Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations

ESA Contract No. 4000135355/21/I-DT-lr

Deliverable D3.3:

Validation Report

Date: 01/06/2025

Version: 1.0

Contributing authors:

M. M. Friedrich, A. Merlaud, G. Pinardi, C. Fayt and M. Van Roozendael (IASB-BIRA)

U. Frieß (IUP-HD)

T. Wagner, S. Ziegler, and S. Beirle (MPIC)

A. Piters (KNMI)

A. Bais, D. Nikolis and D. Karagkiozidis (AUTH)

K. Krause, A. Richter (IUP-B)

Table of Contents

1 Introduction 5
2 Case study MAINZ 5
2.1 Preliminary Tests 5
2.1.1 Improving the time coincidence of the ceilometer measurements with the MAX-DOAS scans 5
2.1.2 Improving the spatial coincidence between scaled ceilometer profiles and MAX-DOAS retrieval profiles 7
2.1.3 ad-hoc filtering for clouds 10
2.2 MMF retrieval tests 11
2.2.1 Investigating the effect of a priori settings 13
2.2.2 Investigating the effect of lowering the scaling factor on the measured DSCDs 17
2.2.3 Investigating the effect of a priori settings when using a fixed scaling factor on the measured DSCDs 21
2.2.4 Final considerations 24
3 Other stations in the UV 25
3.1 Uccle and Cabauw 26
3.2 Xianghe 29
4 Stations in the visible 30
4.1 Uccle 30
4.2 Xianghe 33
5 UV and VIS in UCCLE 35
6 Unified validations 36
6.1 Mainz 38
6.2 Uccle 43
6.3 Cabauw 46
3.4 Xianghe 47
7 Summary and conclusions 49
References 50

Appendix 50
A.1. Mainz-1 MMF original vs MMF v8 extinction layers 50
A.2 Mainz impact of different MAPA flag components on MMF retrieval correlation with AERONET data 53
A.3 Mainz: Tests considering the smallest and biggest RAA of each set in the different directions.. 55
A.4 Unified AOD scatter plots 57

1 Introduction

The technical note [D3.2] highlighted the fact that while AOD values retrieved by MMF showed generally good correlation with AERONET AOD values, it vastly underestimated them in the UV and slightly in the visible. While MAPA VAR showed generally a slightly worse correlation with AERONET AOD compared to MMF, no such underestimation was detected.

A further shortcoming of MMF was found to lie in the sub-optimal filtering of valid data with only a negligible fraction of data being labelled as valid. In the UV, it was found that the MAPA flag works very well, while in the visible, an additional criterion helped improving the results.

To find more suitable retrieval settings for MMF, we first restrict tests to Mainz. Mainz was chosen because it offers simultaneous MAX-DOAS measurements in 4 different azimuthal directions. The tests performed are described in detail in Sect. 2. The important changes for the retrieval settings were then applied to three other stations in the UV: Xianghe, Uccle and Cabauw. The results are presented in Sect. 3, as well as for the visible in Sect.4. We present comparisons for UV and VIS in UCCLE in Sect.5 . The unified validation analysis for all those stations together is presented in Sect.6 and the main findings and conclusions are summarize in Sect.7.

Unless otherwise stated, in this document, MAPA refers to MAPA VAR since it is considered the default version in FRM4DOAS.

2 Case study MAINZ

Prior to testing alternative retrievals for MMF, the robustness of the validation dataset and the mode of comparisons were tested. This is summarised in Sect. 2.1. Sect. 2.2 introduces the different tests for new retrievals.

The measurements in MAINZ are exclusively in UV, therefore only the aerosol at 360 nm is considered. The AERONET AOD is treated in the same way as described in [D3.2] for comparison to the MAX-DOAS retrieved AOD.

2.1 Preliminary Tests

2.1.1 Improving the time coincidence of the ceilometer measurements with the MAX-DOAS scans

To compute the comparison profiles for the MAX-DOAS aerosol extinction profile validation, the default setup used in [D3.2] used the nearest (in time) available ceilometer profile scaled

with the nearest AERONET value for the AOD. The details are outlined in the technical note [D3.2].
A typical MAX-DOAS scan has a time duration of 5-20 min, while the time resolution of ceilometer measurements is on a sub-minute timescale. Hence, a more accurate comparison profile should take this duration into account and consider the time-average of ceilometer measurements during the duration of each scan in question. Additionally, for short scans such as in Mainz (5–10 min), the zenith reference measurements are often not taken in between every scan and hence the total time to consider is even a bit longer than the time span from the first to the last off-axis measurement.

In Mainz, often 2 scans are made in-between two zenith measurements. To account for this, both at the end and at the start of the scan, we added an additional 2 min to the time window used for calculating the average ceilometer profile. Although we could define the length of each scan as the time between the two off-axis measurements, we chose not to, as this would produce identical comparison profiles for multiple scans taken at different points in time.

We tested this approach for the four different measurement directions in Mainz. Fig.1 compares the correlation of the estimated extinction profile with the retrieved profile by MAPA (left panel) and MMF (right panel), using the same baseline data as used in the technical note [D3.2]. The correlation (as a function of height above ground) with the default comparison profiles is shown as dashed lines, the correlation with the new comparison profile is shown with solid lines, the distinct colours represent the different directions. For both MAPA and MMF, directions 2 (north east), 3 (south east) and 4 (south west) show a clear increase in correlation with the comparison profile in all layers below 3 km. For direction 1 (north west), the correlation is in general lower than in the other directions, and the improvement is less pronounced.

Based on this finding, the comparison profiles used in further investigations in this section will use the new estimates for the comparison profiles calculated using a time average of ceilometer profiles over the (slightly extended, +2 min on each side) duration of each off-axis scan.

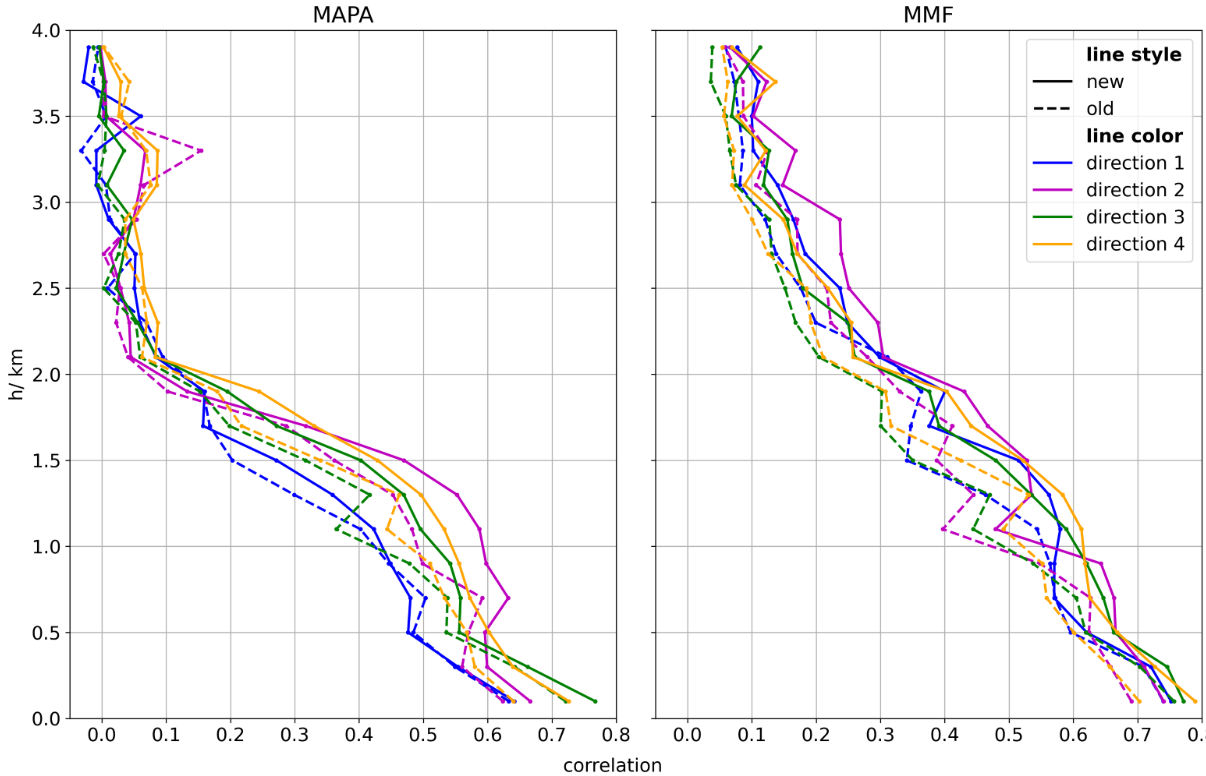


Fig.1: Correlation between scaled ceilometer aerosol extinction profiles and MAX-DOAS retrieved profiles at 360 nm: MAPA (left panel) and MMF (right panel). The dashed lines represent correlations with comparison profiles using the nearest ceilometer backscatter profile, while solid lines show the correlation with profiles using a time-averaged profile during the duration of a scan (+2 min on each side). The distinct colours represent the different azimuthal viewing directions. Note the general improvement of correlation below 3 km when using the time-averaged ceilometer profile, most visible in directions 2, 3 and 4.

2.1.2 Improving the spatial coincidence between scaled ceilometer profiles and MAX-DOAS retrieval profiles

As is evident from Fig. 1, especially for MAPA, there are significant differences in the correlation coefficient up to 2 km considering different directions. This is independent of whether we consider the more or the less accurate time match (see Sect.2.1.1) for the construction of the reference profile.

While the MAX-DOAS measurements were carried out towards these different directions, the comparison profiles calculated from the ceilometer backscatter profiles and the AERONET AODs are representative for the location above the measurement location (for the latter, this is strictly only true for very low solar zenith angles). Hence, the differences in correlation towards the different directions do not necessarily mean that the MAX-DOAS retrievals work better in one direction than in another. The differences in correlation between the different directions are larger for MAPA than for MMF. This might either indicate that more details in the profiles towards different directions are kept in the MAPA data compared to MMF data or that MMF profile retrieval is more stable. Tests increasing the location representative agreement for AOD (i.e. choosing always the direction for which the relative azimuth angle between the measurement direction and the azimuth angle of the sun is the smallest) reveals that especially for MAPA, correlation is worse for small RAA. For MMF this effect is also there, but much less pronounced (compare Appendix A.3). This can likely be attributed to an inappropriate choice of the aerosol phase function (in both cases Henyey Greenstein phase function, Henyey and Greenstein, 1941), resulting in a lower accuracy for forward scattering. Towards all directions, MMF shows a higher correlation at nearly all layers (except for 2 height layers in direction 2, see below) than MAPA.

We cannot directly test which of the two is the case unless we had ceilometer data available in distinct locations in the different lines of sight of the MAX-DOAS azimuth measurement directions. In either case, we should see an improvement in correlation when considering the mean of the MAX-DOAS profiles in the different directions.
Since each MAX-DOAS retrieval from measurements in each direction has its own quality control (we remind the reader that both MAPA and MMF use the MAPA quality flag in this investigation considering the UV aerosol product), requiring all 4 directions simultaneously to be valid, significantly reduces the number of coincident measurements from roughly 4000 (3793, 4991, 3669 and 4022 for directions 1–4, respectively) in each direction separately, to less than 1400 (1395). This is much lower than the number that appears in Fig.A.7 and Fig.A.8 since we require here the existence of a ceilometer backscatter profile; this is not the case in the considerations in Appendix A3.). Hence, we only require at least 3 directions to be valid for calculating the average profile from the different directions at any given time, taking all valid directions (so either all 4 or a subset of 3 directions) into account for the calculation of the mean. This increases the valid coincidences to slightly more than 3000 (3058).

In Fig.2, we show the correlation between the MAX-DOAS retrievals (MAPA in cyan, MMF in red) with the improved time-match comparison profile in each direction separately (first 4 panels) and the correlation considering the mean from at least 3 of the 4 directions in panel 5. The MMF retrieval (except for a small height range in direction 2) shows a better correlation with the comparison profile at all heights than MAPA considering the individual directions. This is not any longer true when considering the mean of the directions: the lowest 1.5 km show similar correlations, and the correlation is improved compared to the separate directions. The thin lines in the last panel show the correlations in separate directions (as in the first four panels) to make this more visible.

Especially for MAPA, using the average of 3—4 directions, the correlation in the lowest 1.5 km is significantly increased. The improvement for MMF is much less pronounced. It is further visible that there is a clear breaking point somewhere between 1.5 and 2 km: the performance (in terms of correlation to the comparison profile) of the two retrieval codes up to this height is remarkably similar. While MMF shows an almost steady decline in correlation above this height, the correlation of the MAPA retrieval declines sharply. Considering the mean of directions, the correlation compared to considering the directions separately, is even reduced.

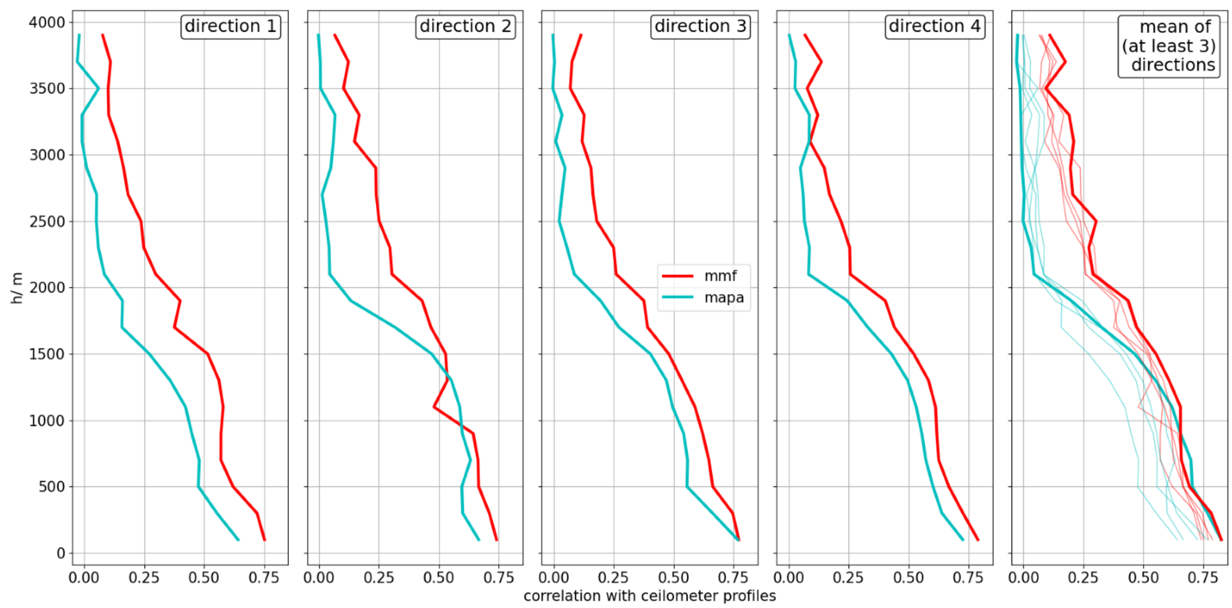


Fig.2: Correlation between scaled ceilometer aerosol extinction profiles with MAX-DOAS retrieved profiles at 360 nm: MAPA (cyan) and MMF (red) in different directions (first 4 panels) and considering the mean of the retrieval profiles in the four directions (last panel). The thin lines in the last panel are the correlations of the retrievals from the measurements in individual directions as shown in the first four panels.

Apart from the correlation with the comparison profiles, it is also insightful to compare the mean and median profiles of MMF, MAPA and ceilometer data, considering only time coincidences (i.e. only considering those times where both valid MAX-DOAS measurements as well as AERONET data and ceilometer backscatter profiles are available). We show these in Fig.3. The minor differences seen in the different directions in the scaled backscatter profiles are solely due to the different sampling, since only those times at which valid MAX-DOAS retrievals in the directions in question are present, were considered.

Since the ceilometer has a blind region that spans most of the lowest layer (0—200m), the lowest layer was constructed from the backscatter around 200m. Hence this is likely an underestimation. Since the whole profile is scaled by the AERONET AOD (compare the validation report D3.3), the remaining layers might be slightly overestimated. Due to the unknown underestimation of extinction in the lowest layer in the scaled backscatter profiles, the following statements exclude the lowest layer.

MMF underestimates the aerosol extinction at all heights. Note that the median and mean MMF profiles are distinctly below the MMF a priori profile (dotted lines) below 2000 m. The latter is in much better agreement with the median backscatter profile than the retrieval. MAPA overestimates the profile below 600 m and underestimates it above. These statements are true for all directions as well as considering the mean of directions.

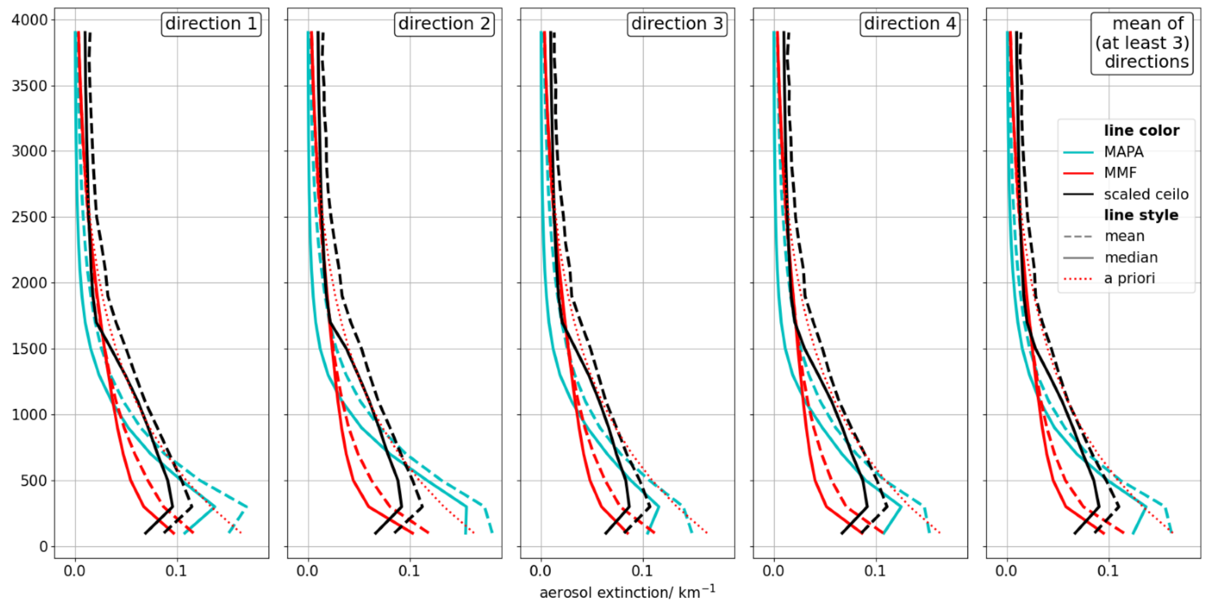
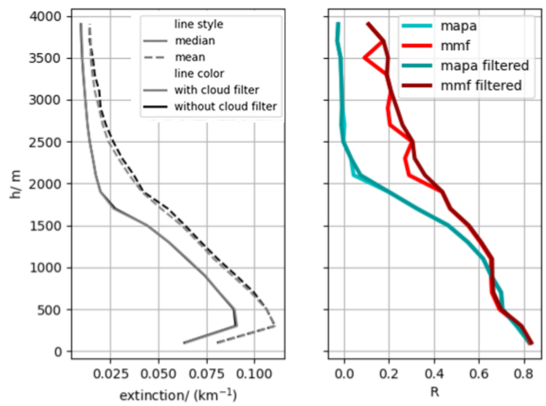


Fig.3: Mean (dashed lines) and median (solid lines) profiles of MAPA (cyan), MMF (red) and scaled ceilometer backscatter profiles (black) in the four different directions (first 4 panels) as well as the mean of the directions (for MAPA and MMF) in the last panel. In each panel, the a priori profile of MMF is included as red dotted line.

