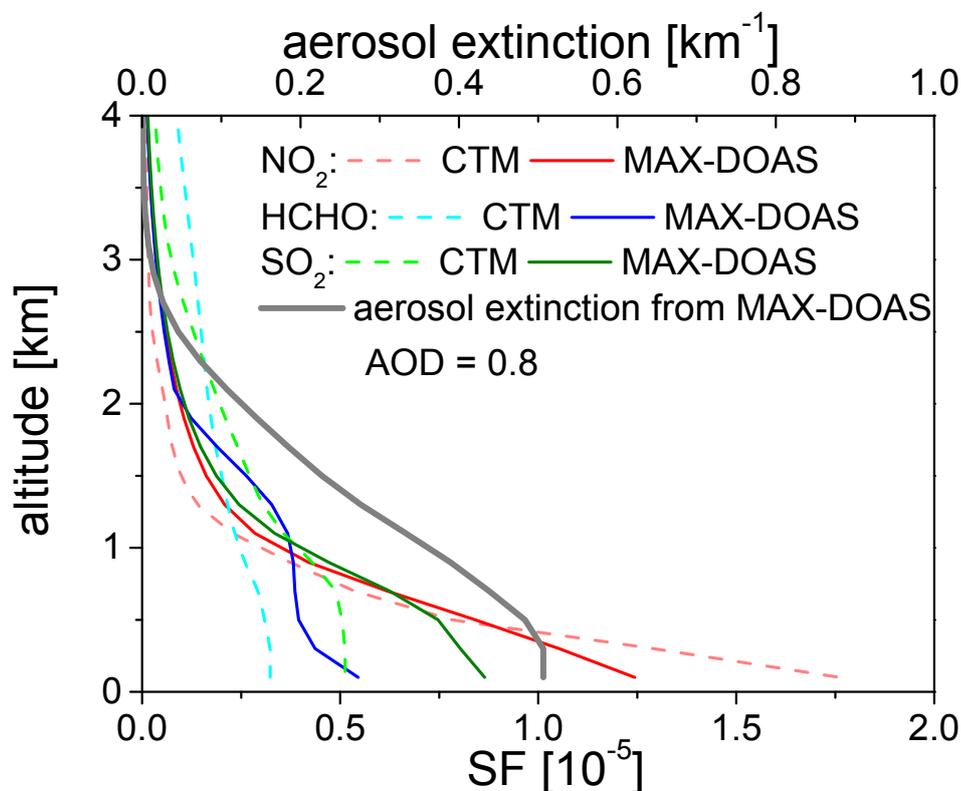


Figure S3S17: For eCF < 30%, weekly cycles of VCDs of NO<sub>2</sub> (a), SO<sub>2</sub> (b) and HCHO (c) derived from different satellite instruments, corresponding coincident MAX-DOAS measurements. In all the subfigures the red and light red lines indicate the improved OMI tropospheric VCDs using the SFs from MAX-DOAS and the original VCDs from OMI products, respectively. The numbers of the available days in each two-month bin from different satellite products are shown in the bottom of each subfigure.



**Figure S18: eCF and CTP from individual OMI observations are plotted against AOD around 360nm derived from AERONET Taihu station during the whole measurement in the condition of eCF<10% and CTP > 900hPa. The red bars on the right and bottom indicate the frequency of eCF, CTP, and AOD in different value intervals. The red lines are the linear regressions of the scatter plots. The correlation coefficients are shown in the plots. The color map in (a) and (b) indicate CTP and eCF, respectively.**



**Figure S19: Averaged aerosol extinction profiles and SF of NO<sub>2</sub>, SO<sub>2</sub> and HCHO derived from all MAX-DOAS measurements under cloud-free sky conditions. The dashed curves indicate the corresponding averaged SF derived from CTM simulations for NO<sub>2</sub> (TM4), SO<sub>2</sub> (IMAGES) and HCHO (IMAGES). (a)**

