Supplement

S1 Pixel screening on low cloud and aerosol

As low clouds and aerosol can induce an error on the column density retrieved by the satellite, a filter on the effective cloud pressure is sometimes employed (Boersma et al., 2018, e.g., only uses pixels with $p_c < 850\,\mathrm{hPa}$). However, a pixel with a low cloud pressure is not necessarily problematic. In this work, we exclude only pixels where the ex-ante uncertainty due to cloud pressure $\mathrm{RMS}(u_{\mathrm{SAT},\,p_c})$ is high. Practically, a one-sided 3-sigma-clipping on $\mathrm{RMS}(u_{\mathrm{SAT},\,p_c})$ is performed. This procedure has a similar effect on $\mathrm{RMS}(u_{\mathrm{SAT}})$ and $\mathrm{RMS}(u_{\mathrm{SAT},\,p_c})$ as a fixed filter on p_c, but has the advantage of removing less pixels. Fig. S1, for the site Xianghe, reveals that most pixels removed by the sigma-clipping approach have indeed a low cloud pressure, but a large portion of pixels with a low cloud pressure are not removed.

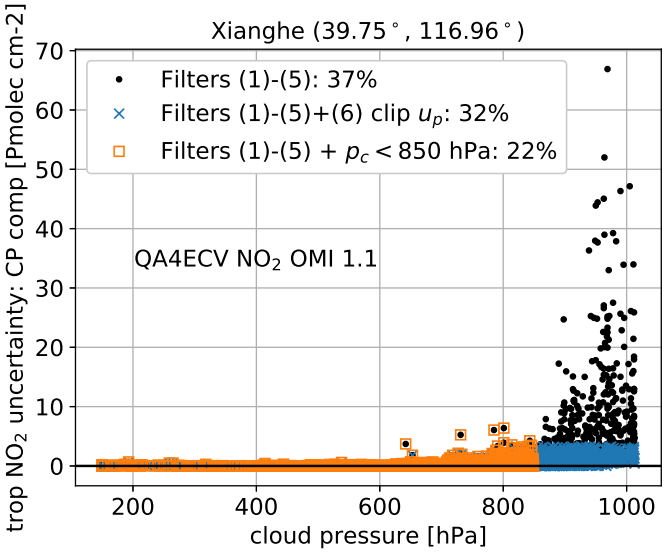

Figure S1. Ex-ante uncertainty due to cloud pressure $u_{\mathrm{SAT},\,p_c}$ vs. cloud pressure, at the site Xianghe, without and with application of flag (6): remove data with $u_{\mathrm{SAT},\,p_c,i} > \mathrm{mean}(u_{\mathrm{SAT},\,p_c,i}) + 3 \times \mathrm{SD}(u_{\mathrm{SAT},\,p_c,i})$ and the alternative of $p_c < 850$ hPa.

S2 Alternative AMF ratio filter

A substantial amount of retrievals are rejected by the lower bound on AMF ratio $\frac{M_{\mathrm{trop}}}{M_{\mathrm{geo}}} > 0.2$ (see Fig. 3). It can be argued that this recommendation from Boersma et al. (2017) is too restrictive. In Fig. S2 the results with an alternative lower bound, 0.05, are presented. There are now significantly more retrievals that are not rejected by the AMF ratio filter (filter (4)) at the sites Bremen, Mainz, OHP, Uccle, Xianghe and (not shown) De Bilt and Cabauw. However, the net data gain after applying the full screening is more modest as many of these retrievals are still rejected by the filter on cloud fraction (filter (5)). For example, at the sites Bremen, Mainz, Uccle, De Bilt, Cabauw now ~10 % data remains (compared to ~8 % with the baseline filtering). Note that especially at the winter months December and January there is a gain for these sites (less than 1% remaining with the

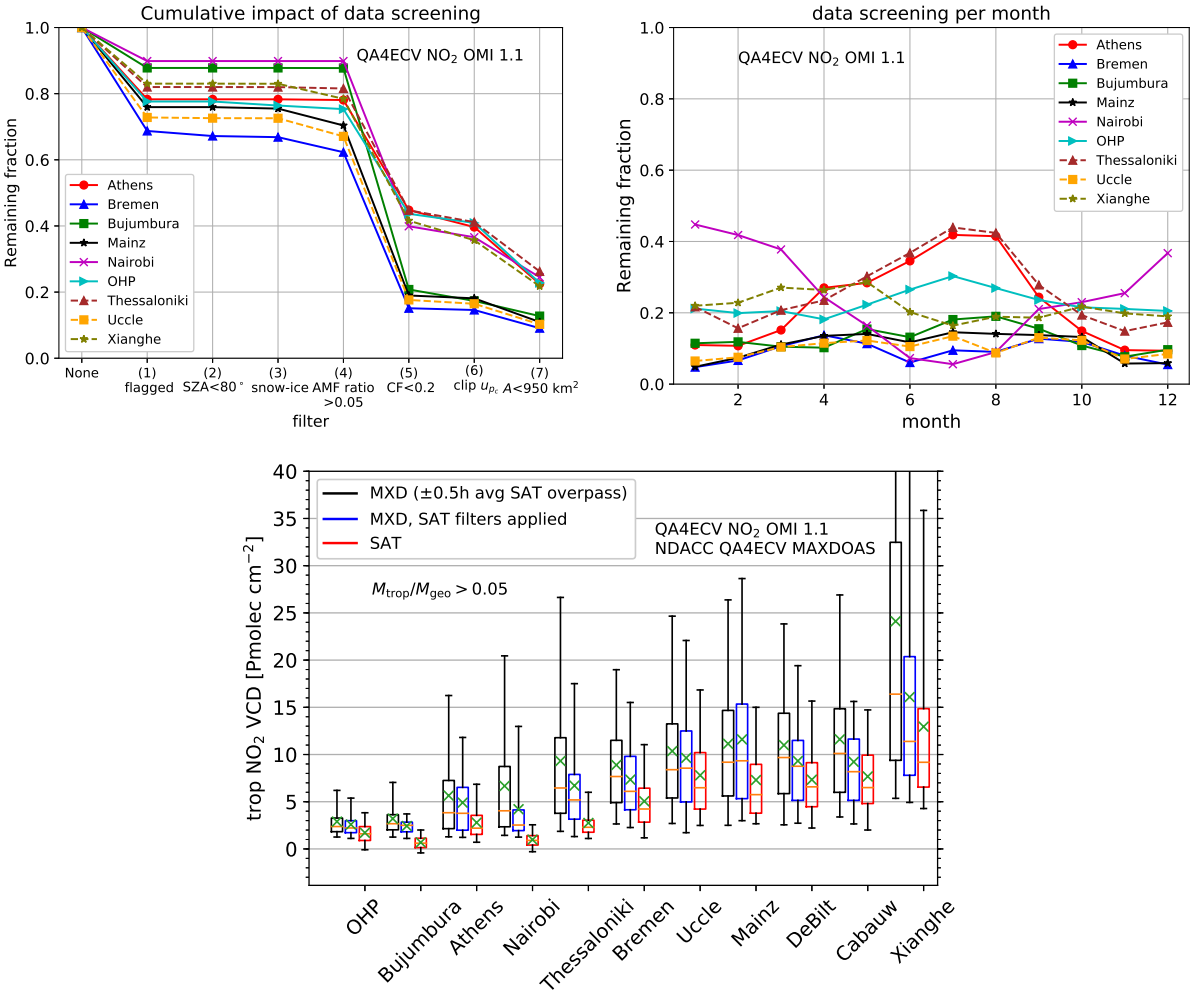

Figure S2. Same as Fig. 3, but with a less restrictive lower bound on $\frac{M_{\mathrm{trop}}}{M_{\mathrm{geo}}}$ of 0.05.