These two preliminary tests show that a good match in both time and space is essential for carrying out the validation analysis. Based on those findings, further analysis in this section will often concentrate on the mean of the retrievals in the different directions as described above. However, even though not always shown, the individual directions were still included in all checks to avoid overlooking outliers in certain directions.

2.1.3 ad-hoc filtering for clouds

Lastly, since the scaling of the backscatter profile with the sun photometer AOD should only be applied in the absence of clouds, since clouds normally have a different backscatter coefficient. In theory, no AOD values from (the cloud screened) AERONET are available in the case of clouds. However, since the AERONET AOD was used as the nearest time to the middle of the



scan time and the viewing directions of the sun photometer are always different if the sun is not in the zenith, we choose to apply an additional soft cloud filtering here by limiting the third derivative of the extinction profile (in extinction per km) to 2. This removes only about 2 % of the scaled ceilometer data, and, if we consider the coincidence with valid mean of the MAX-DOAS directions (i.e. where we also require that the MAPA data is valid in at least 3 directions), it only removes 0.3 % of the data, so the total number after applying the cloud filter is 3050 instead of 3058. The effect on the mean and median profiles is negligible (see Fig.4) as is the effect on the correlation, more so below 1.5 km. Above 1.5 km, the effect is still small, but applying the cloud filter, the decrease of correlation with height is slightly smoother, this is more pronounced for MMF than for MAPA.

Fig.4: Left panel: Median (solid) and mean (dashed) profiles without (black) and with cloud filter. Right panel: correlation of MAPA (cyan) and MMF (red) with the scaled ceilometer profile (as in the last panel of Fig.2) without (light colours) and with (darker colours) cloud filtering. Note the small effect on the mean and median profile and the overall small effect on the correlation below 1.5 km.

2.2 MMF retrieval tests

Before describing the different tests that were performed for MMF, we highlight the main problem of MMF retrievals as already reported in the technical note [D3.2] by concentrating only on the partial AOD up to 4 km instead of the profile. If we do so, we see the same picture as for the profiles. Fig.5 shows scatter plots of AERONET AOD with MMF (red) and MAPA (cyan) in the individual directions (panels a to d) and considering the mean of at least 3 directions, as described in the previous section.

Note that the correlations and the Theil-Senn (TS) slopes for each of the two retrieval codes with AERONET AOD are indicated in the panel legends separately for the individual directions and the mean. Note that, as before, we only consider coincidences of available scaled backscatter profiles and valid MAX-DOAS retrievals, although, we could use a less strict coincidence criterion (only requiring AERONET AOD instead of backscatter profiles). However, for consistency, we choose to keep the same criterion. Also note that, although a part of the MAPA flag flags AOD retrievals larger than 2 as invalid, since we consider the partial profiles up to 4 km (as opposed to the reported AOD), values > 2 can occur as valid (qa flag =0) for MAPA.

MMF shows a better correlation with AERONET than MAPA considering each individual direction for the MAX-DOAS measurements, but a large underestimation, with TS slopes around 0.5, where MAPA shows TS slopes often larger than 0.75. This matches the finding reported in [D3.2].

If considering the mean of at least 3 different directions as the representative value from the MAX-DOAS retrieval (see previous section), the correlation of MAPA increases significantly from around 0.69 to 0.8, while the correlation of MMF increases only slightly from around 0.81 to 0.86. This matches the findings for the profile. While the TS slope increases for both MAPA and MMF in the mean of all directions, it remains small for MMF (TS slope=0.56) compared to MAPA (TS slope=0.81), hence the main issue with MMF is a severe underestimation of the partial AOD up to 4 km.

In Table 1, we summarize the MMF retrieval tests carried out as part of WP3. The tests consider different a priori profiles, as well as different settings for the assumed covariance matrix Sa, as well as different values for scaling the measured DSCD. Note that the a priori covariance matrix cannot currently be provided to the retrieval code completely independent from the a priori profile. A higher scale height (column 2) of the a priori also results in a larger covariances at higher altitudes. The total amount can however be adjusted by the Sa scaling factor (column 3) which gives the percentage of the a priori used to build the diagonal elements of the Sa matrix, and the inter-layer correlation by adjusting the correlation length (column 4); this is used to build the off-diagonal elements of the Sa matrix. Therefore, when we scaled down the a priori profile (tests 2, 3 and 9), we scaled up the Sa scaling factor to achieve similar degrees of freedom under equal measurement conditions. Refer to Friedrich et al. (2019) for the implementation details of the parameters, or in a more concise presentation in the ATBD [D3.4].

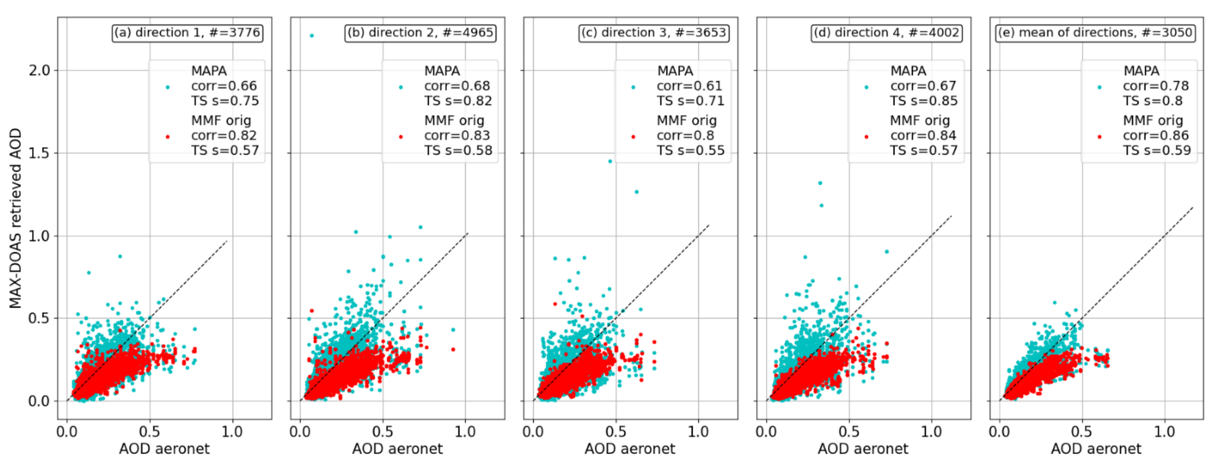


Fig.5: Scatter plots in individual directions (panels a—d) and the mean between the directions (panel e) of retrieved partial AOD (< 4km) from MAX-DOAS measurements, MAPA

(cyan) and MMF (red) with AERONET AOD. The correlations and Theil-Senn slopes are indicated in the legend, the number of valid points is in the title. The 1:1 line is indicated as dashed black line.

Table 1: The tests varied 5 parameters: The scale height (Scale h) used to build the shape of the *a priori* profile, the scaling of the *a priori* that is used to build the diagonal elements on the covariance matrix Sa (Sa diag.), the correlation length that is used to build the off-diagonal elements of the Sa matrix (corr. length), the *a priori* AOD used to scale the *a priori* extinction profile and the scaling factor (SF) used to scale down the measured DSCD that is fitted by MMF. The different background colour separates the separate set of tests (Test 1 and Test 8 belong also to the second set of tests, light grey).

Test #	Scale h/ km	Sa diag./ %	corr. length/ m	AOD a priori	SF (measured DSCD)
original	1	50	200	0.18	1
1	2	60	600	0.18	1
2	2	100	800	0.1	1
3	2	150	1000	0.1	1
4	2	80	1000	0.18	1
5	2	60	600	0.18	0.9
6	2	60	600	0.18	0.95
7	1	60	600	0.18	0.85
8	2	60	600	0.18	0.85
9	2	150	1000	0.1	0.85

As for the tests in the previous section, we only include measurements at times at which ceilometer profiles, as well as AERONET AOD measurements and retrieval profiles (either in the direction in question, or at least in 3 directions if considering the mean) are available. Further, as in the previous section, the statements about the profiles exclude the lower most layer since the extinction in this layer cannot be estimated from the ceilometer backscatter profile in a reliable way.

2.2.1 Investigating the effect of a priori settings

The first set of tests (Test 1 – 4, light grey shading in Table 1) investigates whether with an *a priori* aerosol extinction profile that falls off less steep with height, and hence also giving more freedom from deviation from the aerosol profile at higher altitudes, the mismatch between the AOD (see Fig.5) can be addressed.

In Fig.6, we show a similar plot as in Fig.5 but showing the partial AODs of the default MMF (red) compared to the partial AODs of MMF Test 1 (blue). As can be seen in that figure, the general good correlation R between MMF and AERONET AODs is not affected, and the underestimation is slightly reduced in each direction, as well as when considering the mean of the directions. The Theil-Senn (TS) slope increases from 0.56 to 0.66.

In Table 2, we summarize the TS slopes and correlations of Tests 1—4. It is evident that Tests 1,2 and 4 show similar slopes and correlations. However, Test 3 shows a weaker correlation in all directions, especially in direction 3. This is caused by several outliers, indicating that too much freedom (reduced penalty in cost function for deviation from the a priori profile) is given at higher altitudes, using 150% of the a priori for the construction of the diagonal elements. If a complex real situation (spatial inhomogeneity in horizontal direction as well as temporal inhomogeneity) is modelled with a simple 1D forward model that cannot reproduce this complexity, it can happen that highly oscillatory vertical profiles produce simulated dSCDs that fit the measured dSCDs slightly better than smoother vertical profiles. To avoid these, a better estimate of the true errors that account for temporal and spatial inhomogeneity is needed. If a better way of estimating the true errors (also accounting for time variability during the duration of a scan) on the DSCDs could be found, this might be avoided. However, since the forward model assumes horizontal homogeneity, even if an accurate way to account for time variability (by enlarging the corresponding DSCD errors) was found, the short coming of trying to model a 2D variable field (in radial direction from the instrument) with a 1D layered aerosol profile limits the reproducibility of the true situation. Hence, it is safer to impose a stricter limit on the deviation from the a priori profile. Test 3 will not be investigated any further.

To further investigate where (i.e. from which height regions) the improvement of TS slope for tests 1, 2 and 4 originates from, it is instructive to compare the mean and median profiles of these tests to the ones of MMF (and MAPA and the scaled back-scatter profile), as shown in Fig.3, as well as looking at the correlation as a function of height, as was shown in Fig.2. Appendix A.1 also shows scatter plots for the first 8 layers between the original MMF and MMF Test v1.

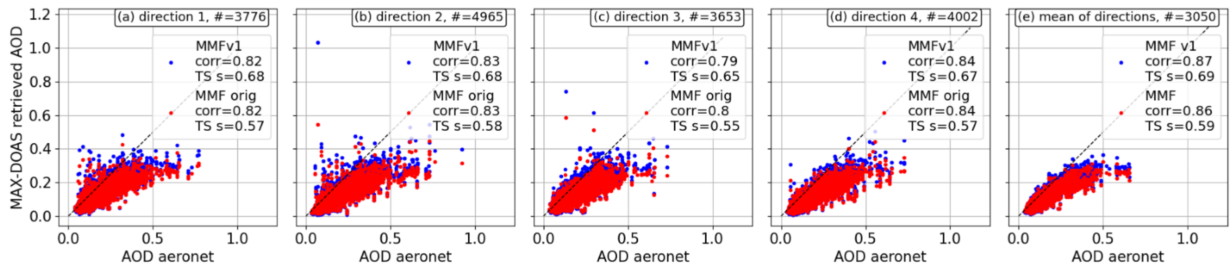


Fig.6: Scatter plots in individual MAX-DOAS directions (panels a—d) and the mean between the directions (panel e) of retrieved partial AOD (< 4km) from MAX-DOAS measurements with MMF Test 1 (blue) and MMF original (red) with AERONET AOD. The 1:1 line is indicated as dashed black line. Note that the TS slope of MMF Test 1 (blue) is increased in each direction as well as in the mean of the directions. The correlation shows some changes in the individual directions but is not heavily affected overall.

In Fig.7, we show the corresponding profiles up to 4 km (median profiles in panels b and c, mean profiles in panels d and e), as well as the correlation with the comparison profile (in panel a) for Tests 1 and 2 (Test 4 is almost identical to Test 1 and is omitted in the plot to increase readability) together with the default original MMF settings in FRM4DOAS. In dashed lines, we show the corresponding a priori profiles.

Table 2: Summary of AOD correlation analysis for the first set of MMF tests: The first column in each test shows the correlation R, the second column of each test shows the TS slope s. For each test, the individual directions 1—4 are shown, as well as the mean of the direction in the last row.

	original		**Test 1**		**Test 2**		**Test 3**		**Test 4**	
	R	s	R	s	R	s	R	s	R	s
Dir 1	0.82	0.57	0.82	0.68	0.81	0.67	0.80	0.69	0.81	0.70
Dir 2	0.83	0.59	0.83	0.68	0.79	0.66	0.77	0.69	0.81	0.69
Dir 3	0.80	0.55	0.79	0.65	0.80	0.63	0.72	0.65	0.79	0.68
Dir 4	0.84	0.57	0.84	0.67	0.83	0.66	0.82	0.68	0.84	0.69
mean	0.86	0.59	0.87	0.69	0.87	0.67	0.85	0.69	0.87	0.70

We choose to show the mean and median extinction profiles once in linear scale (panels b and d) and once in log scale (panels c and e) to show more clearly that even at 4 km, due to the different distribution of freedom in the profile (the covariance matrix scales with the *a priori* profile; if the *a priori* profile falls off slower with height, then this is also true for the covariances), the retrieved profile deviates from the a priori profile. In either of the two representations, it is visible that albeit the vastly different extinctions of the a priori profiles in the lowest 600 m, the retrieved profiles agree and are much lower than the scaled backscatter profile (mind the lowest layer, see above). Further, in all cases, the mean and median profiles at higher altitudes are always below the respective a priori profile. Hence the underestimation of MMF cannot be explained with a too low a priori since the information from the measurements drives the retrieval to lower extinctions.

However, choosing an *a priori* that falls off slower compared to the default one, seems to, at least for MAINZ, match better with reality and hence is recommended. Allowing the covariance matrix to fall off equally slower with height, increases the correlation at higher altitudes (see Test 1 in Fig.7a). However, as is evident from Test 3, and to a lesser extend Test 2, higher layers are sensitive to situations that are not reproducible with a 1D model that is constant over the time of a scan measurement. Similar findings (too low extinction retrievals) were also found for MAPA when using a fixed scaling factor applied to the measured DSCDs of 1 (Beirle et al. 2019). The key points from the first set of tests are:

- Changing the a priori settings does not change the aerosol extinction profile in the lowest 600 m and does hardly affect the correlation of the retrieval with the scaled backscatter profile at those heights. Hence changing a priori settings does not help to increase the partial AOD in this height range.
- Choosing a more adequate a priori profile and covariance matrix that allow for deviations from the a priori at larger altitudes, slightly decreases the underestimation of MMF
- Allowing for too large deviations at high altitudes makes the retrieval prone to outliers

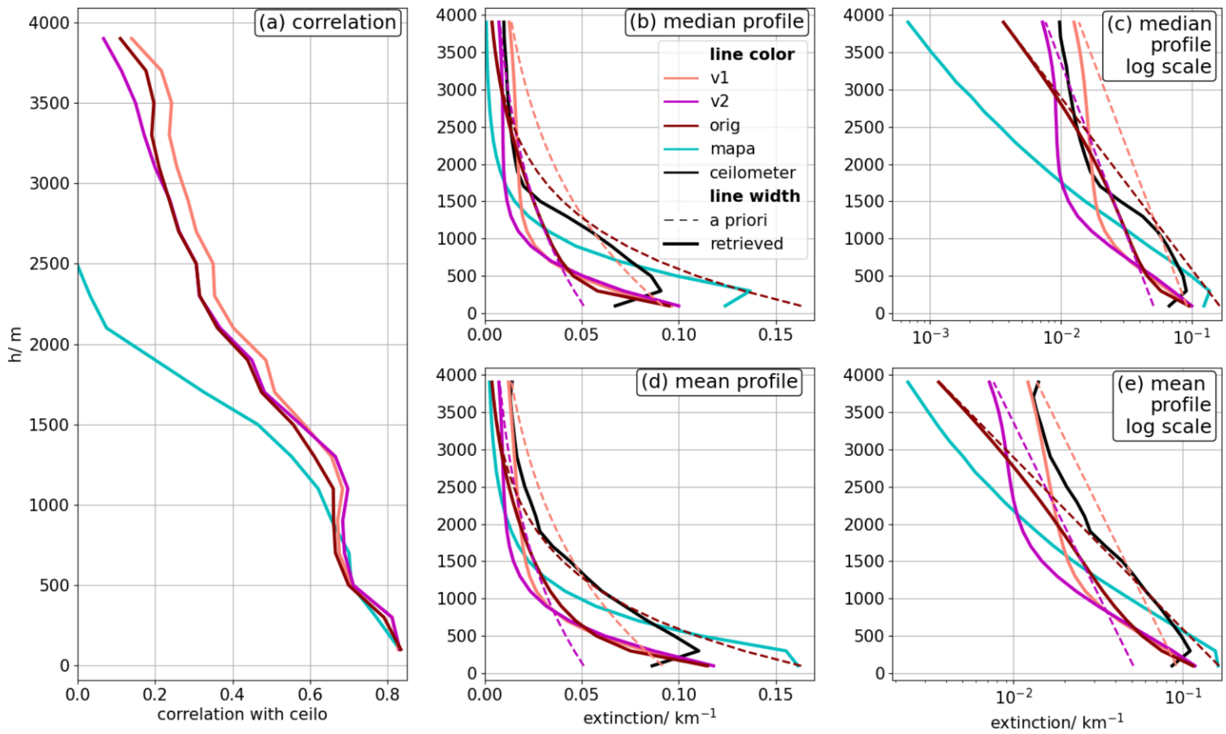


Fig.7: Correlation between scaled ceilometer profiles and MAPA (cyan) and profiles from MMF Tests 1, 2 and the original FRM4DOAS implementation (orange, magenta and brown, respectively) (panel a) and median (panels b and c) and mean (panels d and e) profiles in linear (panels b and d) and log scale (panels c and e) together with the corresponding MMF Test *a priori's* (dashed lines). Note that the mean and median profiles of all test cases agree in the lowest 600 m and are hence much lower than the a priori profiles of Test v1 and the original FRM4DOAS implementation. For those two tests, the retrieved profiles (both mean and median) are below the a priori profiles at all heights.