~~Jan. 26, 2012~~

~~Oct. 28, 2013~~

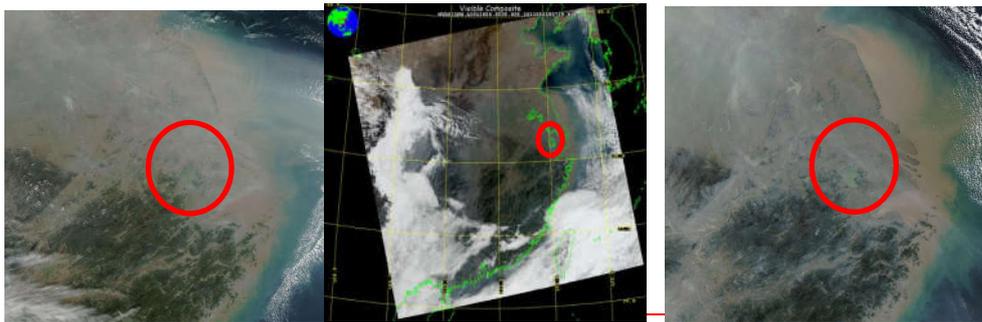
~~Dec. 10, 2011~~



~~Nov. 20, 2013~~

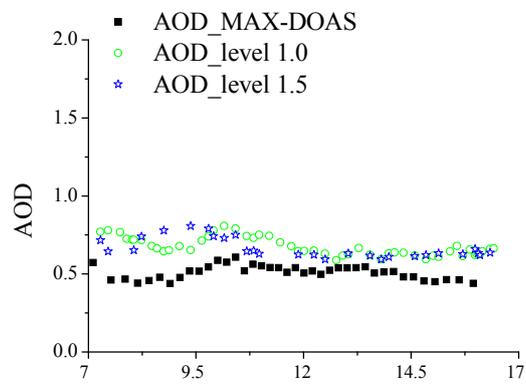
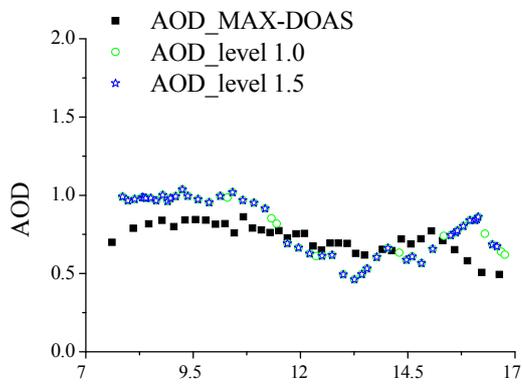
~~Apr. 22, 2012~~

~~Nov. 13, 2013~~



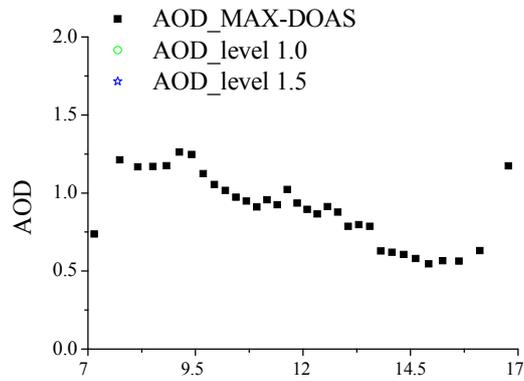
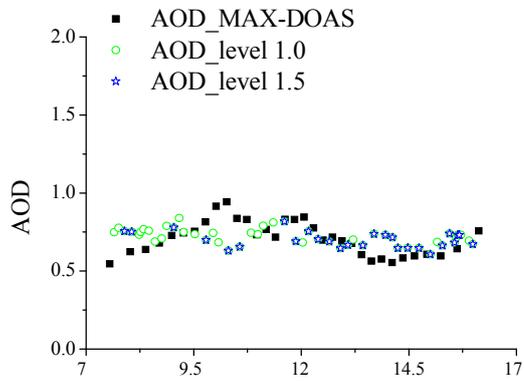
~~(b) Jan. 26, 2012~~