baseline filtering, and ~5% with the modified filtering). The sampling bias at Bremen and De Bilt is reduced with the alternative AMF ratio filter, while it is removed at Mainz.

S3 Horizontal sampling difference error: details of estimation

The RMS of horizontal sampling difference error, $\mathrm{RMS}_{\Delta r_{\mathrm{SAT,GB}}}$, can in principle be estimated by (i) taking the RMS of $N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}$, (ii) subtracting from this in quadrature the standard deviation due to uncorrelated random noise in the satellite measurement, and (iii) finally, assuming that $\mathrm{RMS}_{\Delta r_{\mathrm{SAT,GB}}}$ can be approximated as $\mathrm{RMS}_{\Delta r_{\mathrm{SAT_1,SAT_2}}}$, given that $\Delta r_{\mathrm{SAT,GB}} \approx \Delta r_{\mathrm{SAT_1,SAT_2}}$ (see section 3.4.3)

$$\begin{aligned}\left[\mathrm{RMS}\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right)\right]^2 &= \sigma^2_{\mathrm{SAT_1,noise}} + \sigma^2_{\mathrm{SAT_2,noise}} + \left(\mathrm{RMS}_{\Delta r_{\mathrm{SAT_1,SAT_2}}}\right)^2 \\ &\approx u^2_{\mathrm{SAT_1},N_s} + u^2_{\mathrm{SAT_2},N_s} + \left(\mathrm{RMS}_{\Delta r_{\mathrm{SAT,GB}}}\right)^2 \\ \Rightarrow \left(\mathrm{RMS}_{\Delta r_{\mathrm{SAT,GB}}}\right)^2 &\approx \left(\left[\mathrm{RMS}\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right)\right]^2 - \left(u^2_{\mathrm{SAT_1},N_s} + u^2_{\mathrm{SAT_2},N_s}\right)\right)\end{aligned} \tag{S1}$$

In Eq. (S1), the standard deviation due to uncorrelated random noise is approximated by the ex-ante uncertainty component due to SCD random error. This is justified further below.

Fig. S3 presents, per site, the mean squared difference between $\mathrm{SAT_1}$ and $\mathrm{SAT_2}$, split in a squared bias and a variance component. The variance component is by far the dominant one. The bias reaches at most ~-0.6 $\mathrm{Pmolec\,cm^{-2}}$ (at Athens, Bremen and Mainz), and is therefore only a very minor contributor to $\mathrm{RMS}\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right)$ and to the observed bias between OMI and MAX-DOAS (discussed in section 3.4.6).

The difference in tropospheric column between both neighbouring pixels can be ascribed to differences in total slant column, in stratospheric slant column and in tropospheric AMF (see Eq. (1)). The contribution to the dispersion due to slant column variation can be approximated as $\sigma\left((N_{s,\mathrm{SAT_2}} - N_{s,\mathrm{SAT_1}})/M_{\mathrm{SAT_1}}\right)$ and the dispersion due to tropospheric AMF variation as $\sigma\left(-N_{v,trop,\mathrm{SAT_1}} \times (M_{\mathrm{trop,SAT_2}} - M_{\mathrm{trop,SAT_1}})/M_{\mathrm{SAT_1}}\right)$. This is also shown in Fig. S3. It can be concluded that variation in the total slant column explains the largest part of the tropospheric column dispersion. As stratospheric slant column and tropospheric AMF do not vary much between neighbouring pixels, they can also not be a large source of uncorrelated random error. This, in turn, means that u_{SAT,N_s}, the uncertainty due to random error in the slant column, can be used approximately as the uncertainty due to all uncorrelated random error, justifying the approximation in Eq. (S1).

For the sites Bujumbura, Nairobi and OHP the dispersion $\sigma\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right)$ is roughly equal to the ex-ante noise uncertainty $\sqrt{u^2_{\mathrm{SAT_1},N_s} + u^2_{\mathrm{SAT_2},N_s}}$ (in the order of 1 $\mathrm{Pmolec\,cm^{-2}}$), implying that horizontal sampling difference error is not significant compared to the random noise from the satellite measurement. Specifically for Nairobi and OHP, one has even $\sigma\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right) < \sqrt{u^2_{\mathrm{SAT_1},N_s} + u^2_{\mathrm{SAT_2},N_s}}$, indicating that the uncertainty due to random noise is overestimated at these sites, and/or that there is some error correlation between neighbouring pixels.

For the other sites $\sigma\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right) > \sqrt{2}u_{\mathrm{SAT},N_s}$, i.e., there is a measurable contribution of horizontal sampling difference error. However, it is always much smaller than $\mathrm{RMS}\left(N_{v,\mathrm{trop,SAT}} - N_{v,\mathrm{trop,REF}}\right)$. E.g., at the site Xianghe, $\sigma\left(N_{v,\mathrm{trop,SAT_2}} - N_{v,\mathrm{trop,SAT_1}}\right) \approx 2.5\,\mathrm{Pmolec\,cm^{-2}}$ and $\sqrt{u^2_{\mathrm{SAT_1},N_s} + u^2_{\mathrm{SAT_2},N_s}} \approx 1\,\mathrm{Pmolec\,cm^{-2}}$ (see Fig. S3), therefore $\sigma_{\Delta r_{\mathrm{SAT,GB}}} \approx 2.2\,\mathrm{Pmolec\,cm^{-2}}$. The RMS of the OMI vs. MAX-DOAS comparison at Xianghe, $\mathrm{RMS}\left(N_{v,\mathrm{trop,SAT}} - N_{v,\mathrm{trop,REF}}\right)$, is 7.9