The results of these tests motivated us to perform a second set of tests for MMF, using the same retrieval settings as in Test 1, but varying the scaling factor (SF) applied on the measured DSCD that is passed to the retrieval code for fitting, see Table 1.

2.2.2 Investigating the effect of lowering the scaling factor on the measured DSCDs

Test 1, Test 5, Test 6 and Test 8 (compare Table 1) all use the same retrieval settings for MMF. However, the measured DSCD SF is changed. While Test 1 uses no scaling factor (i.e. SF=1) on the measured DSCD (as the default FRM4DOAS MMF settings), Test 5 uses 0.9, Test 6 uses 0.95 and Test 8 uses 0.85. The AOD correlation analysis is summarized in Table 3.

Table 3: Summary of AOD correlation analysis for the second set of MMF Tests: The first column in each test shows the correlation *R*, the second column of each test shows the TS slope *s*. For each test, the individual directions 1—4 are shown, as well as the mean of the direction (last row).

	Test 1		Test 5		Test 6		Test 8	
	R	s	R	s	R	s	R	s
Dir 1	0.82	0.67	0.81	0.74	0.82	0.71	0.81	0.77
Dir 2	0.83	0.68	0.82	0.77	0.83	0.72	0.79	0.81
Dir 3	0.79	0.65	0.79	0.72	0.79	0.68	0.78	0.76
Dir 4	0.84	0.67	0.84	0.77	0.84	0.72	0.83	0.82
mean	0.87	0.69	0.87	0.77	0.87	0.73	0.86	0.81

It is evident that the TS slope *s* increases (for all directions separately, but also for the mean in the directions) with decreasing SF. For Test 8 (the lowest tested SF, 0.85), the slope for the mean increases to 0.81 which is comparable to that of MAPA (s=0.8, see Fig.5). The correlation R from Test 8 (0.86 considering the mean of the directions) is slightly better than that of MAPA (0.80).

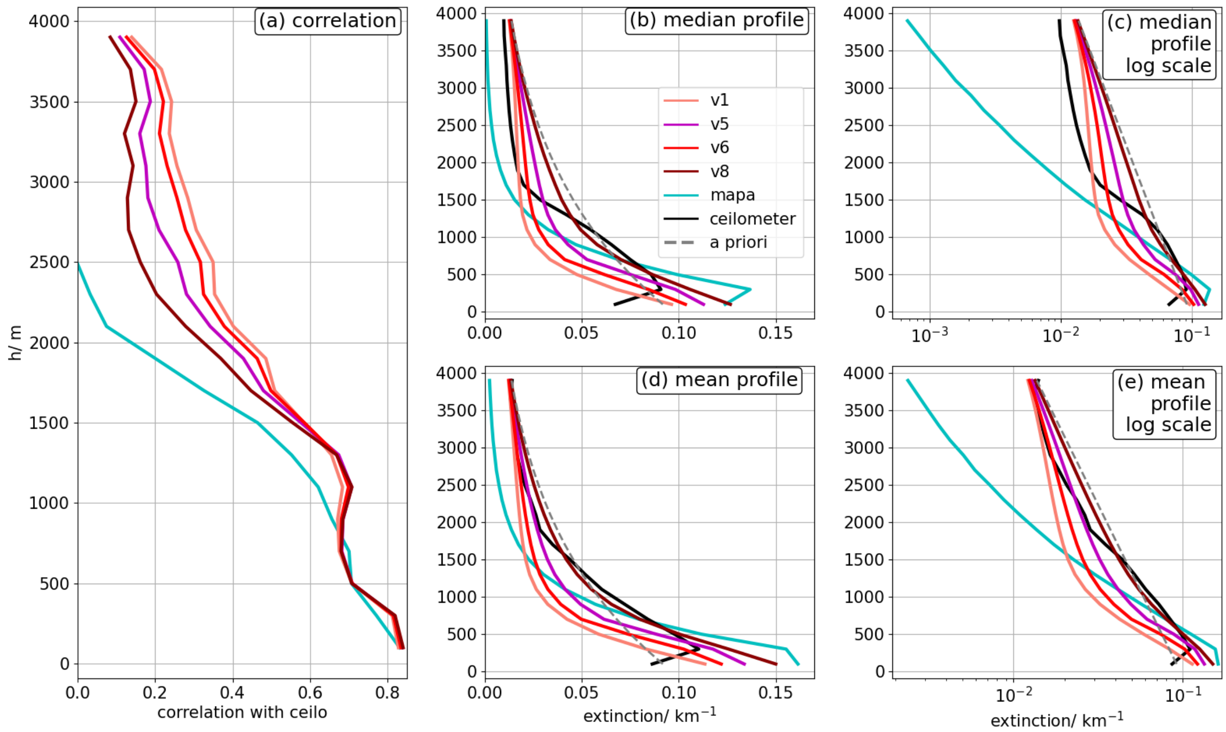


Fig.8: Correlation between scaled ceilometer profiles and retrieved profiles: MAPA (cyan), MMF Tests 1, 5, 6 and 8 (orange, red, magenta, and brown, respectively) (panel a) and median (panels b and c) and mean (panels d and e) profiles in linear (panels b and d) and log scale (panels c and e) together with the MMF Test a-priori (dashed line). Note the almost identical correlation with the scaled ceilometer profiles up to about 1.5 km.

In Fig.8, we show the correlation of the retrieved extinction profiles with the scaled ceilometer profile for those 4 tests (panel a), as well as the median (panels b and c) and mean (panels d and e) profiles of those tests and MAPA and the scaled ceilometer profile. All 4 tests show an almost identical correlation with the scaled ceilometer profile up to a height of approximately 1.5 km. However, the mean and median profiles are vastly different and the total AOD (< 4km) scales almost inversely with the scaling factor applied on the measured DSCDs in the scaling factor range tested, see Fig.9. It is noteworthy to mention that the correlation above 1.5 km is reduced with decreasing SF.

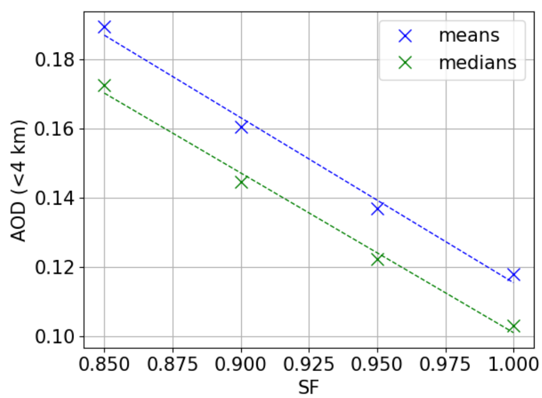


Fig.9: Median (green) and mean (blue) AODs from retrieved profiles applying a different SF to the measured DSCDs. Here, SF 0.85 corresponds to Test 8, SF 0.9 corresponds to Test 5, SF 0.95 corresponds to Test 6, and SF 1 corresponds to Test 1. Note the almost inverse scaling of AOD with SF.

While the total AOD (< 4km) shows an almost linearly inverse scaling with the SF, there are significant differences between which parts of the profile are affected. Looking at the percentual difference of extinction between the tests with SF < 1 (Tests 6, 5 and 8) with the one with SF = 1 (i.e. Test 1), compare Fig.10, shows that the largest increase happens at around 1 km altitude: Applying a scaling factor of 0.85 results in an increase of extinction of more than 100% around 1 km of altitude, while the increase close to the surface is only a few tens of percent.

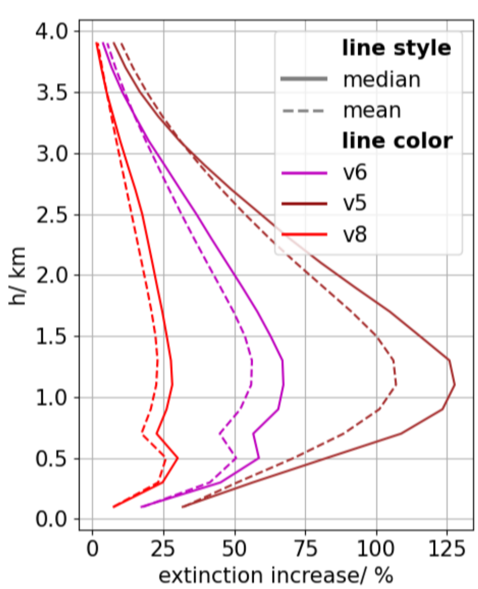


Fig.10: Percentual differences in extinction profiles for Tests 5,6 and 8 compared to Test 1. While the AOD scales inversely with SF, the effect on the profile is highly height dependent: it is much stronger in the range 0.5—2 km than at the surface. Note that the median shows a larger increase than the mean.

The key points from this set of tests are:

- The correlation between the scaled backscatter profile and the retrieved profile in the lowest most 1.5 km is hardly affected by scaling the measured DSCD
- The retrieved AODs scale almost inversely with the applied scaling factor.

- The higher the applied scaling factor, the lower the correlation at altitude above ~1.5 km.
- A scaling factor of 0.85 removes the underestimation in AOD and a similar agreement with the AERONET AOD is reached as for MAPA.

2.2.3 Investigating the effect of a priori settings when using a fixed scaling factor on the measured DSCDs

The last set of tests (Tests 7 – 9, indicated in Table 1 by a dark grey background) concentrated on keeping the DSCD scaling factor (applied on the measured DSCD) constant at 0.85 and to vary the a priori conditions (both profile and Sa), similar as was done in the first set of tests, but now with a different SF. Table 4 summarizes the correlation analysis of the AODs w.r.t the AODs from AERONET.

Table 4: Summary of AOD correlation analysis for the first set of MMF tests: The first column in each test shows the correlation R, the second column of each test shows the TS slope s. For each test, the individual directions 1—4 are shown, as well as the mean of the direction in the last row.

	original		Test 7		Test 8		Test 9		MAPA	
	R	s	R	s	R	s	R	s	R	s
Dir 1	0.81	0.54	0.81	0.74	0.81	0.77	0.77	0.81	0.66	0.75
Dir 2	0.82	0.52	0.80	0.77	0.79	0.81	0.51	0.85	0.68	0.82
Dir 3	0.79	0.49	0.78	0.73	0.78	0.76	0.72	0.80	0.61	0.71
Dir 4	0.83	0.53	0.83	0.77	0.83	0.82	0.67	0.86	0.67	0.85
mean	0.86	0.56	0.86	0.77	0.86	0.81	0.83	0.84	0.78	0.80

All tests show similar slopes for the AOD. This slope is considerably improved compared to the original implementation. However, Test 9, having the same a priori conditions as Test 3, shows less good of a correlation. As already discussed for Test 3, this is due to too much freedom at high altitudes which leads to outliers in case of real situations that cannot be modelled by a time constant 1D model. However, Test 9 also shows the largest slopes, exceeding those of MAPA (for convenience included in the last column) in all directions, as well as in the mean. The correlation of the mean of the directions with AERNET AOD is slightly higher than that of MAPA.

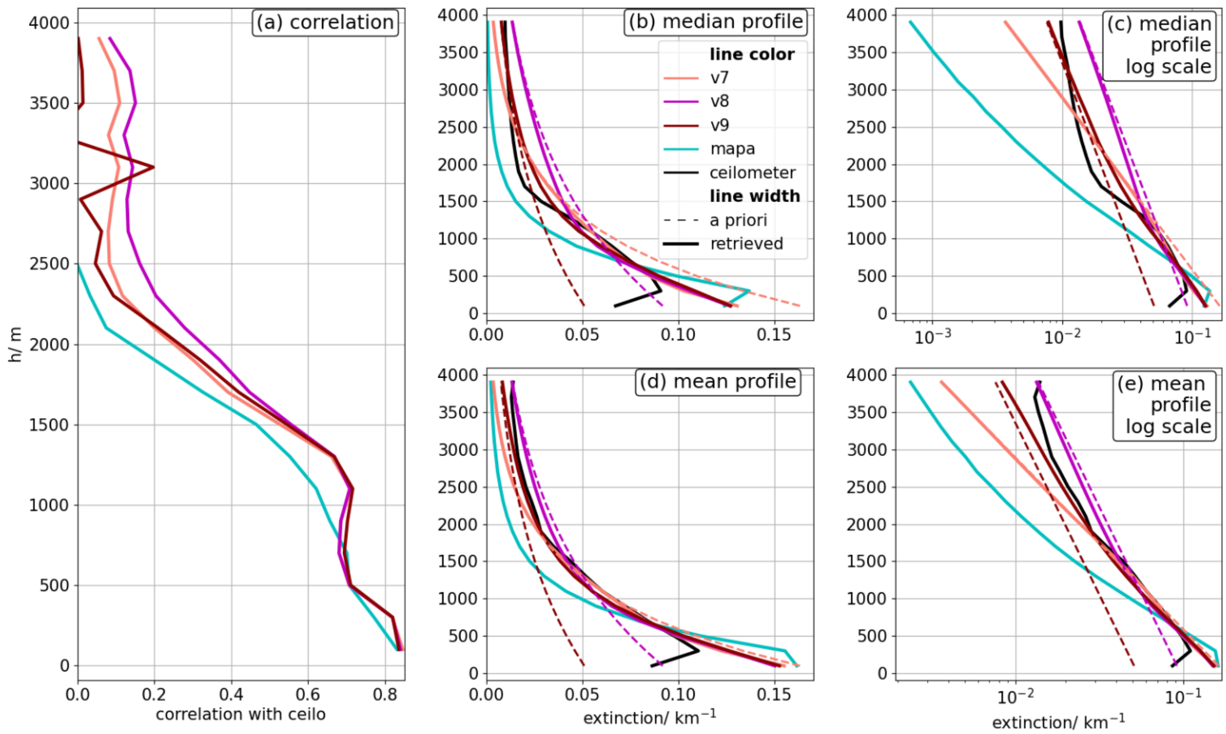


Fig.11: Correlation between scaled ceilometer profiles and MAPA (cyan) and profiles from MMF Tests 7, 8 and 9 (orange, magenta and brown, respectively) (panel a) and median (panels b and c) and mean (panels d and e) profiles in linear (panels b and d) and log scale (panels c and e) together with the corresponding MMF Test *a priori*'s (dashed lines). Note the good agreement between both the median and mean profiles of the different MMF tests, as well as the almost identical correlation with the scaled ceilometer profiles.

In Fig.11, we show the correlation of the retrieved MAPA (cyan) and MMF Tests 7, 8 and 9 (orange, magenta and brown, respectively) with the scaled ceilometer profile (panel a, ceilometer profile in black in the remaining panels), as well as the median (panels b and d) and mean (panels c and e) profiles individually. We also include the corresponding a-priori profiles for the different MMF tests in linear scale (panels b and d) as well as in log-scale (panels c and e). As for the first set of tests, the a priori profiles span a wide range of extinction in the lowest 4 km. Still, both the median as well as the mean retrieved profiles of the different MMF tests agree very well in the lowerest 1.5 km. They also agree well with the scaled ceilometer profile in this height range. Also, the correlation (panel a) in the lowest 1.5 km is hardly affected by the change in a priori settings. As is evident from Table 1, not only the a priori profile but also the constructed a priori covariance matrix is vastly different between the tests. Fig.12 compares the AOD scatter plots of MAPA-AERONET AOD and MMF Tests 8 – AERONET AOD (similar as in Fig.5

for the default MMF retrieval). Using a SF 0.85 and an a priori profile with a scale height of 2 km shows less scatter and often a better slope than MAPA in individual directions as well as in the mean of the directions.

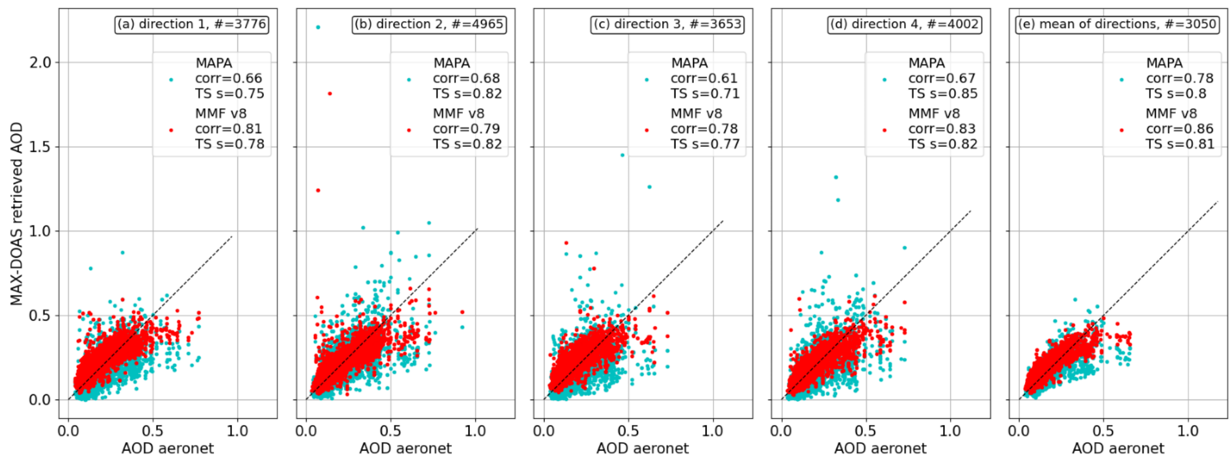


Fig.12: Scatter plots in individual MAX-DOAS directions (panels a—d) and the mean between the directions (panel e) of retrieved partial AOD (< 4km) from MAX-DOAS measurements with MAPA (cyan) and MMF Test 8 (red) with AERONET AOD. The correlations and Theil-Senn slopes are indicated in the legend. The 1:1 line is indicated as dashed black line. The number of valid coincidences is given in the title.

We conclude therefore:

- the true information content in the MAX-DOAS measurements about the lowest 1.5 km is high enough to estimate the extinction profile and the resulting profile is independent of the a priori profile
- A scaling factor of approximately 0.85 seems necessary to retrieve the correct order of magnitude of extinction in the lowest most 1.5 km.
- A covariance matrix that offers a too low constraint at high altitudes is not recommended

Good retrieval settings are hence:

- Using a scaling factor of around 0.85
- Using an *a priori* profile that matches closely the median extinction profile above 1.5 km. A scale height of 2 km and an AOD of 0.1 (as in TEST 9) -- 0.18 (as in Test 8) seems to be a viable choice for MAINZ.
- Using a Sa matrix that does not use a constrained that is not too loose above 1.5 km, even though, under favourable conditions (temporal and horizontal spatial homogeneity) some information about aerosol extinction at higher altitudes might be

available. For MAINZ, using 600 m correlation length and a scaling on the diagonal of 60% of the a priori (if the a priori is chosen with 2 km scale height and 0.18 AOD) seems appropriate, using 150 % and a correlation length of 1 km when using an a priori AOD of 0.1 seems too high.