~~Oct. 28, 2013~~



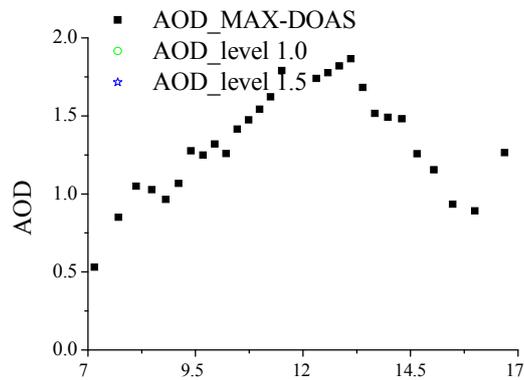
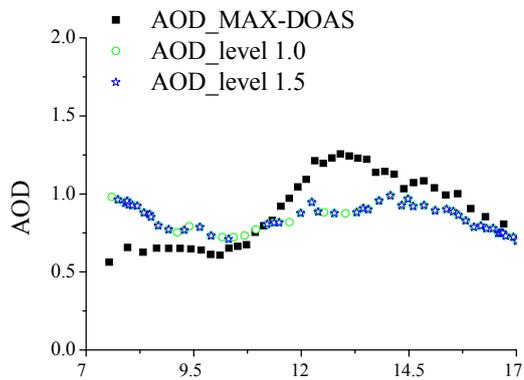
~~Dec. 10, 2011~~

~~Nov. 20, 2013~~

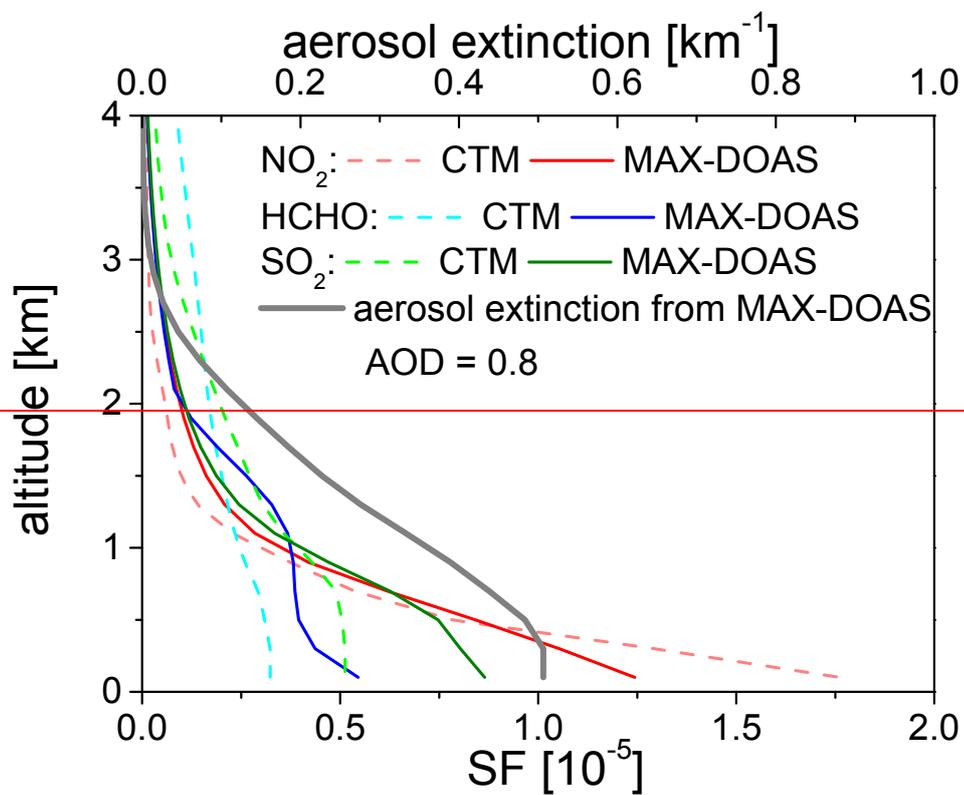


~~Apr. 22, 2012~~

~~Nov. 13, 2013~~



~~**Figure S4: (a) visual images from MODIS on the Aqua satellite on the six days with strong aerosol pollutants, obtained from the MODIS Rapid Response website, NASA/GSFC ([http://aeronet.gsfc.nasa.gov/cgi-bin/bamgomax\\_interactive](http://aeronet.gsfc.nasa.gov/cgi-bin/bamgomax_interactive)); (b) AODs from MAX-DOAS and the nearby Taihu AERONET station on the six days.**~~



**Figure S5: Averaged aerosol extinction profiles and SF of NO<sub>2</sub>, SO<sub>2</sub> and HCHO derived from all MAX-DOAS measurements under cloud free sky conditions. The dashed curves indicate the corresponding averaged SF derived from CTM simulations for NO<sub>2</sub> (TM4), SO<sub>2</sub> (IMAGES) and HCHO (IMAGES).**