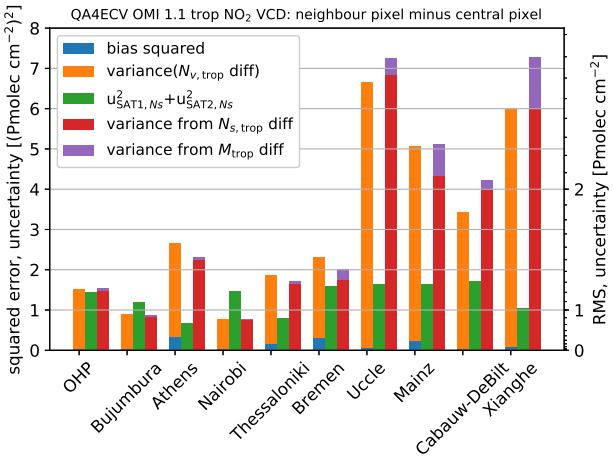

Figure S3. Per site three bar plots are provided. The left bar shows the mean squared difference of tropospheric VCD sets of QA4ECV OMI NO_2 pixels, where for set 1 the ground-pixel covers the MAX-DOAS site and for set 2 the pixel is next-nearest to the MAX-DOAS site. The mean squared difference is divided in two components: a bias component and a dispersion component. The middle bar shows the squared combined uncertainty due to slant column density random error. The right bar components give an estimate of the column variance between both pixels due to variation of the slant column (estimated to first order as $(N_{s,\mathrm{SAT}_2} - N_{s,\mathrm{SAT}_1})/M_{\mathrm{SAT}_1}$) and due to variation of the tropospheric AMF (estimated to first order as $-N_{v,\mathrm{trop},\mathrm{SAT}_1} \times (M_{\mathrm{trop},\mathrm{SAT}_2} - M_{\mathrm{trop},\mathrm{SAT}_1})/M_{\mathrm{SAT}_1}$). Note that the sum of variances due to SCD and tropospheric AMF difference can overshoot the variance of trop VCD difference (e.g., at Uccle), due to correlation and due to the first order approximation.

$\mathrm{Pmolec\,cm^{-2}}$ (see Fig. 12). Subtracting in quadrature $\sigma_{\Delta r_{\mathrm{SAT,GB}}}$ would reduce this to 7.6 $\mathrm{Pmolec\,cm^{-2}}$. A similar reasoning for Uccle leads to a reduction of $\sigma_{\Delta r_{\mathrm{SAT,GB}}}$ from 4.9 $\mathrm{Pmolec\,cm^{-2}}$ to 4.4 $\mathrm{Pmolec\,cm^{-2}}$. For the other sites the contribution of horizontal sampling difference error is smaller. Horizontal sampling difference error is therefore a minor contributor to $N_{v,\mathrm{trop},\mathrm{SAT}_2} - N_{v,\mathrm{trop},\mathrm{SAT}_1}$.

S4 NO_2 monthly fields

For a selection of sites, spatial NO_2 fields were constructed for each month from the QA4ECV OMI 1.1 data, averaged over years 2004-2016. A variant of the temporal averaging approach of Wenig et al. (2008) is used. Per 0.02×0.02 grid cell the arithmetic average of covering satellite ground pixels is taken. Note that only ground pixels with area $< 950\,\mathrm{km}^2$ are selected. Selection of months is based on minimal monthly bias (left) and maximal monthly bias (right) from Fig. 10.

The true resolution of these NO_2 fields is lower than the $0.02° \times 0.02°$ grid. The theoretical spatial smoothing window (taking into account the ground pixel size) is provided in Fig. S4. The NO_2 fields are shown in S5.

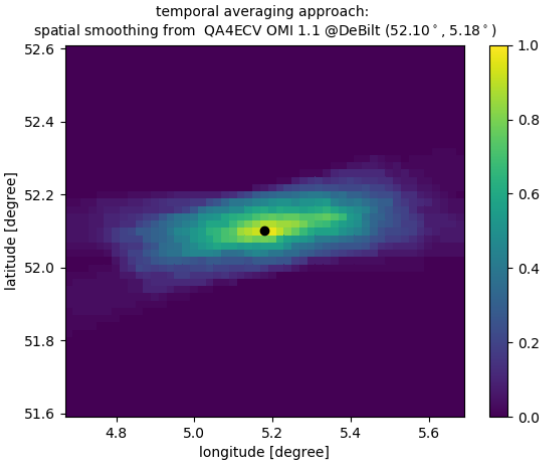

Figure S4. Spatial smoothing window of the NO_2 field constructed from QA4ECV OMI 1.1 data.