2.2.4 Final considerations

Before presenting the results for the other stations, we re-introduce on top of the MAPA flag, the following components from the MMF flag:

- Check on the RMS: We require the RMS (in units of DSCD error) between simulated and measured DSCD to be less than 3 (more than 2 gives a warning).
- Check on the DOF: the DOF needs to be larger than 1.3 (less than 1.6 gives a warning).
- Consistence check: running the retrieval with an a priori AOD with 1/3 of the original a priori AOD, the retrieved AOD must agree within ± 30 % (more than ± 10 % gives a warning)
- AOD limit: The retrieved AOD should be smaller than 5 (more than 2 gives a warning)
- No heavy oscillations: We require the third derivative of the profile to be less than 0.5.
- No extreme top-down profile: We require more than 75% of the extinction to be below 3 km.

Introducing these checks on top of MAPA flag, reduces the number of valid retrievals (for MMF) by about 2%, however, only about 0.05% of the data that has backscatter and AERONET data is affected. In Fig.13, we reproduce Fig.12 but applying these extra flags. This does not greatly influence the results for MAINZ (since we only consider times at which scaled backscatter profiles are available, only two times for direction 2 are filtered out). In the remaining part of the report, MAPA data only uses the MAPA flag.

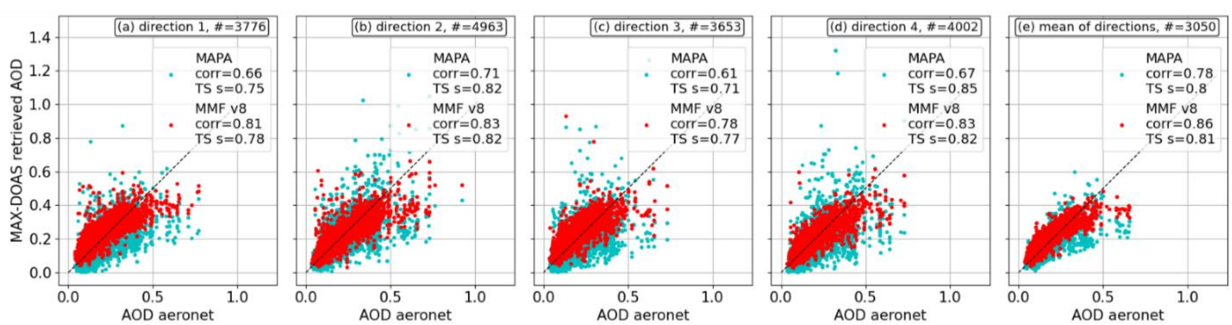


Fig.13: As Fig.12, but with adding the above mentioned MMF flags on top of the MAPA flag.

Fig.8 and Fig.11 show that applying a scaling factor (different from 1) to the measured DSCD removes the already low information at altitudes above 1.5 km even further. Hence, it might be insightful to make a similar plot as Fig.13 but only considering the partial AOD up to 1.5 km (middle height of layer, so up to the 8th layer). We present this in Fig.14. Note the excellent slope and the good correlation between the partial AERONET AOD and the partial AOD from both retrieval codes when considering the mean of the directions, as well as the good agreement between the extinction up to this height retrieved from the two codes.

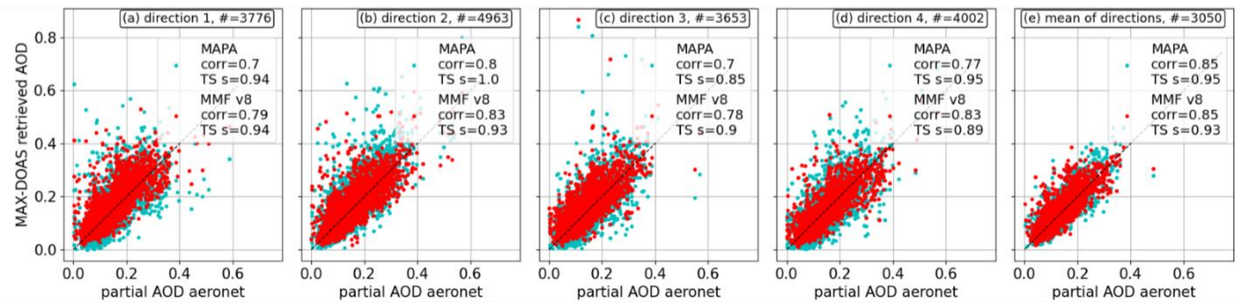


Fig. 14: As Fig.13 but showing only partial AODs up to 1.5 km (layer 8).

3 Other stations in the UV

As for Mainz, extinction profiles for Uccle, Xianghe and Cabauw were retrieved with the retrieval settings that were found to be more favourable for MAINZ to test whether the same settings serve for these stations.

The settings are summarized in Table 5.

Table 5: original and new retrieval settings for MMF

Test #	Scale h/ km	Sa diag./ %	corr. length/ m	AOD a priori	SF (measured DSCD)
original	1	50	200	0.18	1
new	2	60	600	0.18	0.85

Contrary to MAINZ, the stations investigated in this section do not offer the MAX-DOAS measurements in different directions at the same time, hence a similar treatment as for MAINZ (calculating the mean of different directions for comparison with AERNET and scaled ceilometer data) is not possible. However, for all three stations, there are measurements in different directions, and they are all treated together, i.e. not split up in different directions.

3.1 Uccle and Cabauw

If we apply the same filtering as described above and consider the same partial AOD (i.e. up to 1.5 km, layer 8) and perform the analysis with the new settings, then we see somewhat comparable results for Uccle and Cabauw (in the UV) as was achieved for Mainz. This is shown in Fig.15, comparing the new settings (v8, left panels in blue), MAPA (middle panel in cyan) and the original MMF (right panel in red). The latter two use the MAPA flagging, the first uses the combined MAPA and MMF flagging as outline above.

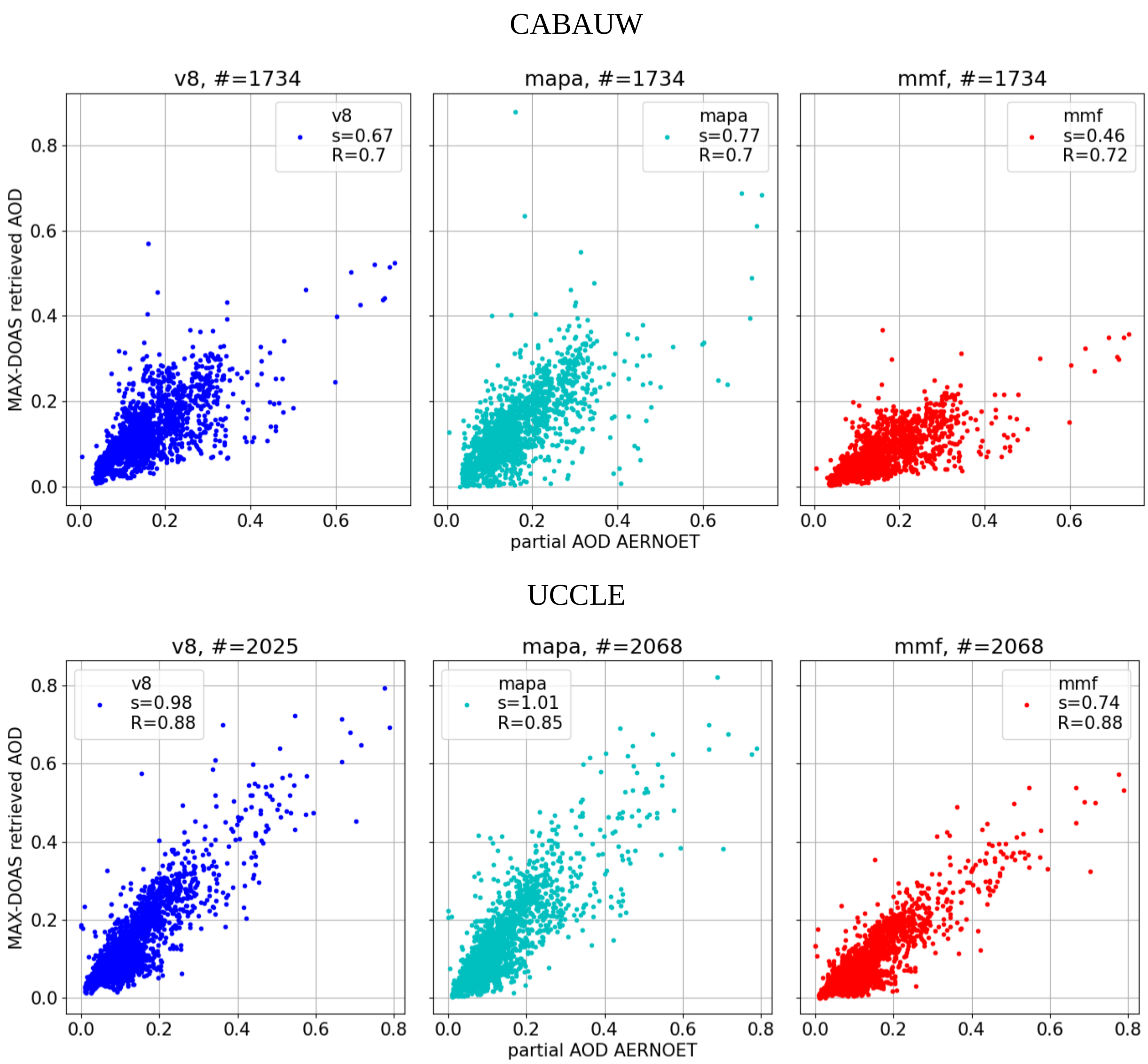


Fig.15: Partial AOD (up to 1.5 km, layer 8) from MAX-DOAS retrievals compared to AERONET, as Fig. 13, but for stations Cabauw (upper panels) and Uccle (lower panels).

However, while the slope for Uccle is close to 1 for both MAPA and MMF, the slope for Cabauw is substantially below 1, especially for MMF, even if considering the partial AOD up to 1.5 km.

For Mainz (compare Fig.13 up to 4 km, and Fig.14 up to 1.5 km), the change in correlation between considering the full or partial AOD was minor (MAPA increased from 0.78 to 0.85, MMF decreased from 0.86 to 0.85), the effect on the slope was considerable (for MAPA from 0.8 to 0.95 and for MMF from 0.81 to 0.93).

If considering the AOD up to 4 km (compare Fig.13 for Mainz) for Uccle and Cabauw, see Fig.16, then we see a similar picture for Uccle: considering the partial AOD up to 1.5 km instead of 4 km increases the slope from 0.85 to 0.98 (MMF) and from 0.86 to 1.01 (MAPA). The correlation is largely unaffected and increases from 0.81 to 0.88 (MMF) and from 0.83 to 0.85 (MAPA).

For Cabauw, the slope (MMF: from 0.50 to 0.67, MAPA: from 0.62 to 0.77) and the correlation (MMF: from 0.59 to 0.70, MAPA: from 0.61 to 0.70) also increase when considering the AOD up to 1.5 km instead of up to 4 km but are overall much lower.

MAPA retrieves aerosols above 4 km, hence we can also compare the total AOD instead of the partial AOD up to 4 km. In that case, we do not any longer need the existence of a backscatter profile, hence this increases the number of coincidences. For MMF, nothing changes since the reported AOD equals the AOD up to 4km.

Doing so does not change the slope for MAPA for Cabauw, it stays at 0.62, however the correlation reduces from 0.61 to 0.54. The contribution of the extinction profile above 4 km to the total AOD seems hence not to fill the gap between MAPA retrieved and sun-photometer measured AOD. Likewise, for Uccle, the slope only increases slightly from 0.86 to 0.87 while the correlation decreases from 0.83 to 0.82.

For MMF (only change is the slightly larger number of data since the MMF AOD equals the partial AOD up to 4 km), the slope values are 0.5 for Cabauw (compared to 0.5) and 0.84 for Uccle (compared to 0.85) and the correlations are 0.61 (compared to 0.59) for Cabauw and 0.8 (compared to 0.81) for Uccle. These changes are all rather insignificant. All these results are summarized in Table 6 (for Cabauw) and Table 7 (for Uccle).

It was further investigated whether it might be possible to relax (for MMF) the MAPA flag to also include warnings (and the additional MMF flag). This increased the number of valid values to 2228 (for Cabauw) while reducing the slope to s=0.46 and the correlation to R=0.56 and for Uccle to 2168 while reducing the correlation to 0.78 with not changing the slope (s=0.85).

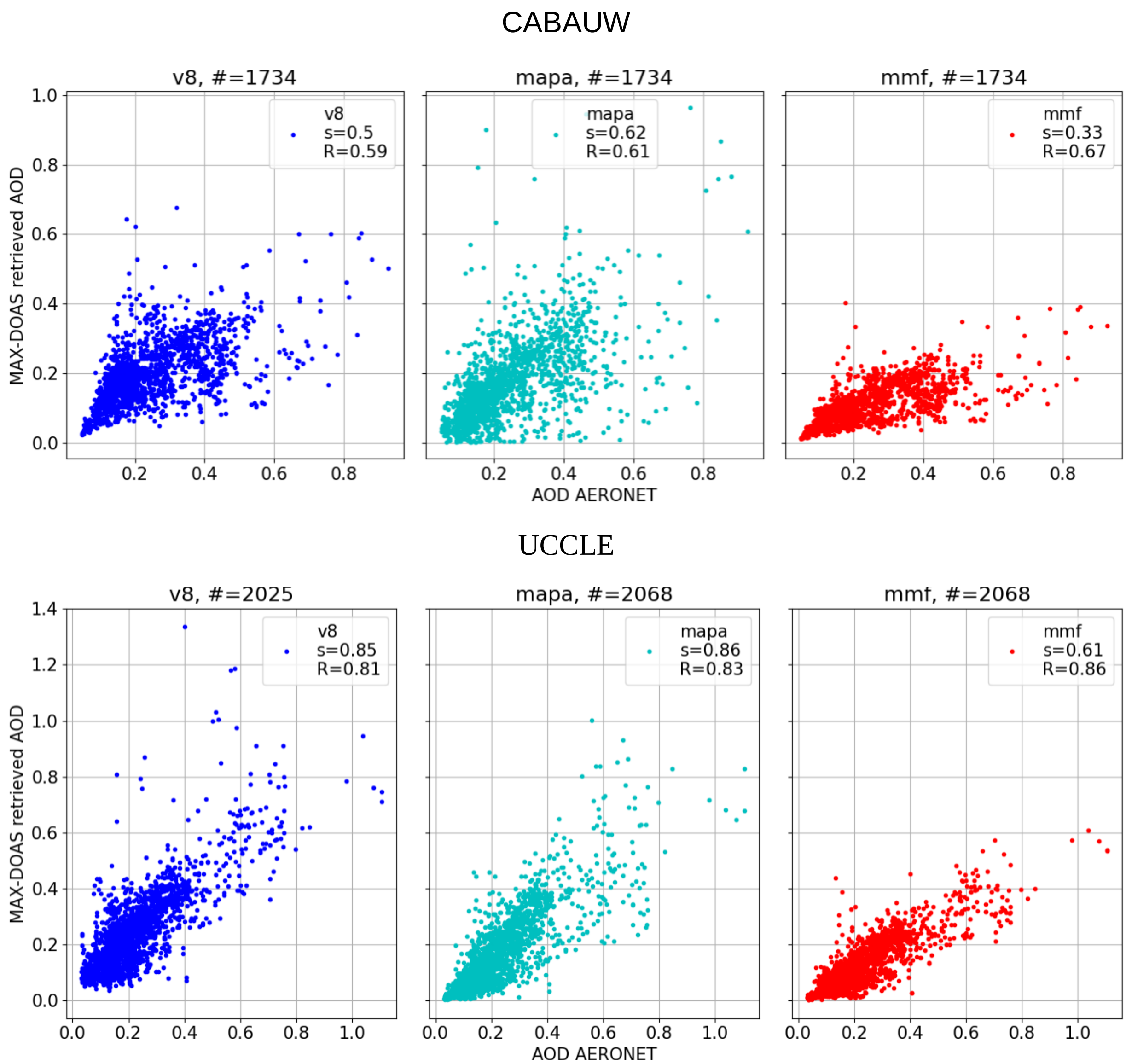


Fig. 16: Cabauw (upper panels) and Uccle (lower panels) MAX-DOAS retrieved AOD up to 4 km vs. AOD from AERONET.

Table 6: Overview of correlation analysis for Cabauw

	Original MMF		Test 8 MMF		MAPA	
	R	s	R	s	R	s
part. AOD 1.5km	0.72	0.46	0.70	0.67	0.70	0.77
part. AOD 4km	0.67	0.33	0.59	0.5	0.61	0.62
AOD	0.69	0.33	0.61	0.5	0.54	0.62

Table 7: Overview of correlation analysis for Uccle

	Original MMF		Test 8 MMF		MAPA	
	R	s	R	s	R	s
part. AOD 1.5km	0.88	0.74	0.88	0.98	0.85	1.01
part. AOD 4km	0.86	0.61	0.85	0.81	0.83	0.86
AOD	0.86	0.61	0.84	0.80	0.82	0.87

3.2 Xianghe

For Xianghe, there is no ceilometer data available (see [D3.2]), hence we only perform comparisons on the AOD only, so for MMF this equals the AOD up to 4 km, but for MAPA, we use the total AOD value reported in the files. A corresponding figure for comparing the MAX-DOAS retrieved AOD to the AERONET AOD is presented in Fig.17 and summarized in Table 8.

While for Uccle, the additional flag for MMF filtered an additional ~2 % of the data (as was found for Mainz), for Xianghe, slightly more than 3% are additionally invalid for MMF. Both the slope as well as the correlation of MAX-DOAS retrieved and AERONET is comparable for MAPA (s=0.73, R=0.88) and MMF (s=0.71, R=0.91). Further, it was found for MMF that relaxing the MAPA flag to also include warnings (see Sect. 3.1), increases the number of valid points to 4250 while reducing the slope and correlation only slightly (to s=0.68 and R=0.87).

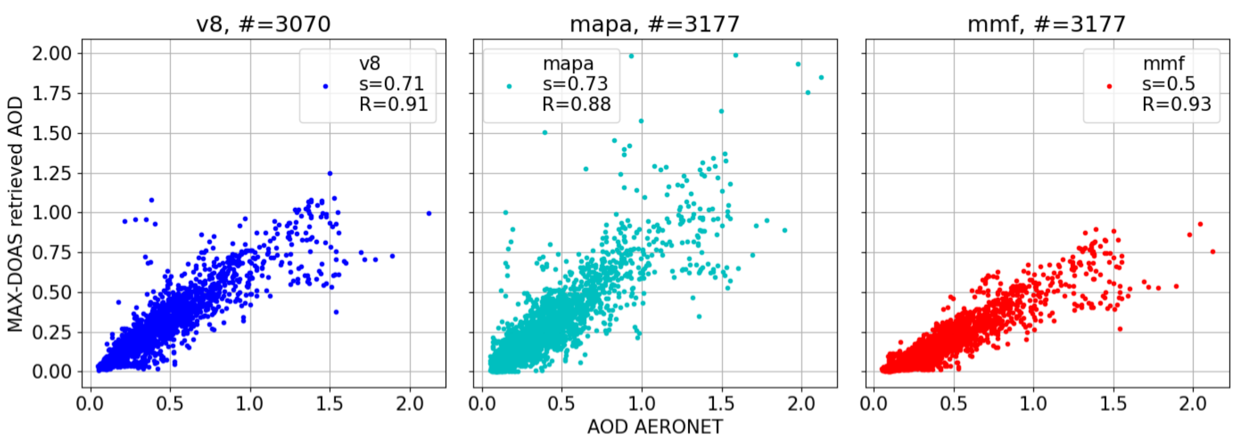


Fig. 17: MAX-DOAS retrieved AOD vs. AOD from AERONET at Xianghe, from left to right: MMF v8, MAPA and original MMF.

Table 8: Overview of correlation analysis for Xianghe

	Original MMF		Test 8 MMF		MAPA	
	R	s	R	s	R	s
part. AOD 2km	n/a	n/a	n/a	n/a	n/a	n/a
part. AOD 4km	n/a	n/a	n/a	n/a	n/a	n/a
AOD	0.93	0.5	0.91	0.71	0.88	0.73

For MAINZ (median of directions), using the stricter MAPA flag combined with MMF flag yielded 3050 coincidences and a correlation of 0.86 and a TS slope of 0.81. Relaxing the MAPA flag increases the number of coincidences to 4334 but reduces the correlation to 0.83 and the slope to 0.78.

Hence, we conclude that as MMF flag (in the UV), a flag built from MAPA warning (i.e. using MAPA qa-flag < 2) and the above described new MMF works well for Mainz, Xianghe, Uccle and Cabauw.

4 Stations in the visible

4.1 Uccle

[D3.2] found that for the MMF VIS retrieval, the MAPA flag did not work well, and a new flag was added to the original MMF flag to flag-out retrievals with oscillations. This flag was adopted above for UV (see MMF flag description above and [D3.2]). In this section, when showing MMF original, this flag was applied. The new versions use the same flag as used for UV in previous sections, i.e. MMF flag as described above and MAPA warning. Further, [D3.2.] found that the underestimation of MMF in VIS is less than for UV.

Applying the same retrieval settings as indicated in Table 5 (as “new”) to UCCLE VIS, results in an improved slope for MMF, but a reduced correlation. In Fig.19, we show the equivalent plot as was shown in the lower panel of Fig.16 (Uccle UV) but for the retrieval in the visible.

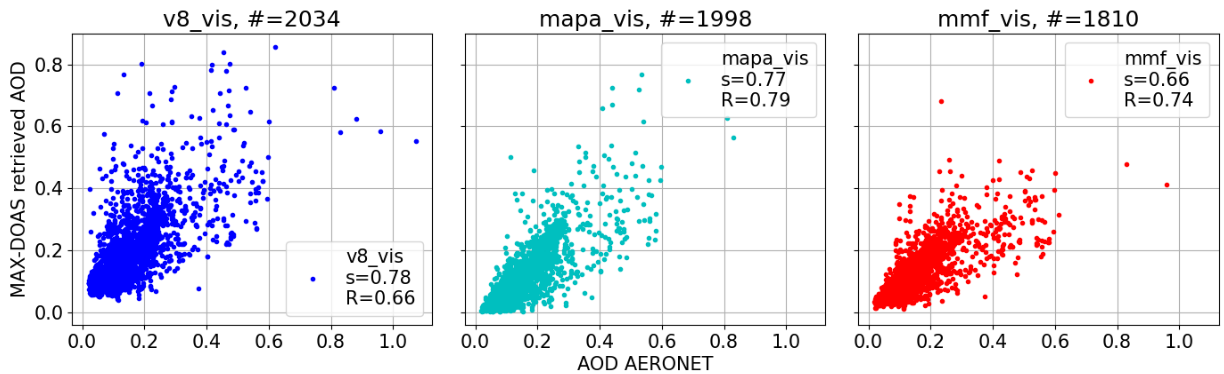


Fig. 19: As Fig. 16 upper panel, but VIS channel for Uccle.

These settings are not improving the retrieval, especially low AOD conditions are overestimated. Comparing the distribution of the MAPA fitted SF (only valid co-incidences between UV and VIS measurements) shows that the distribution of SF in VIS is quite different from the SF distribution in the UV, see Fig.20.

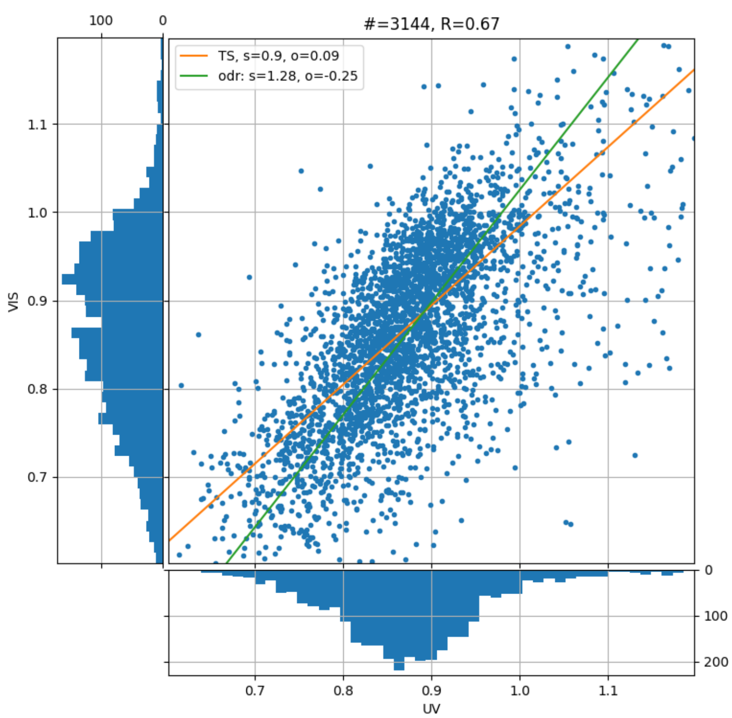


Fig. 20: Comparison of MAPA fitted scaling factors of coincidences of valid UV and VIS retrievals at Uccle. The center panel shows a scatter plot of SF-UV on x and SF-VIS on y. The attached plots left and below show histograms of VIS and UV SF, respectively. Note that the SF distribution in VIS is flatter than the SF distribution in UV and shows a double peak.

While the SF distribution in the UV is concentrated around 0.85 (the value chosen for the retrieval tests), the distribution of SF in the visible is much flatter with more very small (<0.8) and large (>1.0) SF. Motivated by this finding, a test run with applying MAPA fitted SF (limited to max SF=1.2 and min SF=0.7) was performed. We show the resulting scatter plot (only the new MMF version) in Fig.21. Note that the correlation improves, as well as the number of valid points increases.

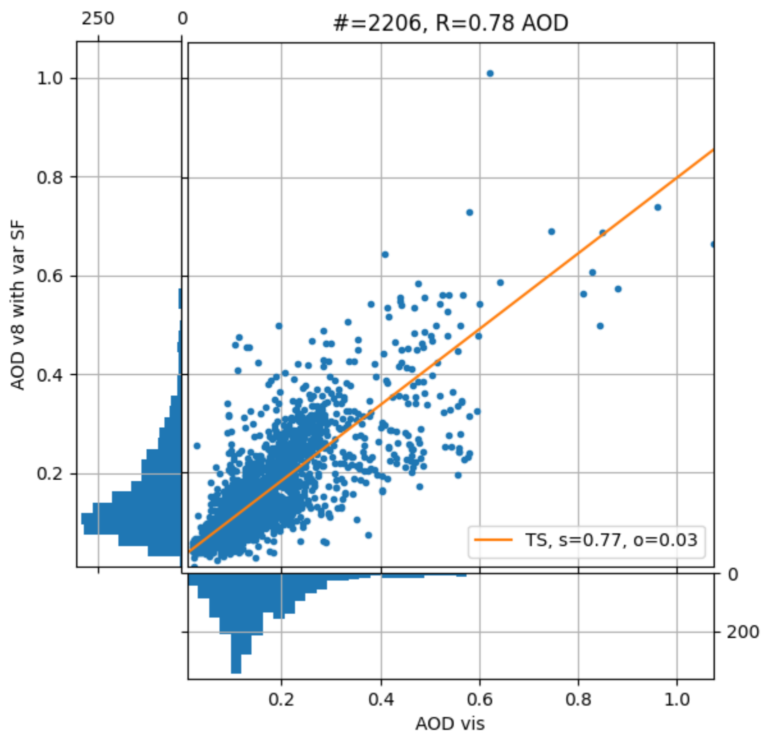


Fig. 21: Scatter plot of MMF retrieved AOD, applying the MAPA variable scaling factor on the measured DSCDs before retrieval, against AERONET AOD for Uccle VIS. Note the improved correlation, the diminished offset and the increased number of valid points compared to Fig.19.

It is instructive to compare this version of MMF retrieval (i.e. with applying the fitted SF from MAPA to the measured DSCD before retrieval) to MAPA. In Fig.22, we show the scatter plot of the partial AOD up to 2 km (left), and the scatter plot of the AOD (what is reported as AOD) in the right panel. Note the almost identical MAPA- and MMF- retrieved partial AODs up to 2 km. We can conclude that if the same scaling as fitted by MAPA is pre-applied to the measured DSCDs before retrieving the extinction profile with MMF, the extinction retrieved in the lowest 2 km agree almost perfectly between MAPA and MMF for those retrievals where both codes yield a valid retrieval. This yields 5576 coincidences that show a correlation of 0.97 and a slope of 0.97 considering the lowest 2 km.

If we drop the (strict qa-flag =0) MAPA flag and apply the new MMF flag on both (not shown), we have 7131 coincidences a correlation of 0.92 and a TS slope of 0.96; if we drop the MMF flag (not shown) and only use the (strict) MAPA flag, we have 5703 coincidences, a correlation of 0.94 and a slope of 0.97.

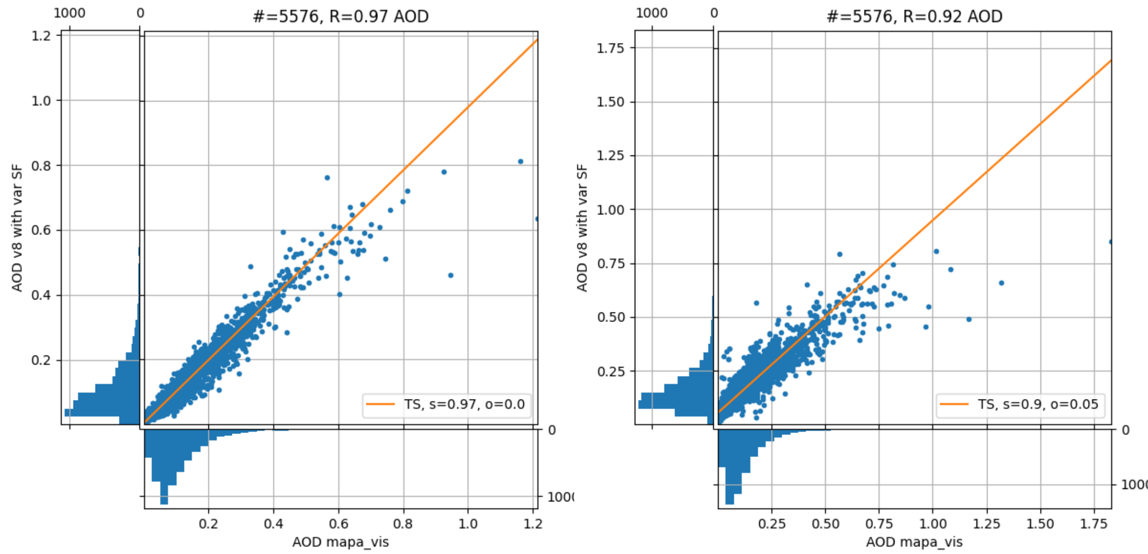


Fig. 22: Scatter plot between MAPA (x-axis) and MMF (y-axis) retrieved partial AOD (up to 2km, left) and AOD (right) for UCCLE VIS. Note the very high degree of correlation and slope considering only the lowest 2 km.

4.2 Xianghe

For Xianghe VIS, we tested the approach of retrieving the aerosol extinction with MMF from the measured DSCDs on which the MAPA fitted SF were applied, as performed in the last part of Sect. 4.1. for Uccle.

In Fig.23, we show the comparison of MAPA (left) retrieved AOD and MMF v8 with pre-scaled measured DSCDs using MAPA fitted SF (right) and the AERONET AOD. Note the comparable correlation for the two retrieval codes.

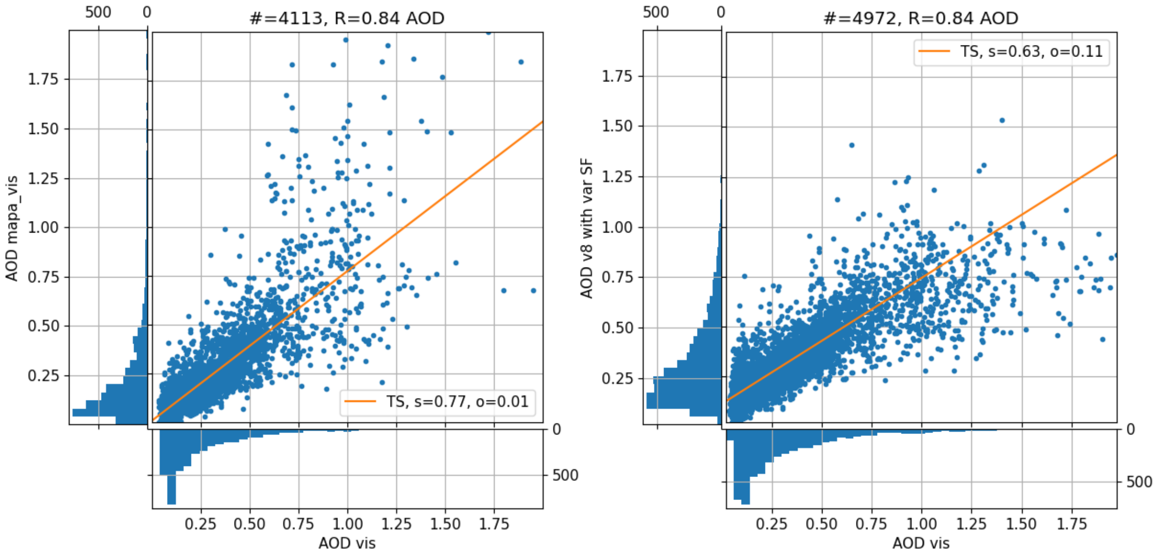


Fig. 23: MAPA retrieved AOD (left) and MMF v8 with pre-applying the MAPA fitted SF retrieved AOD compared to AERONET AOD, using their individual flagging.

Comparing the AOD of MAPA to MMF, as was presented in the right panel of Fig.22 for Uccle, shows comparable correlation and slope for Xianghe VIS, to what was found for Uccle VIS (Fig. 22 right panel), see Fig.24.

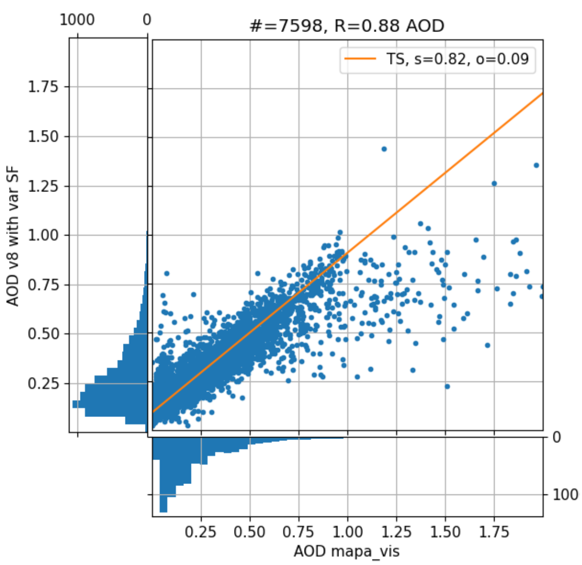


Fig. 24: MAPA retrieved AOD (x-axis) compared to MMF v8 with pre-applied MAPA fitted SF (y-axis).

5 UV and VIS in UCCLE

Before presenting the unified validation approach of all stations, it is worth to consider both UV and VIS together at one station.

For UV, the equivalent figure to Fig.22 (left, up to 2 km, right up to 4 km, using the pre-applied variable SF) comparing MMF v8 with variable scaling factor from MAPA compared to MAPA is shown in Fig.25. The agreement, both in terms of slope, as well as in terms of correlation is even better in UV than in VIS.

The same analysis (both for UV, upper panels, and VIS, lower panels) but with MMF version 8 with the fixed SF of 0.85 is shown in Fig.26. The correlation is much worse than when applying the SF prior to MMF retrieval. Hence, we conclude that pre-applying the MAPA fitted SF to the measured differential slant column densities (dSCDs) before MMF retrieval greatly increases the agreement between the two retrieval codes both in the UV as well as in the visible.

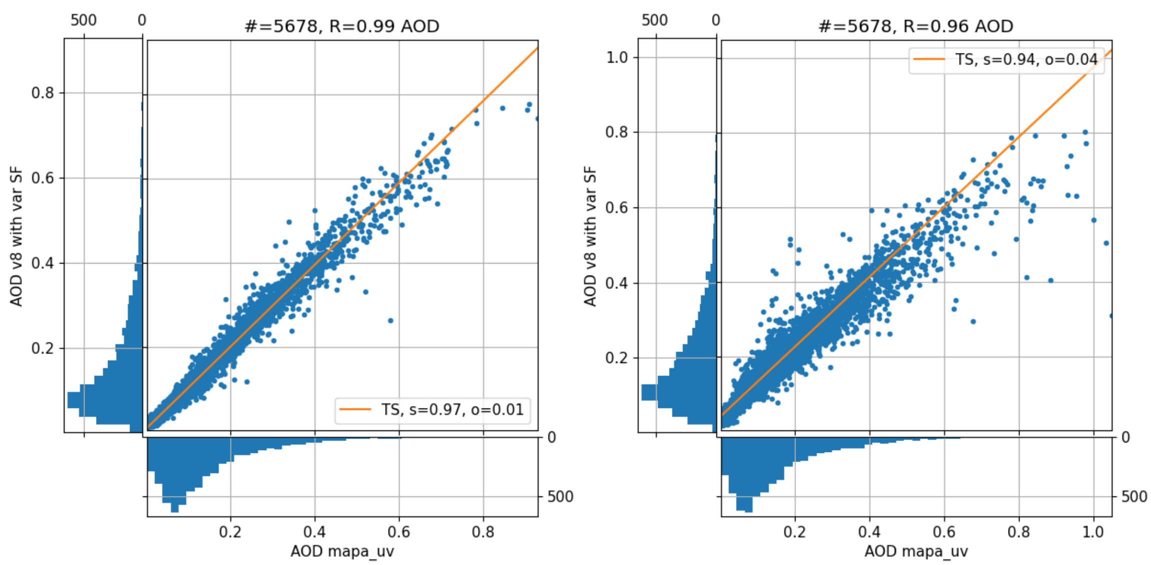


Fig.25: As Fig. 22 but for UV: partial AOD comparison between MAPA (x-axis) and MMF version 8 with pre-applied SF from MAPA (y-axis) for Uccle UV up to 2 km (left) and 4 km (right).