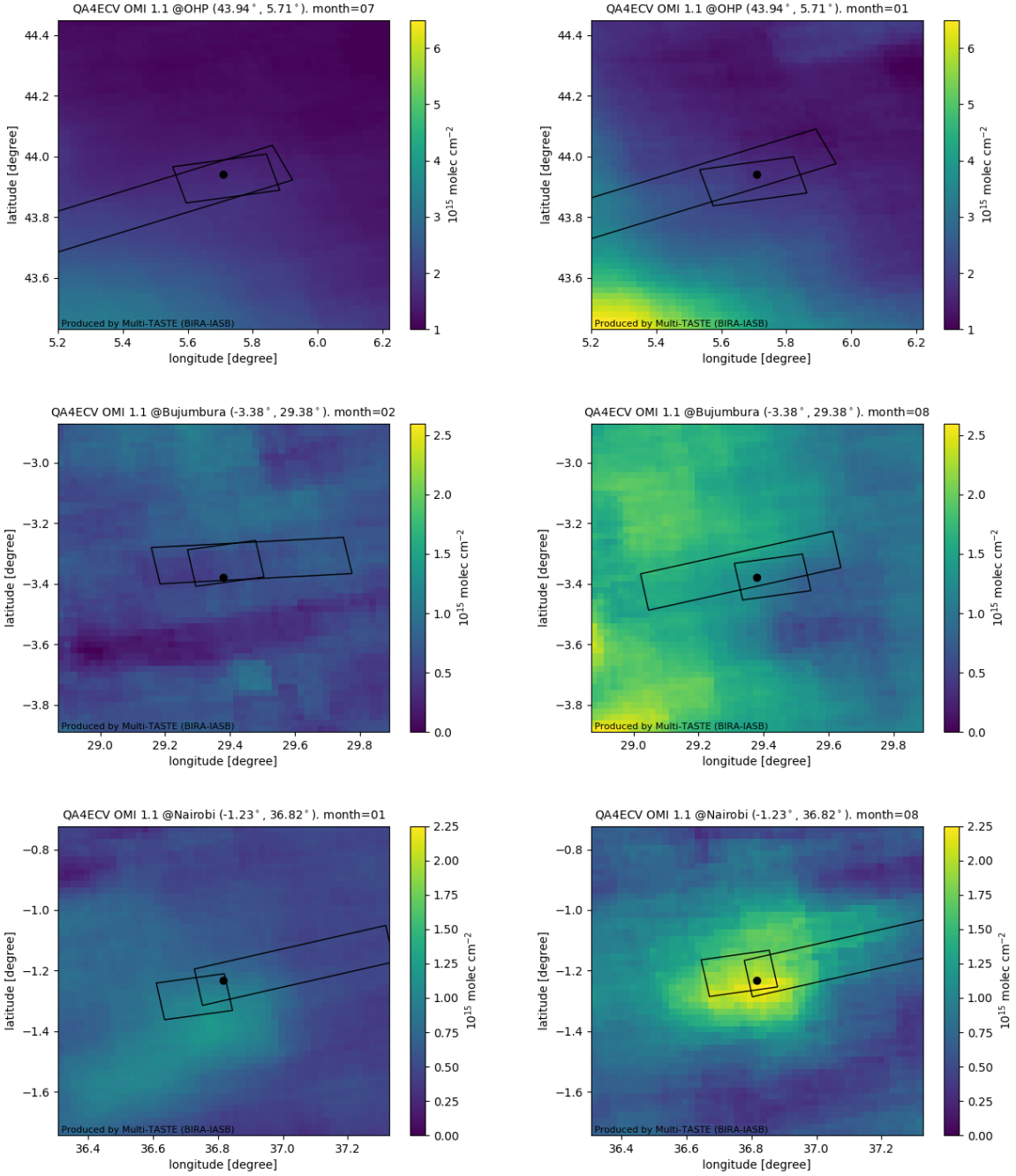

Figure S5. NO_2 monthly fields on a $0.02° \times 0.02°$ grid, obtained from QA4ECV NO_2 OMI 1.1 by averaging over all years, for a selection of sites. Selection of months is based on minimal monthly bias (left) and maximal monthly bias (right) from Fig. 10.

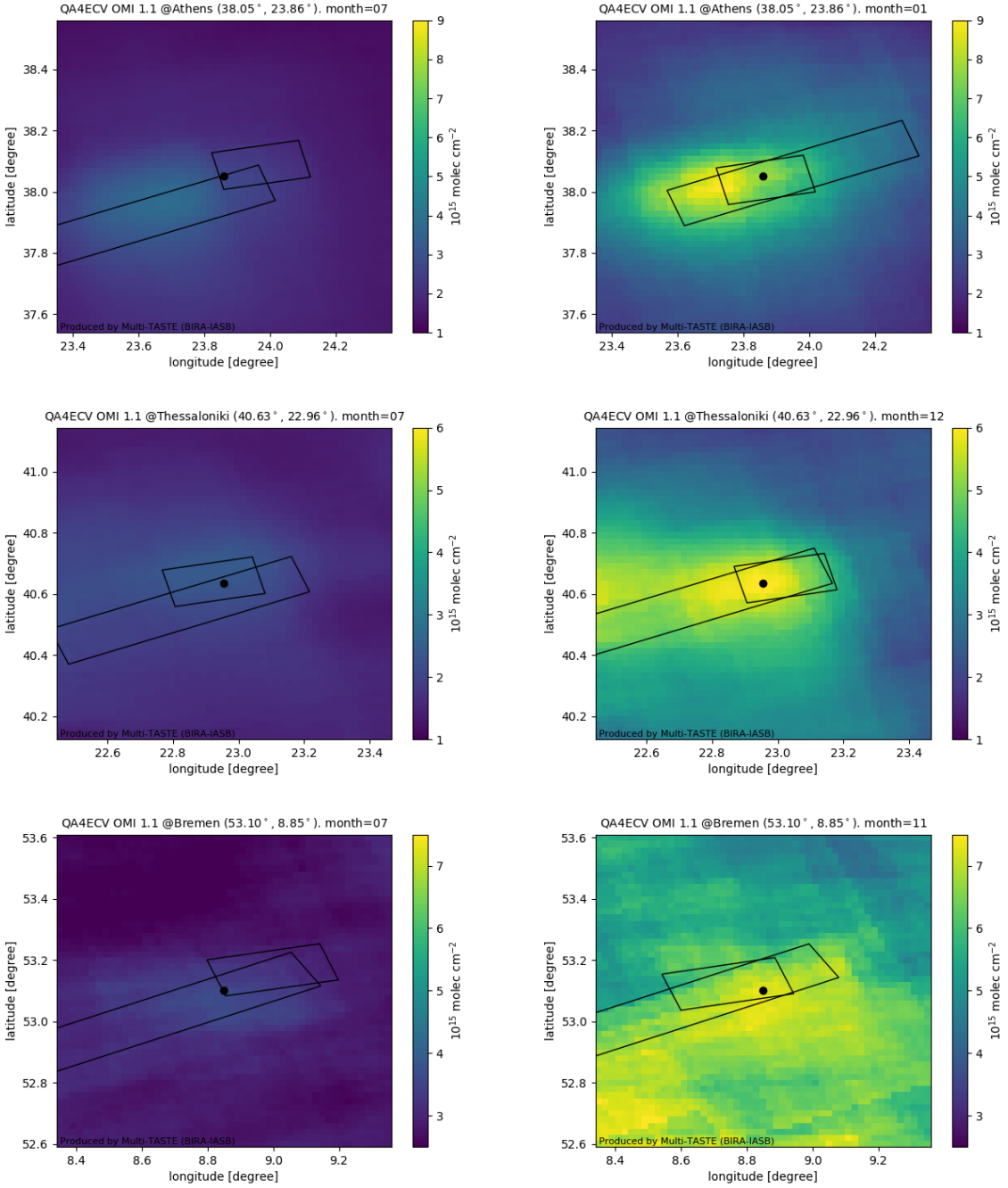

Figure S6. Continuation of Fig. S5.