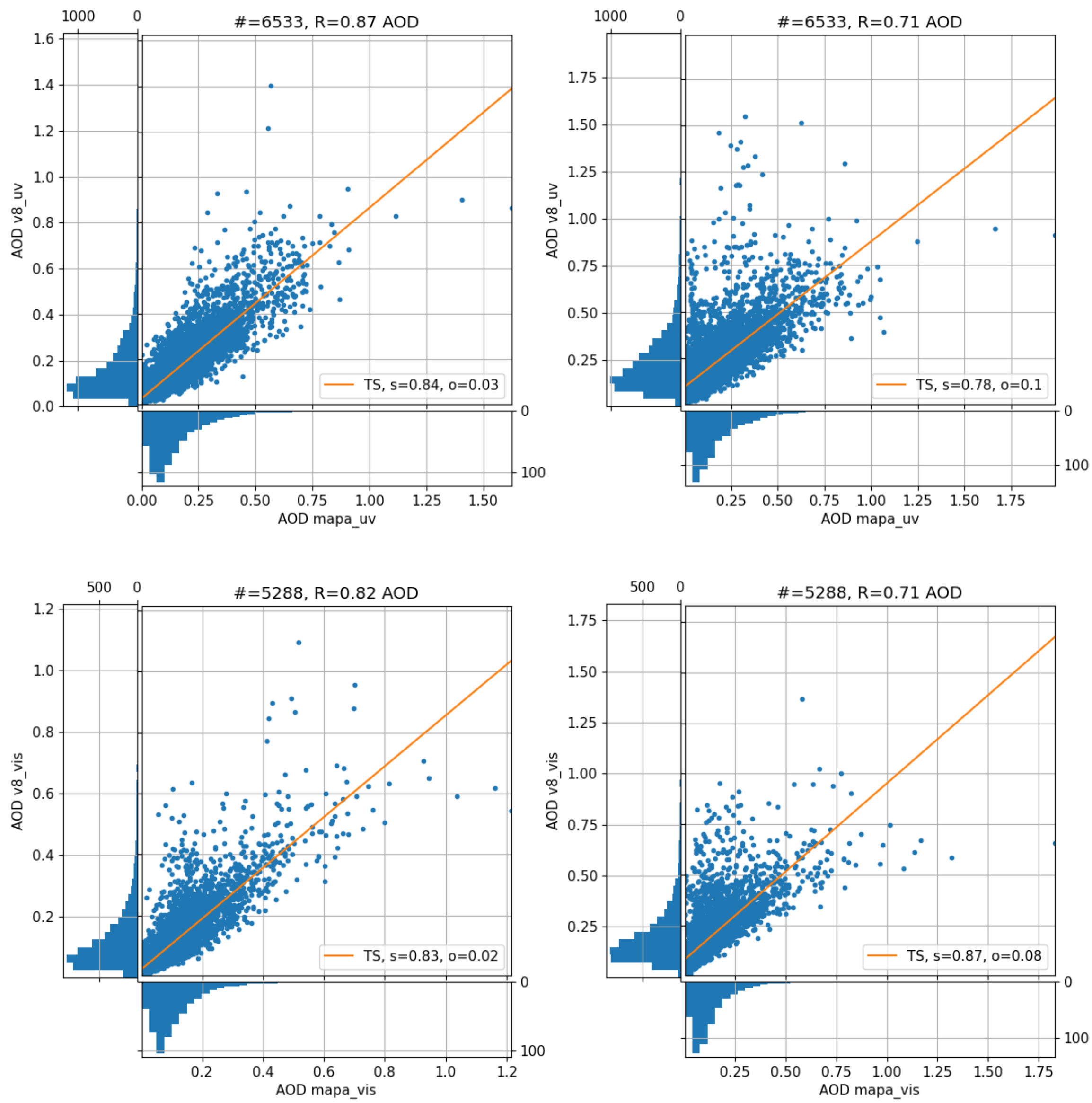


Fig. 26: MAPA retrieved AOD (x-axis) compared to MMF v8 with fixed SF 0.85 (y-axis) partial AOD up to 2 km (left) and 4 km (right) in UV (upper panels) and VIS (lower panels).

6 Unified validations

Until now, the assessment has been performed in terms of correlations and slope between MAX-DOAS retrieval on the one hand and AERONET AOD and/ or scaled backscatter profiles on the other hand. However, as was shown in Sect. 2.1.1, the correlation depends heavily on the exact time match and small time-shifts can lead to significant differences.

We propose to consider instead the absolute and relative difference in AOD between MAX-DOAS retrieved AOD and AERONET retrieved AOD, where we use the AERONET retrieved AOD as the base for the relative comparisons.

In this Section, we will use the AERONET AOD data (independent on the existence of the presence of backscatter profiles). This increases the number of coincidences compared to other parts of the report. This also means that no further filtering is applied. It also enables to treat all stations in the same way, also those for which no backscatter profiles are available.

All MMF tests use a flag constructed in the same way (as described in Sect.2.2.4), including the original MMF. However, since the limits on relative RMS as well as on the consistency of the two runs with different a priori settings were different for the original MMF run, the flags are not exactly comparable and often result in a lower number of valid values for the original MMF. However, this reflects the state of original MMF as was shown in [D3.2].

From the MAPA flag, we only keep the part about the SF. This is motivated by tests performed (presented at PM 10, see Appendix A.2) that investigated which parts of the MAPA flag is responsible for working so well in UV on MMF data. However, MMF applies an even stricter limit than the strict warning flag of MAPA, it is required that the SF is within [0.7, 1.2].
In this section, we use all MMF data that is not flagged as error, and for MAPA, MAPA data that is flagged as valid.

We consider all tests presented so far.
For MAINZ UV: MAPA, MMF original and 9 tests (see Table 1)

For Uccle UV and VIS: MAPA, MMF original and test 8 once with a fixed SF (0.85) and once with the fitted SF from MAPA.

For Xianghe UV and VIS: MAPA, MMF original and test 8 (in UV, with a fixed SF, in VIS with the variable SF from MAPA)

For Cabauw UV: MAPA, MMF original and test 8 with a fixed SF.

For each instrument, we will show a Figure with 4 panels:

A histogram showing the valid number of points for MAPA (MAPA, unless otherwise stated, refers to MAPA VAR), MAPA SF 1 (in the following MAPA1), MAPA SF 0.8 (in the following MAPA08), MMF original (MMF) and the various MMF tests (v1... v9; if a scaling factor different from 1, c.f. Table 1, this is also directly indicated). This is shown in the upper left panels of Fig. 27 – 37.

2 plots showing the percentage of valid data in bins of absolute AOD difference (upper right panel) and relative AOD difference (lower right panels) between MAX-DOAS retrieved AOD and AERONET AOD.

Lastly, we show the cumulative percentage of valid measurements over the absolute AOD difference (lower left panels). For completeness, we add the corresponding scatter plots in the Appendix A.4.

For those stations that were tested with the favoured new settings, i.e. MMF v8 with MAPA fitted SF, we include the statistics of MMF original, MMF v8 variable SF and MAPA for the different quality flags, since the second point highlighted by the technical note [D3.2] for MMF was the low number of valid data. Note, that the baseline (i.e. 100%) refers here to all coincidences between MAX-DOAS measurements and sun-photometer measurements. As outlined above, MAPA considers only truly valid data as valid, while MMF and MMF v8 variable SF consider warning and true valid as valid. The part that is indicated with a grey background in Tables 9—12 contains the data that is shown in the corresponding figures.

6.1 Mainz

For Mainz, many tests were performed. We show all tests in Fig.27 and Fig.28. for the combined MAINZ data (i.e. the combined directions); for the individual directions, we only show a subset of tests (the ones shown in Fig.27 for MAINZ combined), these are shown in Fig.29 -- Fig.32. Note however that for MAINZ, no tests with MMF applying MAPA's fitted SF were performed.

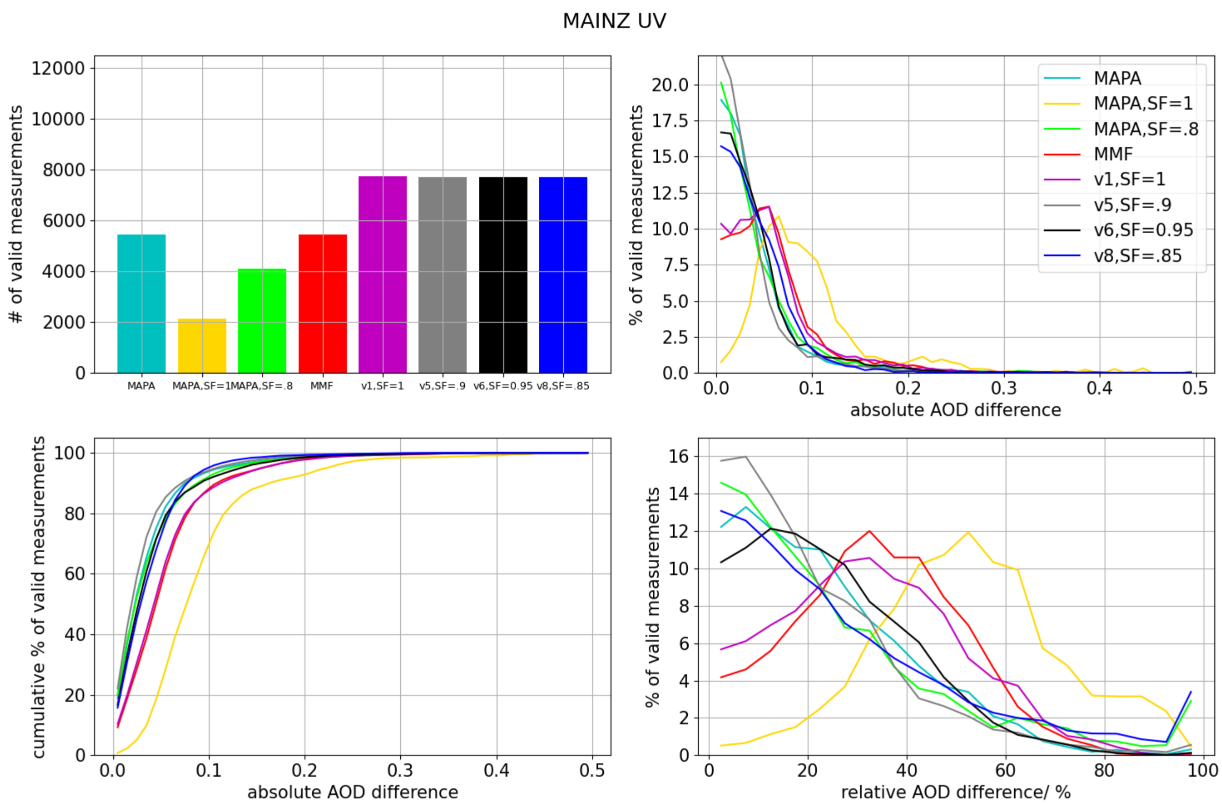


Fig. 27: Histogram of valid coincidences for the different retrievals (upper left), percentage of valid measurements as a function of absolute (upper right) and relative (lower right)

AOD differences between retrieval and AERONET. Cumulative percentage of valid coincidences as function of absolute AOD difference.

For the mean of the directions and the main tests (Fig.27), we find the following:

- MMF and MAPA have similar number of valid coincidences, new MMF versions have considerably more valid data
- MAPA, MAPA08, v5, v6 and v8 all show similar satisfactory performance in terms of absolute and relative difference between MAX-DOAS retrieved and AERONET AOD. Those are retrievals using a SF<1.
- MMF and MMF v1 show a better performance and a higher number of valid points than MAPA1, but for both the absolute as well as the relative differences are much worse than the retrievals using a SF < 1.

Considering MAINZ direction 1 (Fig.29), the absolute difference between the retrieved AOD and the AERONET AOD for v8, v1 and MAPA are similar. Note that v1 uses a scaling factor of 1. Direction 3 (Fig.31) shows a similar picture. However, for the remaining 2 directions (Fig.30 and Fig.32), using a scaling factor different from 1 is necessary to close the gap in AOD.

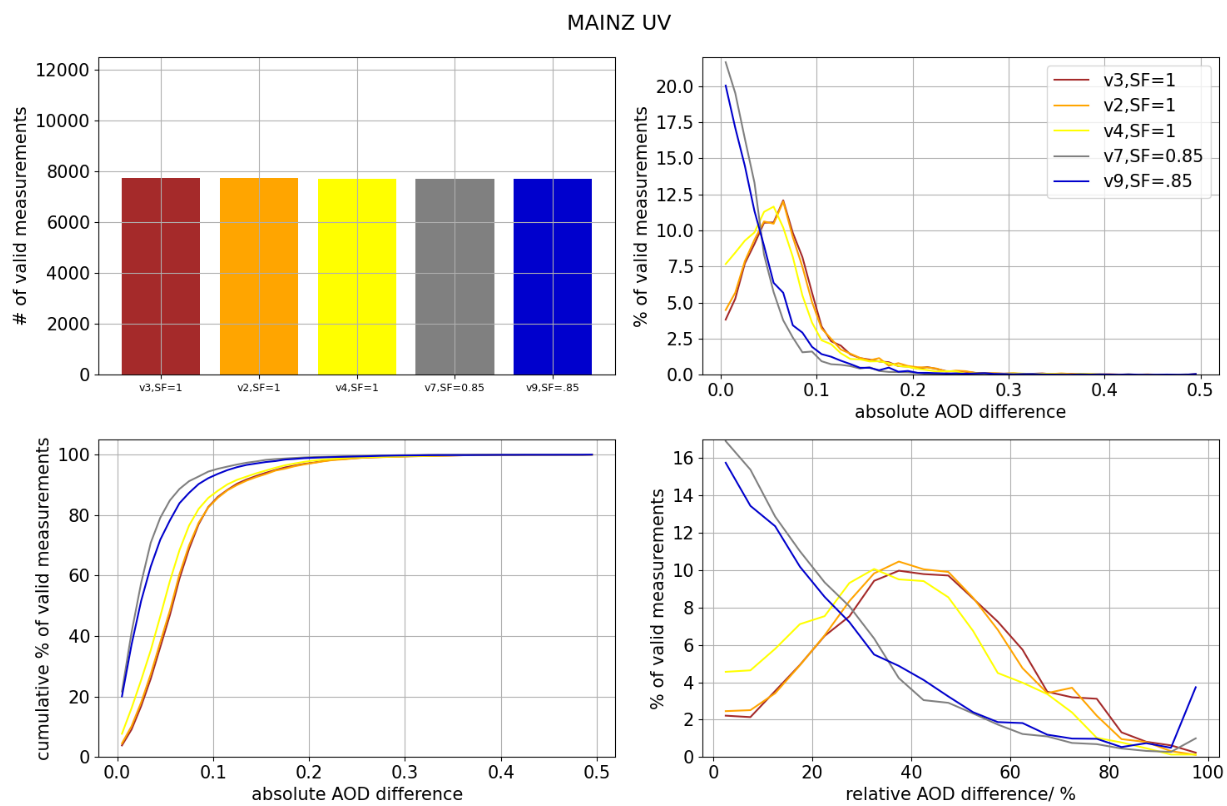


Fig. 28: Same as Fig. 27, but for tests that will not be shown for the individual directions.

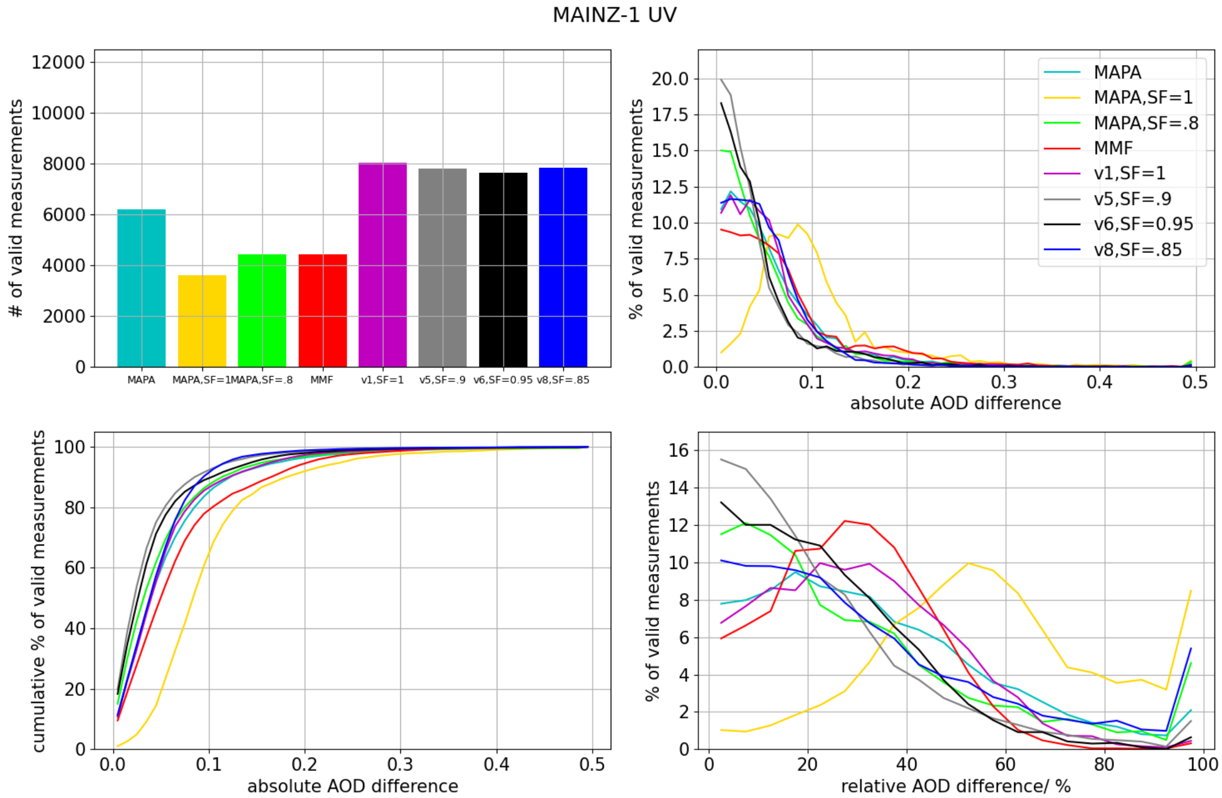


Fig. 29: As Fig. 27, but for MAINZ direction 1.