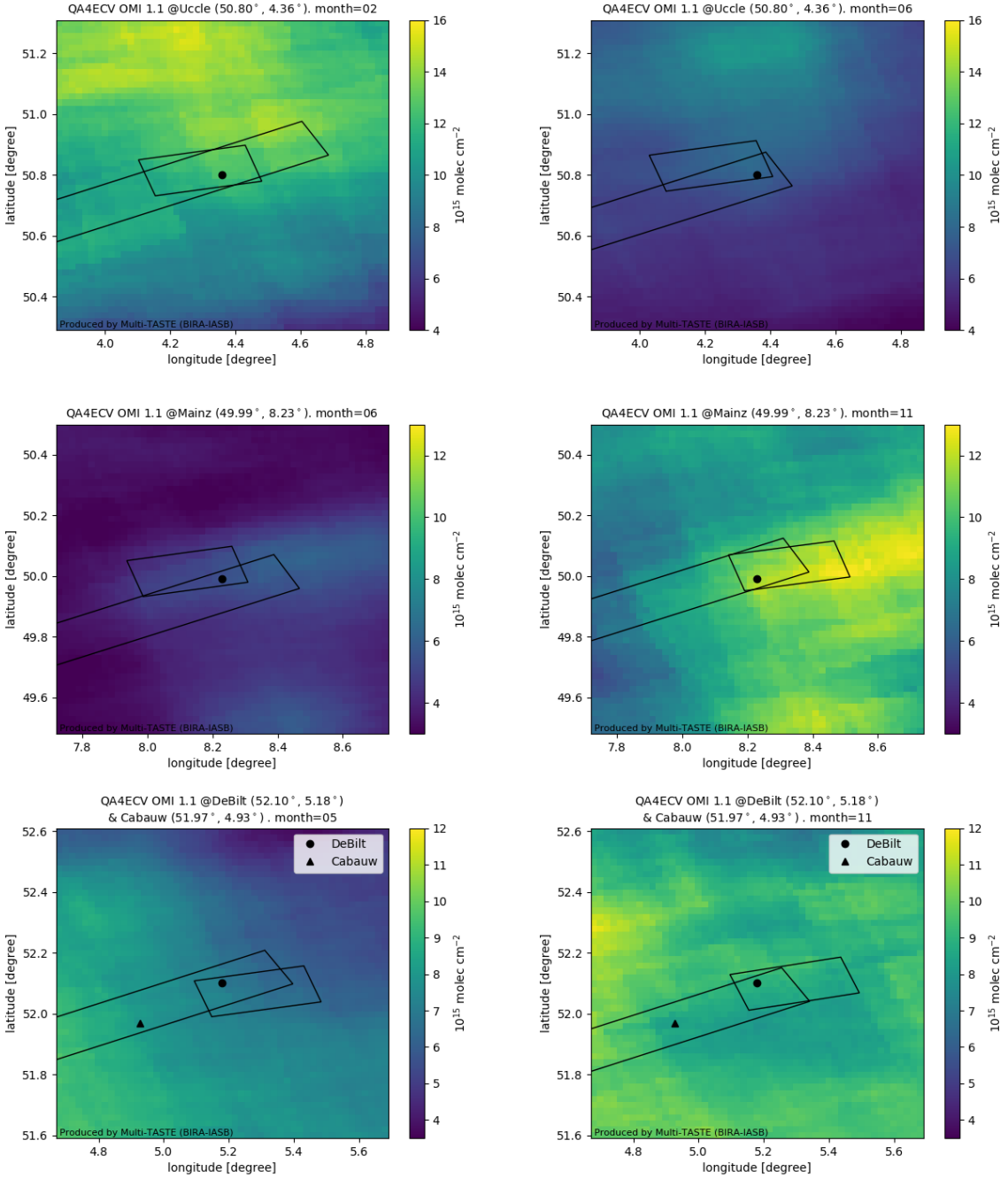

Figure S7. Continuation of Fig. S5.

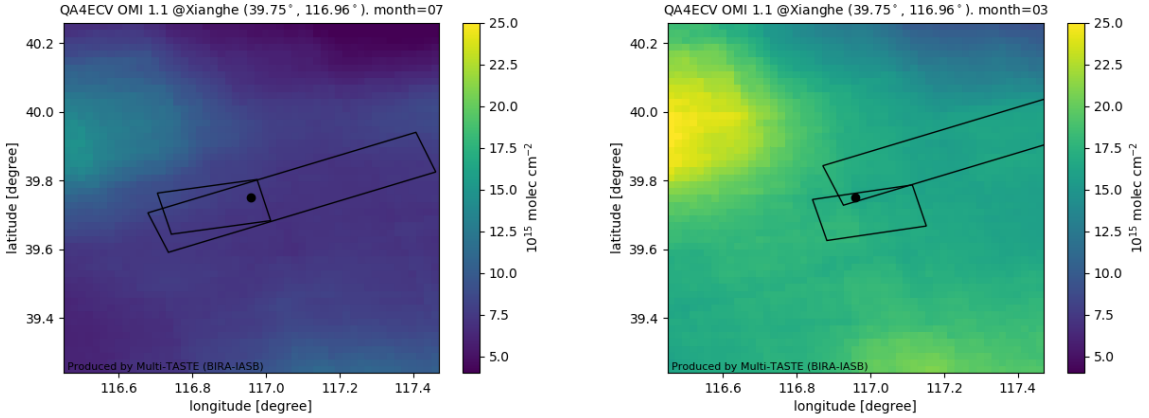

Figure S8. Continuation of Fig. S5.

S5 Dependence on screening criteria

Fig. S9 displays the dependence of mean difference between QA4ECV OMI tropospheric and MAX-DOAS (left column) and standard deviation of the difference (right column) on varying screening criteria. The impact of 4-σ clipping is also presented.

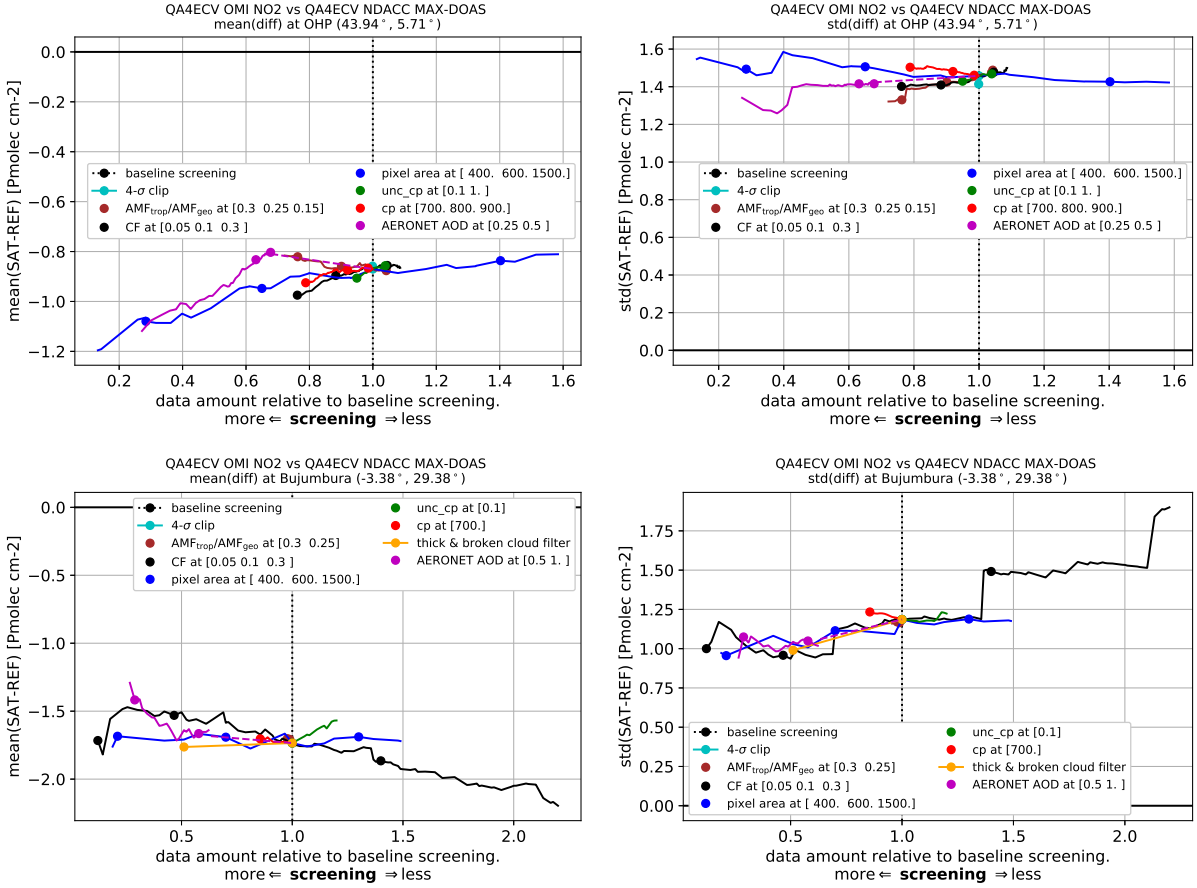

Figure S9. Mean difference (left) and standard deviation (right) in function of varying the screening criteria, for QA4ECV OMI NO_2 vs QA4ECV MAX-DOAS, for the sites OHP and Bujumbura. The x-axis shows the amount of selected (OMI, MAX-DOAS) pairs relative to the baseline screening (indicated with a vertical dashed line). Screening criteria varied here are minimum $\frac{\mathrm{AMF_{trop}}}{\mathrm{AMF_{geo}}}$, maximum ground pixel area, cloud fraction, uncertainty due to cloud pressure and AERONET AOD. Note that co-location with AERONET implies a subsetting of data, here indicated with a dashed line. Also provided are the impact of 4-σ clipping and, if available, the rejection of scenes flagged according to the MAX-DOAS data as 'broken cloud' or thick cloud'.

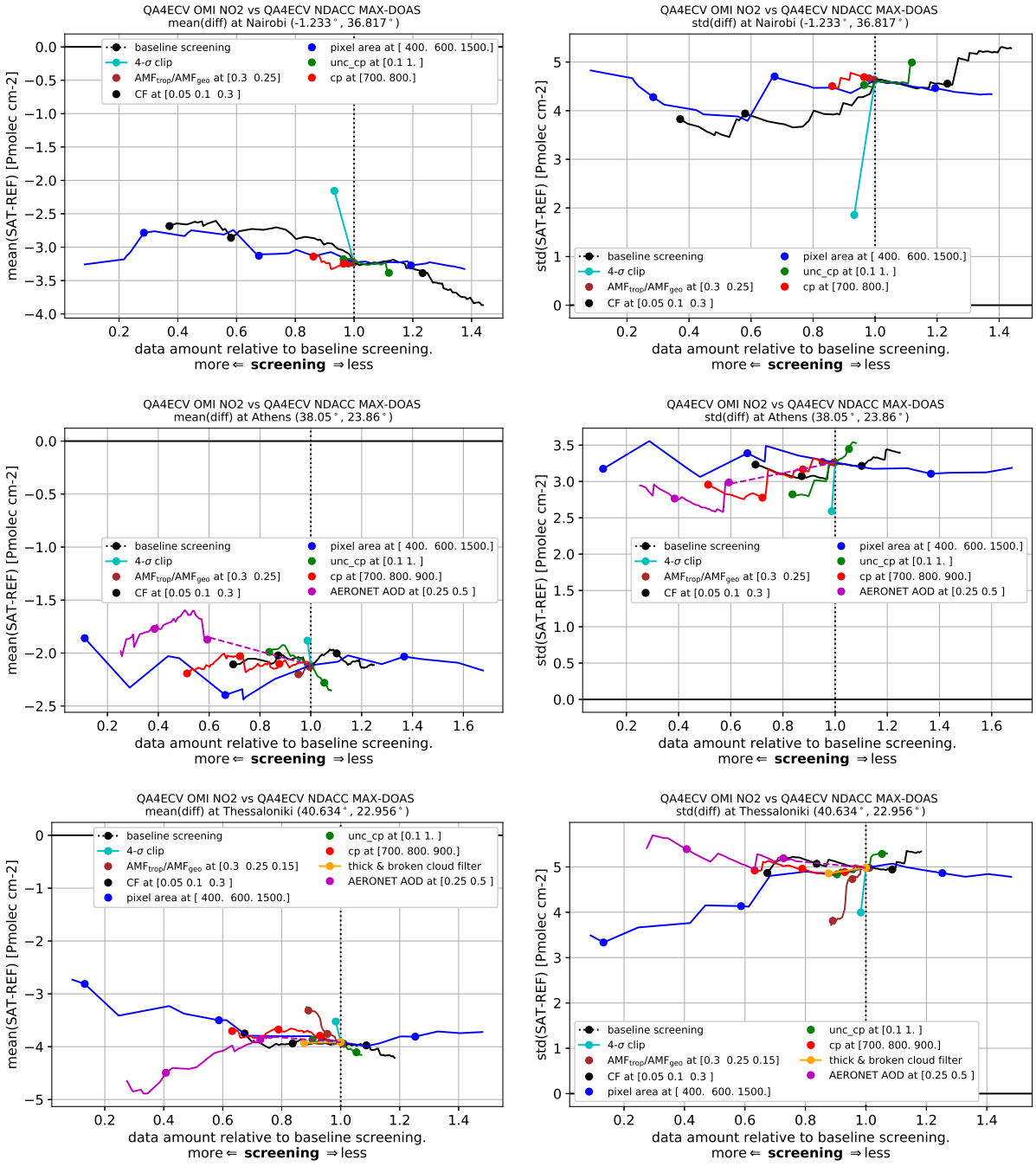

Figure S10. As Fig. S9, but for the sites Nairobi, Athens and Thessaloniki.