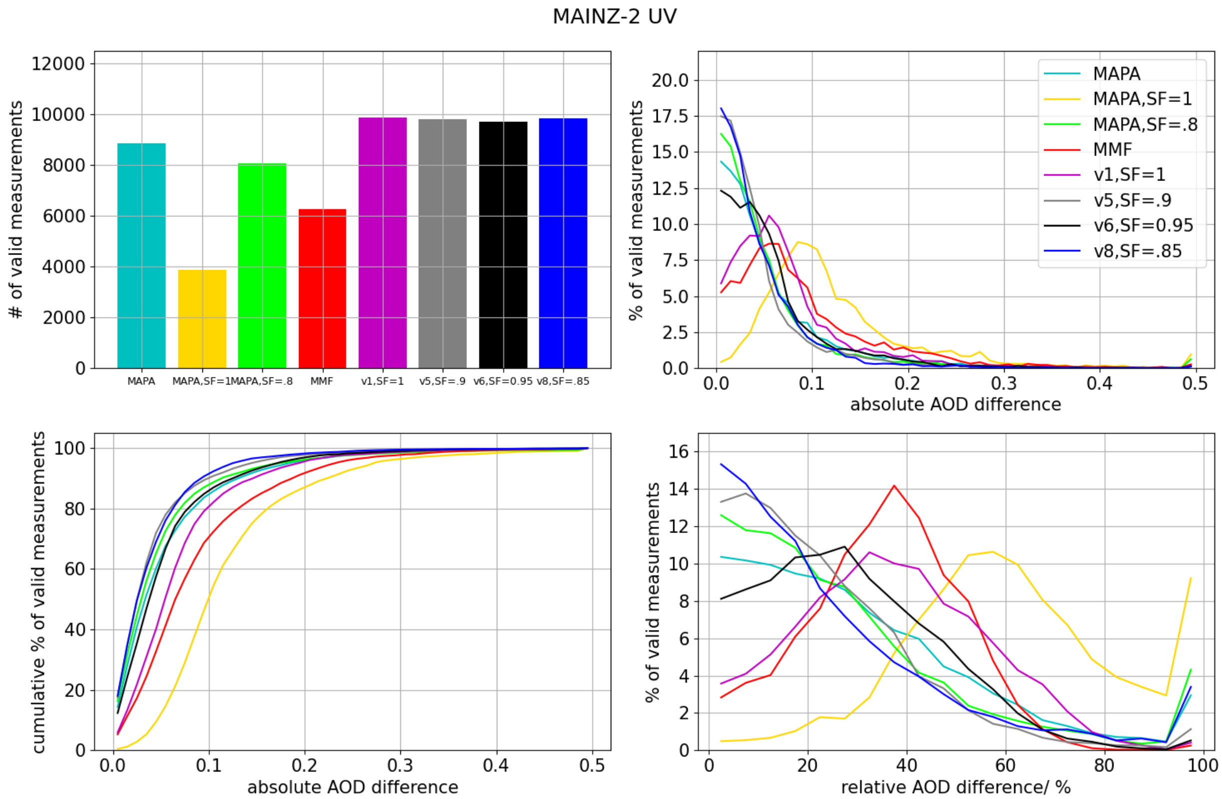


Fig. 30: As Fig. 27 but for Mainz direction 2.

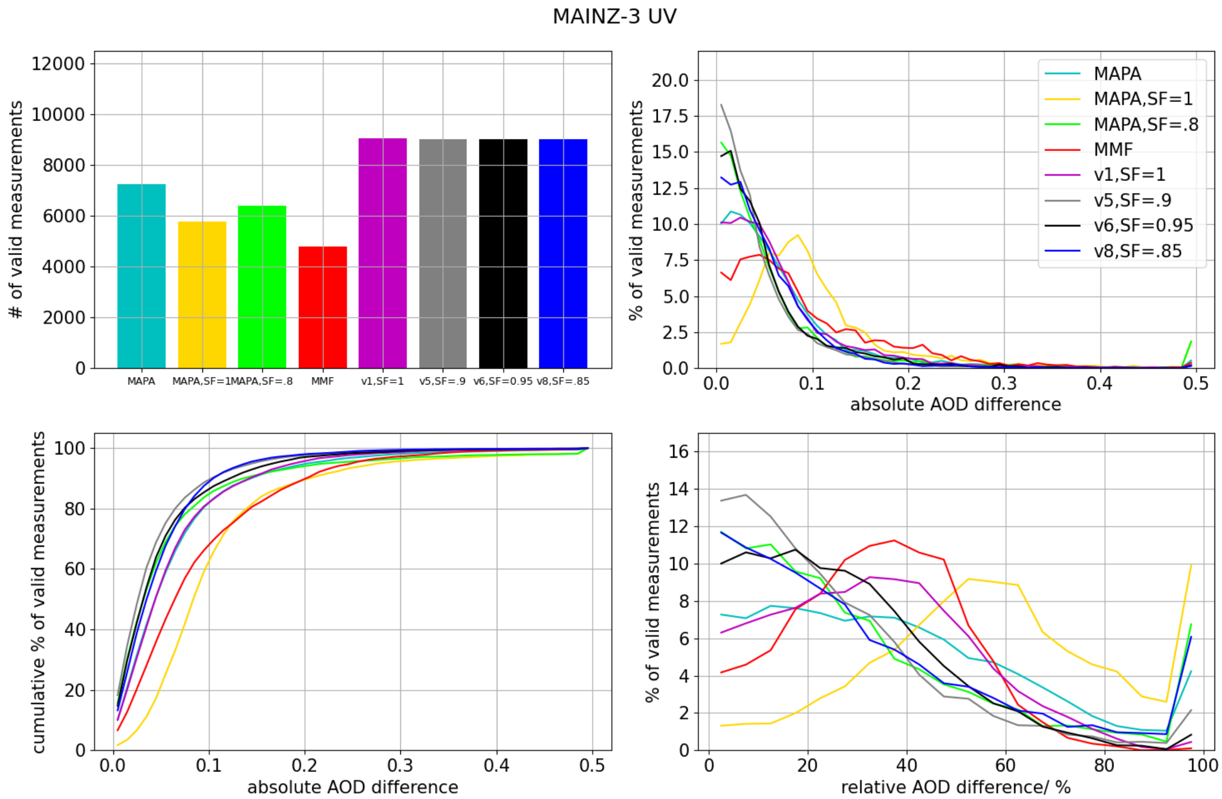


Fig.31: As Fig.27, but for MAINZ direction 3.

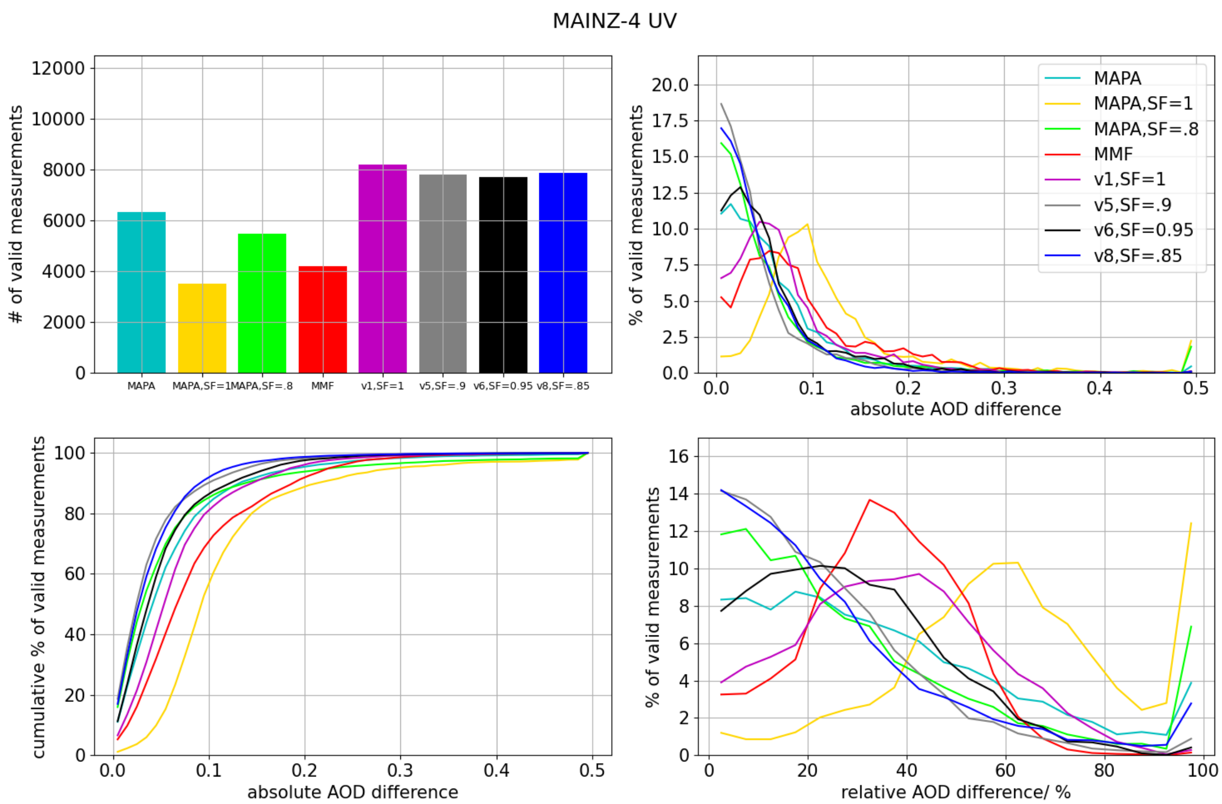


Fig. 32: As Fig.27 but for MAINZ direction 4.

We conclude that a scaling factor smaller than 1 is necessary for Mainz, however whether a scaling factor of 0.8 (as for MAPA0.8), 0.85 (as for MMF v8) or 0.9 (as for MMF v5) does not make a significant difference. MAPA (VAR) does not perform better than MAPA 0.8 for Mainz.

6.2 Uccle

In Sect. 4.1 it was found that for UCCLE in the visible channel, applying a constant scaling factor of 0.85 was not improving the results, but applying the variable SF from MAPA did. Considering the AOD differences (Fig.33), we see the following:

- MAPA08, MMF and v8 with the variable SF fitted by MAPA show equally satisfactory results and perform better than MAPA (VAR)
- MMF original performs equally good as v8 with MAPA fitted variable SF, hence no SF is needed for MMF to close the AOD gap.
- v8 with SF fitted by MAPA retains most data, MMF the least data. Hence, while both perform similar in terms of correlation with AERONET, v8 with a fitted scaling factor is preferred due to the larger number of valid data.
- v8 with a fixed SF of 0.85 performs, in terms of relative or absolute AOD difference, equally bad as MAPA1

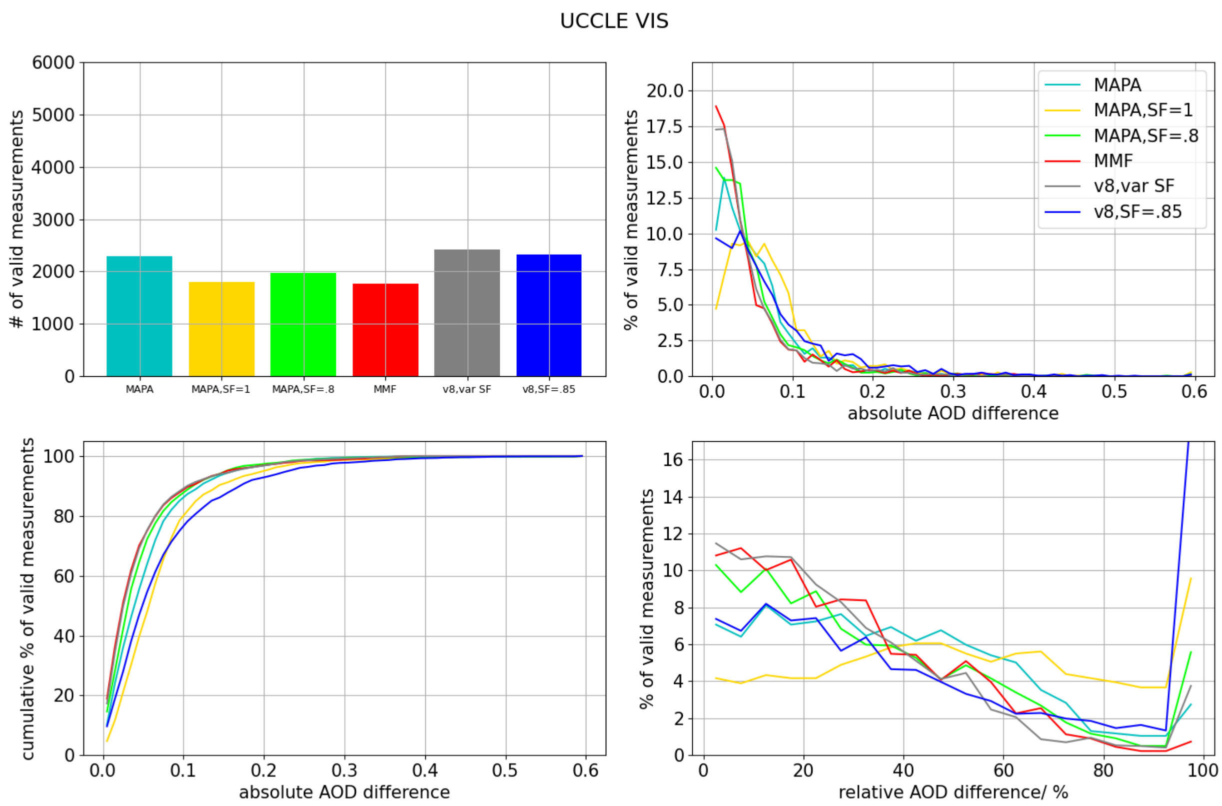


Fig. 33: As Fig. 27, but for UCCLE VIS

UCCLE in the UV (Fig.34) shows a quite different picture than in the visible. As was found for MAINZ (average of directions), a SF smaller than 1 seems necessary to close the AOD gap:

- There is a clear gap between retrievals applying a SF different from 1 (MAPA, MAPA08, v8 with fixed and variable SF) and retrievals using no SF (i.e. SF=1, MMF, MAPA1)
- MAPA08 performs only slightly better than MMF with MAPA fitted SF and considerably better than MAPA (fitting the SF).

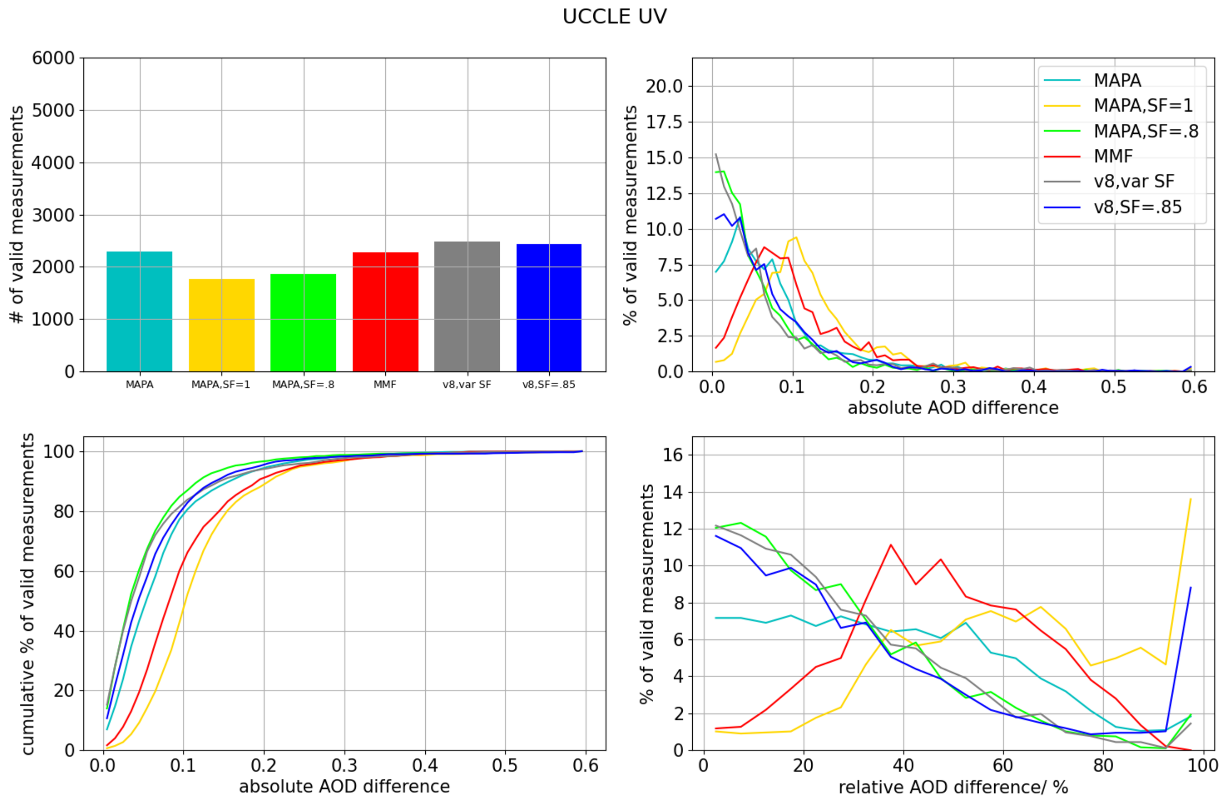


Fig. 34: As Fig. 27, but for UCCLE UV

Table 9: validity statistics for UCCLE UV

	Percentage valid	Percentage warning	Percentage error
MMF original	0.01	73.07	26.92
MMF v8	70.81	8.67	20.52
MAPA	73.27	7.49	19.24

Table 10: validity statistics for UCCLE VIS

	Percentage valid	Percentage warning	Percentage error
MMF original	0.03	51.67	48.30
MMF v8	49.62	21.33	29.05
MAPA	67.06	12.00	20.95

6.3 Cabauw

For Cabauw (Fig.35), a test for MMF with MAPA's variable scaling factor was performed, as well as a test with a fixed SF (0.85).

The main findings are:

- V8 applying a fixed SF of 0.85 or a variable SF outperforms both MMF, as well as MAPA and MAPA08 (the latter two however not significantly), while also retaining more data
- Opposed to MAINZ UV and UCCLE UV, original MMF performs only insignificantly better than MAPA1.