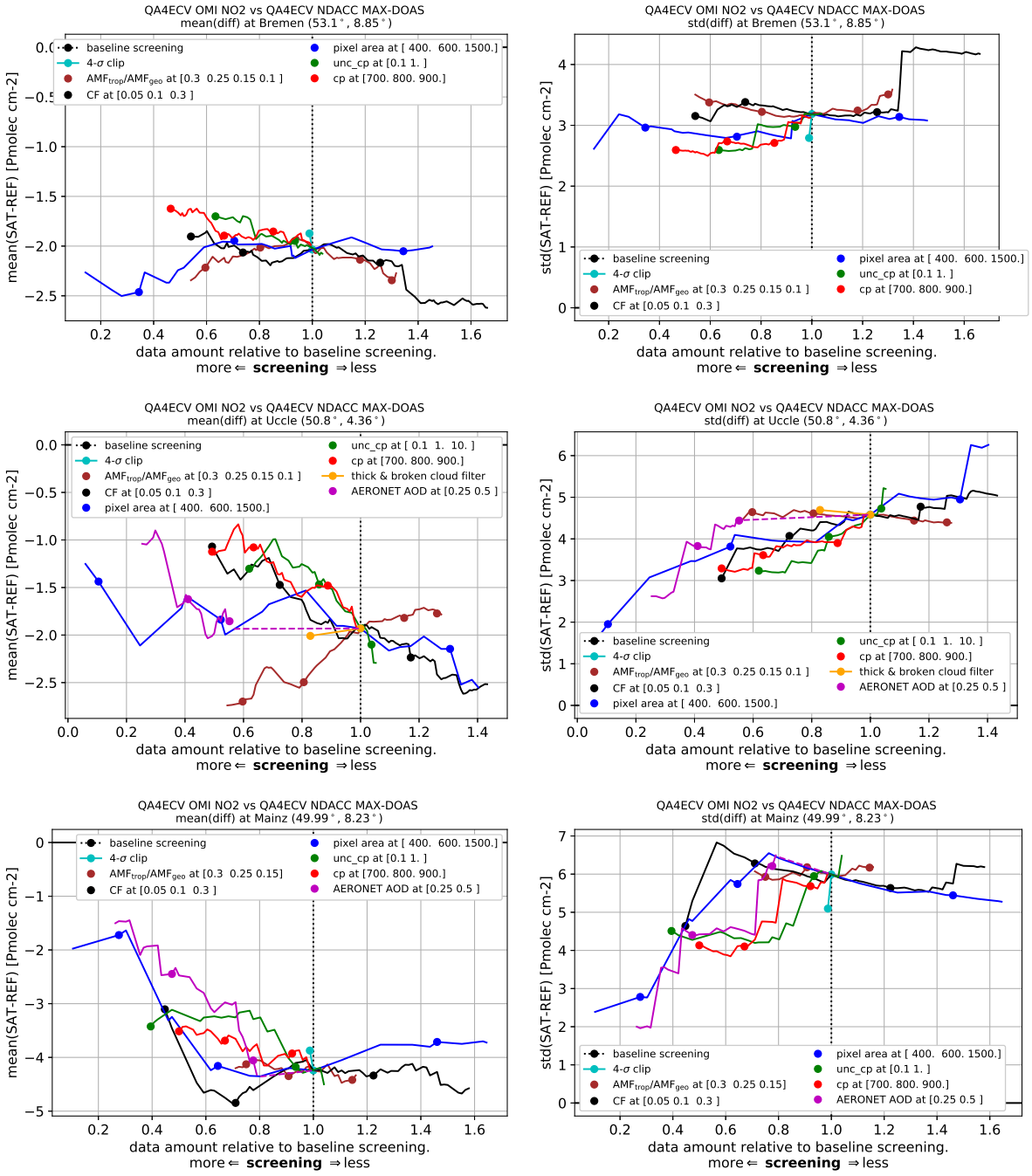

Figure S11. As Fig. S9, but for the sites Bremen, Uccle and Mainz.

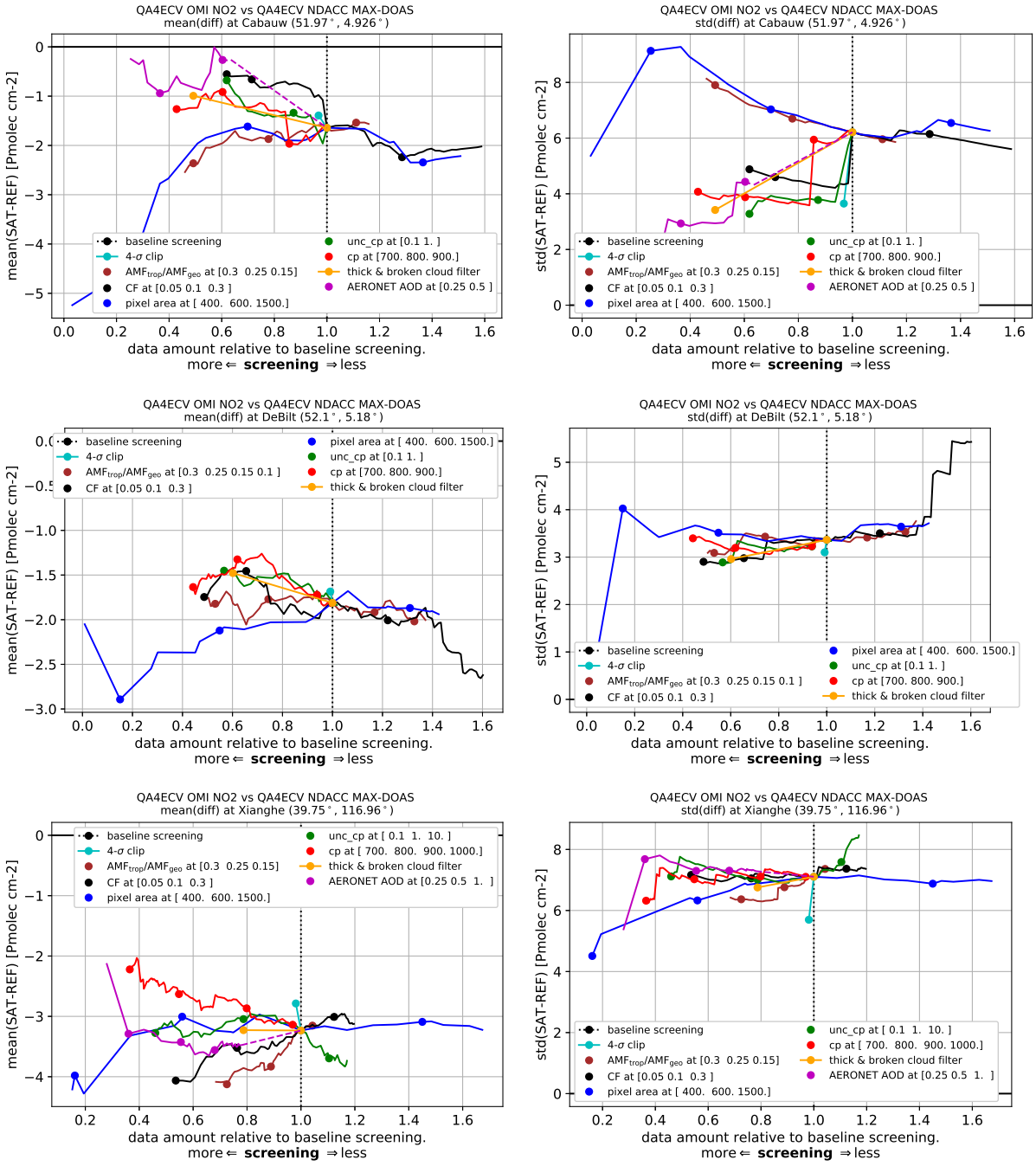

Figure S12. As Fig. S9, but for the sites Cabauw, De Bilt and Xianghe.

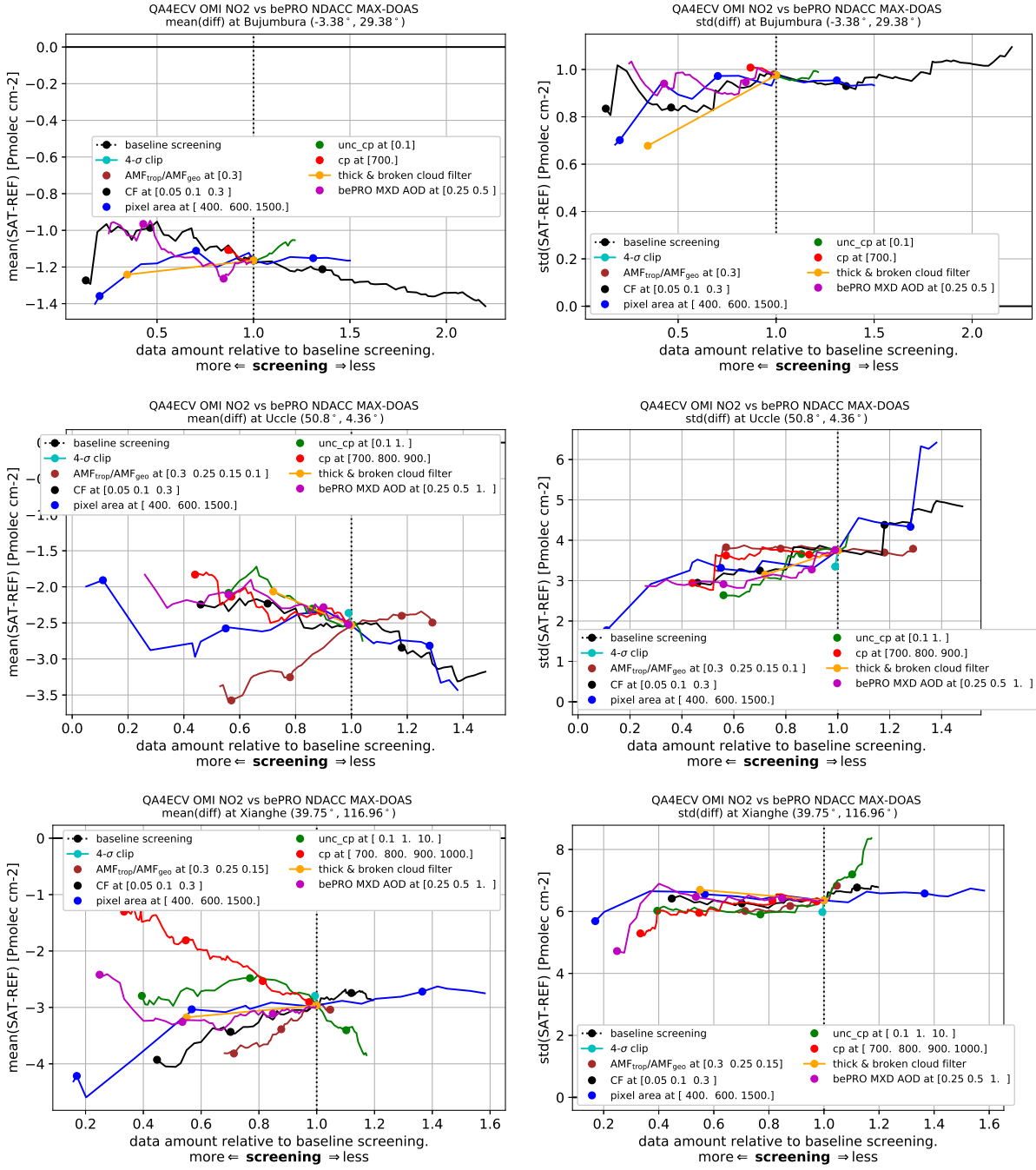

Figure S13. As Fig. S9, but for the bePRO MAX-DOAS processing of Bujumbura, Uccle and Xianghe. Note that for AOD, bePRO MAX-DOAS is used instead of AERONET.

S6 Vertical a priori profile harmonization and smoothing

Fig. 13 presents the mean squared deviation of QA4ECV OMI vs. bePRO MAX-DOAS at Uccle and Xianghe, using (i) the OMI satellite and MAX-DOAS data, (ii) direct smoothing of the MAX-DOAS retrieved profile using the satellite averaging kernel, (iii) replacement of the MAX-DOAS a priori profile shape with that of the satellite, (iv) a combination of first (iii), then (ii).

In more detail:

- **(ii) Direct smoothing of the MAX-DOAS retrieved profile.** This is obtained as follows: the MAX-DOAS retrieved partial column density profile is shifted to the surface pressure level of the satellite (method of Zhou et al., 2009) and then mass-conserved regridded (Langerock et al., 2015) to the QA4ECV OMI pressure grid.

- **(iii) Replacement of the MAX-DOAS a priori profile shape with that of the satellite.** This is obtained as follows: the OMI a priori profile (from TM5) is shifted to the surface pressure of MAX-DOAS (method of Zhou et al., 2009), then regridded (Langerock et al., 2015) to the MAX-DOAS pressure grid, and rescaled to match the a priori VCD size of the original MAX-DOAS a priori. The resulting profile, defined on the MAX-DOAS grid, is then used to replace the original MAX-DOAS a priori, using Eq. (10) of Rodgers and Connor (2003), thereby modifying the MAX-DOAS retrieved profile. With this procedure we have modified the shape of the MAX-DOAS a priori profile but not the a priori VCD size.

  The motivation is as follows: the *column density size* of the MAX-DOAS a priori profile is obtained from a geometrical approximation MAX-DOAS measurement, therefore likely more correct than the TM5 modelling. The MAX-DOAS a priori *profile shape* however is modelled by a simple exponential, and therefore likely to be less correct than the TM5 modelling.

  As the OMI retrieved column depends on the a priori profile shape but not on its size, and since the satellite a priori shape is unaltered in the above procedure, the OMI retrieved VCD is not modified.

- (iv) **Combination of first (iii), then (ii).** First all steps of (iii) are applied to obtain a harmonized MAX-DOAS retrieved profile. Then this profile is shifted (method of Zhou et al., 2009) and regridded (Langerock et al., 2015) to the OMI pressure grid. Then Eq. (3) is applied to this profile using the OMI averaging kernel.