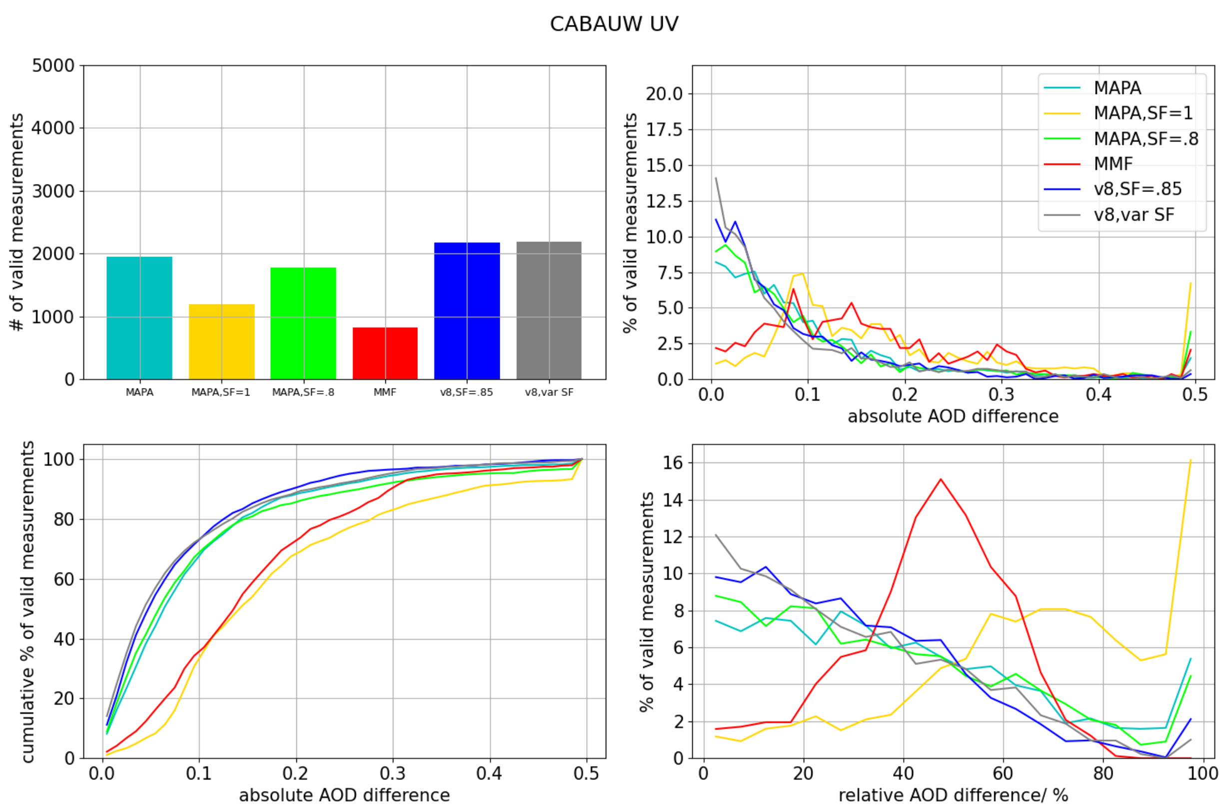


Fig. 35: As Fig. 27, but for CABAUW UV

Table 13: validity statistics for CABAUW UV

	Percentage valid	Percentage warning	Percentage error
MMF original	0.01	19.90	80.09
MMF v8	35.92	17.19	46.79
MAPA	47.29	15.16	37.55

3.4 Xianghe

For Xianghe VIS (Fig.36), a test was performed for v8 with a variable SF (as fitted by MAPA). For UV (Fig.37), both a fixed SF (0.85) as well as the variable SF as fitted by MAPA were tested. While MMF original (apart from a lower count of valid data) at UCCLE worked very well in the visible, the situation is different in Xianghe. Already at moderate AODs (>0.25), an underestimation is visible (c.f. Fig.38). High AOD values in Xianghe are often caused by transport of aerosols from close by heavy pollution sides and originate hence at higher altitudes.

Unfortunately, no aerosol profile data for Xianghe was available. Considering the relative AOD difference plots in the visible and in the UV (lower right panels of Fig.36 and Fig.37, respectively) shows that both retrieval codes with any settings miss a certain percentage of AOD. Likely because it is located at altitudes where MAX-DOAS lacks sensitivity. However, MMF v8 with MAPA variable scaling factor performs reasonably well in both UV and visible.

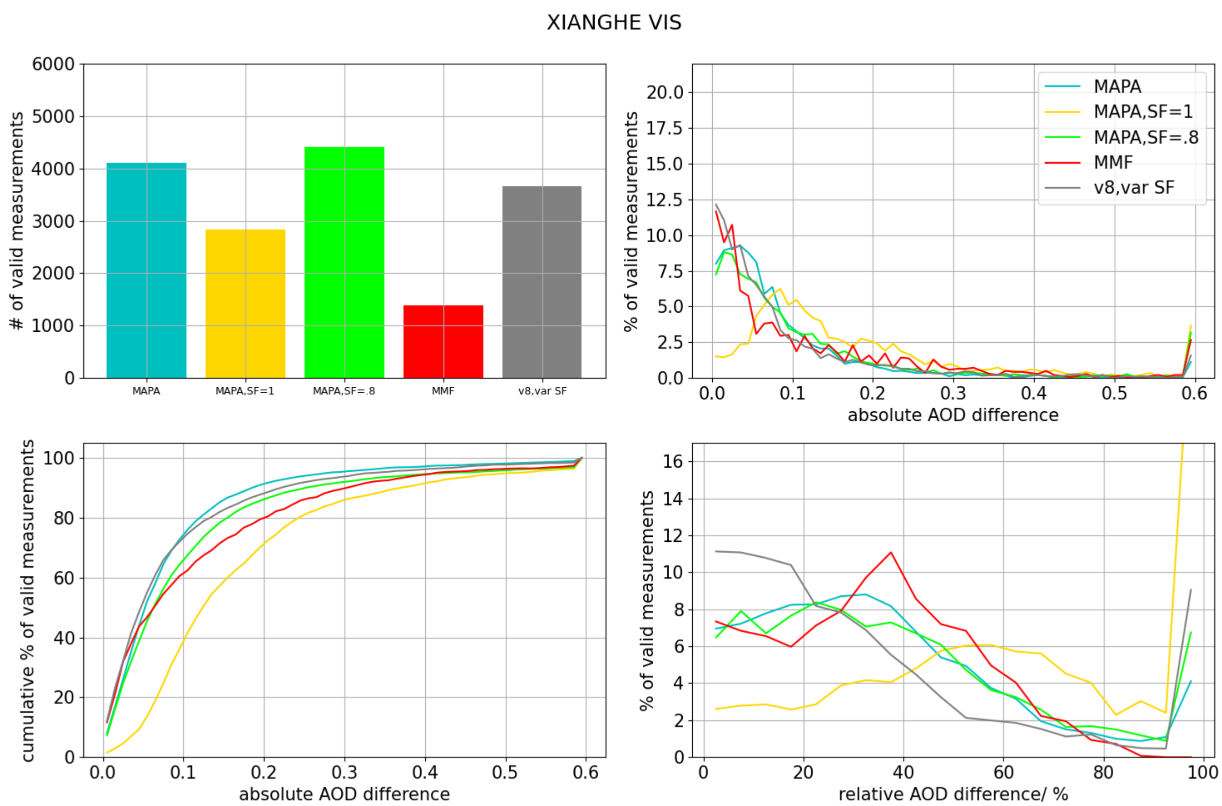


Fig. 36: As Fig. 27, but for XIANGHE VIS

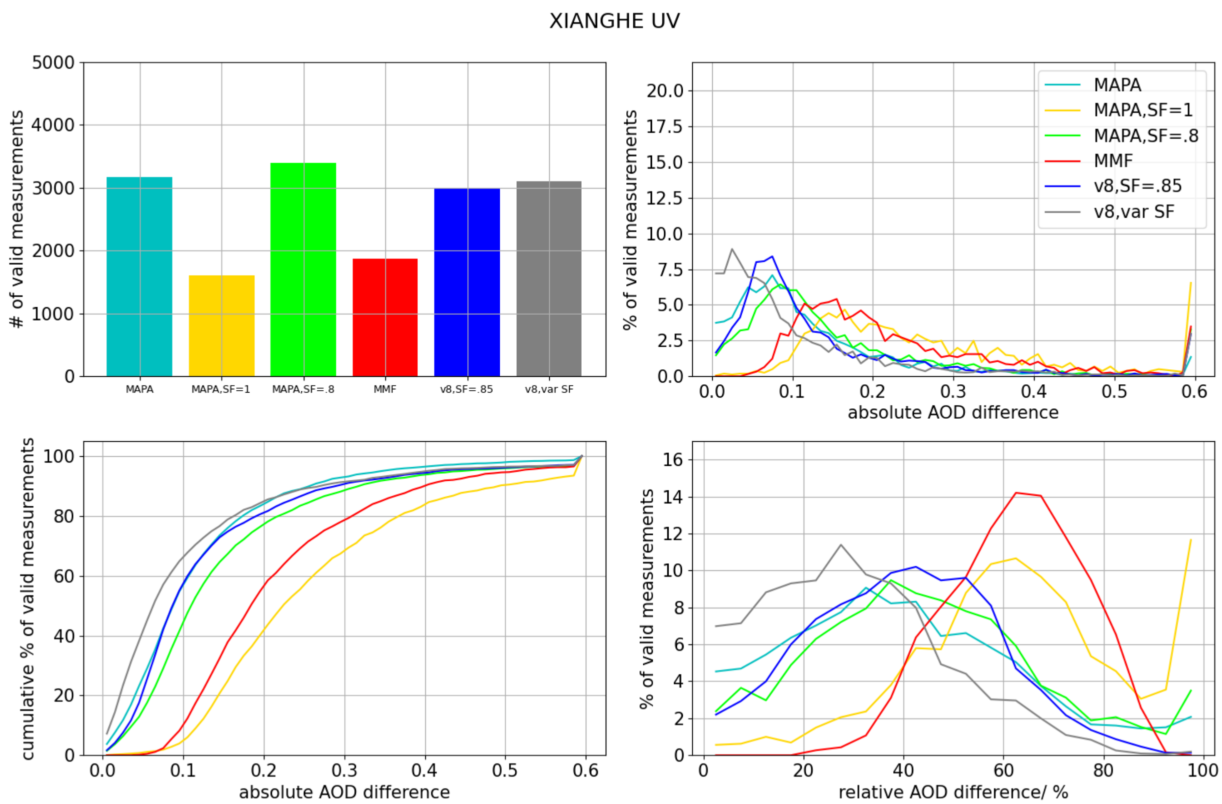


Fig. 37: As Fig. 27, but for XIANGHE UV

Table 11: validity statistics for XIANGHE UV

	Percentage valid	Percentage warning	Percentage error
MMF original	0.00	24.04	75.96
MMF v8	29.34	10.71	59.95
MAPA	40.94	17.77	41.33

Table 12: validity statistics for XIANGHE VIS

	Percentage valid	Percentage warning	Percentage error
MMF original	0.00	16.60	83.40
MMF v8	36.58	7.36	56.16
MAPA	49.09	19.78	31.13

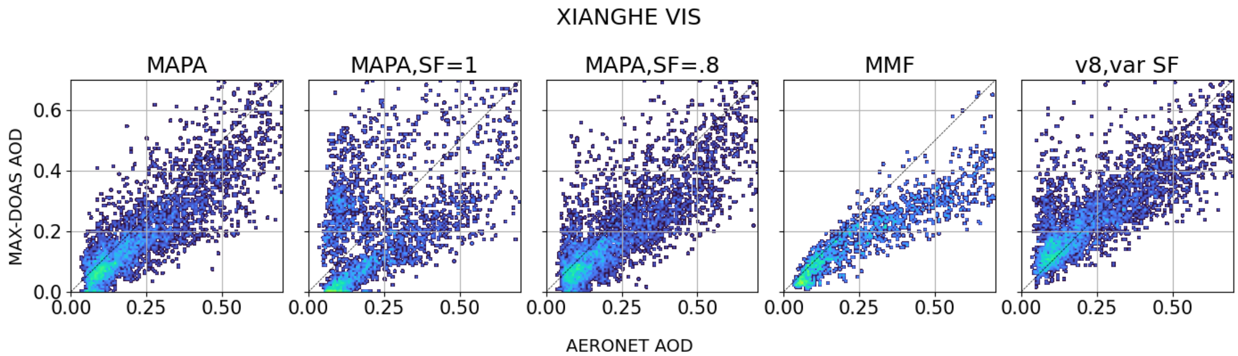


Fig. 38: Zoomed scatter plot of MAX-DOAS retrieved AOD (left to right: MAPA, MAPA1, MAPA08, MMF and v8 var SF). For a full-scale version, see Appendix A.5.

7 Summary and conclusions

The TN [D3.2] showed that AODs retrieved by MMF showed an equally good correlation with sun-photometer AODs as MAPA VAR retrieved AODs, however, especially in the UV, it underestimated the AODs severely. Further, the flagging scheme in MMF was very strict, resulting in virtually no truly valid data. These shortcomings were addressed by an updated flagging scheme, a change of a priori settings and the use of MAPA VAR fitted O4 scaling factors. The main points from the investigations are:

- Applying a SF to the measured DSCDs prior to the retrieval hugely impacts the magnitude of extinction below 1.5 -- 2 km and can close the gap between MMF retrieved AOD and sun-photometer measured AOD
- The correlation between MMF MAX-DOAS retrieved and AERONET scaled backscatter profile in this altitude range is hardly affected at all by applying a SF lower than 1.
- With a constant SF (either 1 or 0.8), the choice of the a priori profile does not affect the profile retrieved with MMF in the altitude region where the measurement yields sufficient information about the profile.
- Allowing more freedom at higher altitudes (by means of using an a priori profile and hence a covariance matrix that falls of less steep with altitude), helps but cannot solve AOD underestimation alone in the UV for MMF.
- In UV, applying SF of 0.85 for MMF closes the gap, applying the same fixed SF to the visible worsens the correlation.
- Applying a SF lower than 1 reduces the correlation at higher altitudes.
- Applying the same SF as fitted by MAPA (limiting to the range between 0.7 and 1.2), yields very good agreement between MAPA and MMF, especially in the lowest 2 km with slopes and correlations very close to 1.
- Applying a variable SF for MMF (as fitted by MAPA VAR) works in both wavelength ranges and sometimes outperforms MAPA (VAR) and hence will for the moment be a good way forward. However, the need for the application of a SF remains unclear and

should be the subject of further investigations. The currently proposed way forward should thus be seen as a pragmatic way forward.

References

Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y., & Wagner, T. (2019). The Mainz profile algorithm (MAPA). Atmospheric Measurement Techniques, 12(3), 1785–1806. https://doi.org/10.5194/amt-12-1785-2019

Henyey, L. G., & Greenstein, J. L. (1941). *Diffuse radiation in the galaxy*. The Astrophysical Journal, **93**, 70–83.

Friedrich, M. M., Rivera, C., Stremme, W., Ojeda, Z., Arellano, J., Bezanilla, A., García-Reynoso, J. A., & Grutter, M. (2019). NO_2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City. *Atmospheric Measurement Techniques*, 12(4), 2545–2565. https://doi.org/10.5194/amt-12-2545-2019

[D3.4] ATBD, FRM4DOAS-2.0. ATBD. Deliverable D3.4. FRM4DOAS-2.0

[D3.2] Technical Note, FRM4DOAS-2.0. TN. Deliverable D3.2. FRM4DOAS-2.0

Appendix

A.1. Mainz-1 MMF original vs MMF v8 extinction layers

We show here the scatter plots for the lowest 8 layers comparing the original MMF implementation to Test 8 which is the basis for the proposed new implementation, however with a fixed scaling factor. Each subsection presents correlation scatter plots of MMF v8 compared to the original MMF implementation. While the correlation is high in each layer, the slope (both Theil-Senn and orthogonal distance regression) is larger than 1.

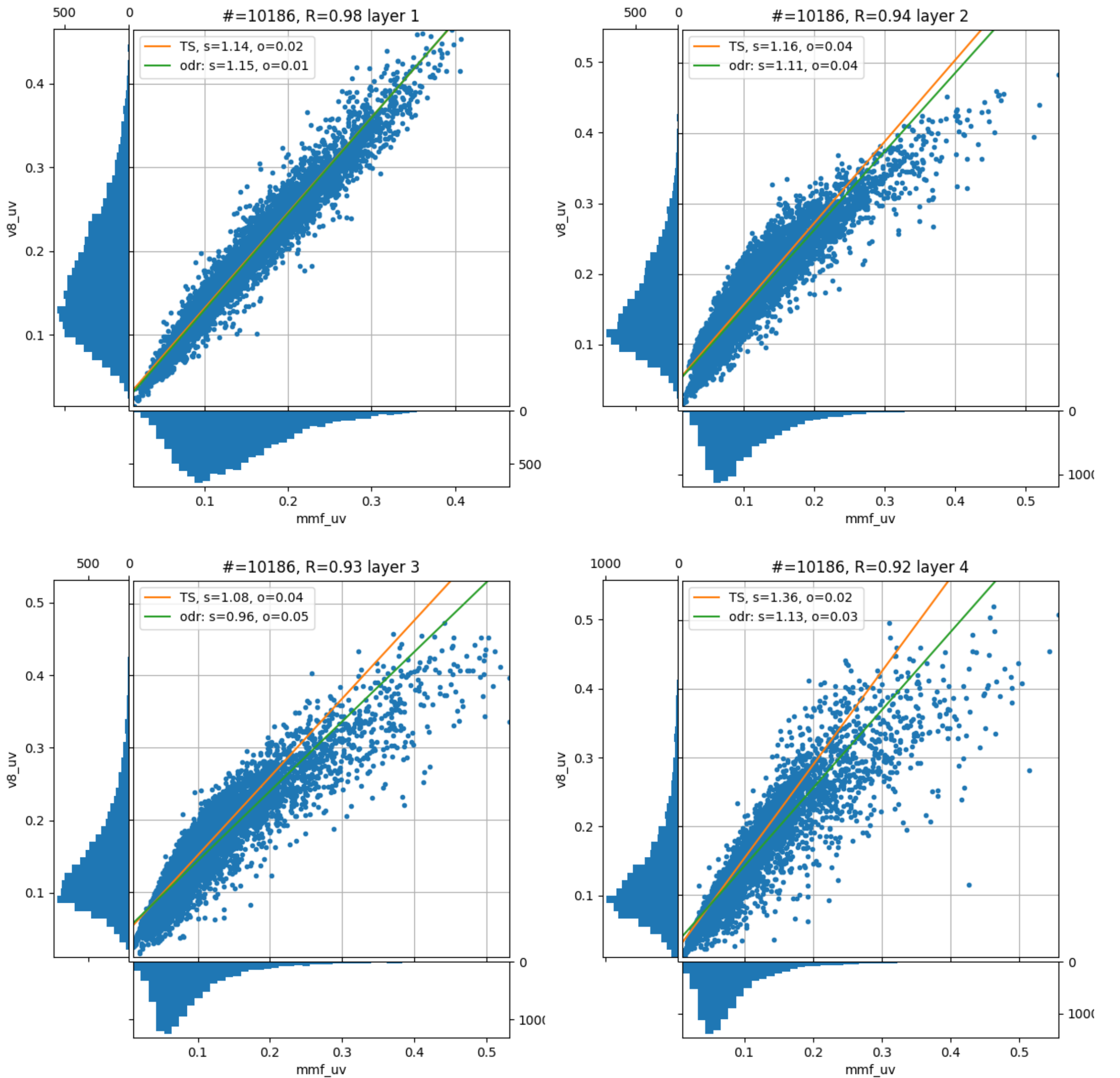


Fig. A1: Extinction per layer for MMF v8 compared to MMF original in the lowest 4 layers for MAINZ-1. Indicated in the title is the number of valid coincident retrievals and the correlation. Two estimators (orthogonal distance regression and Theil Senn slope) for the slope and offset are indicated, see legend.

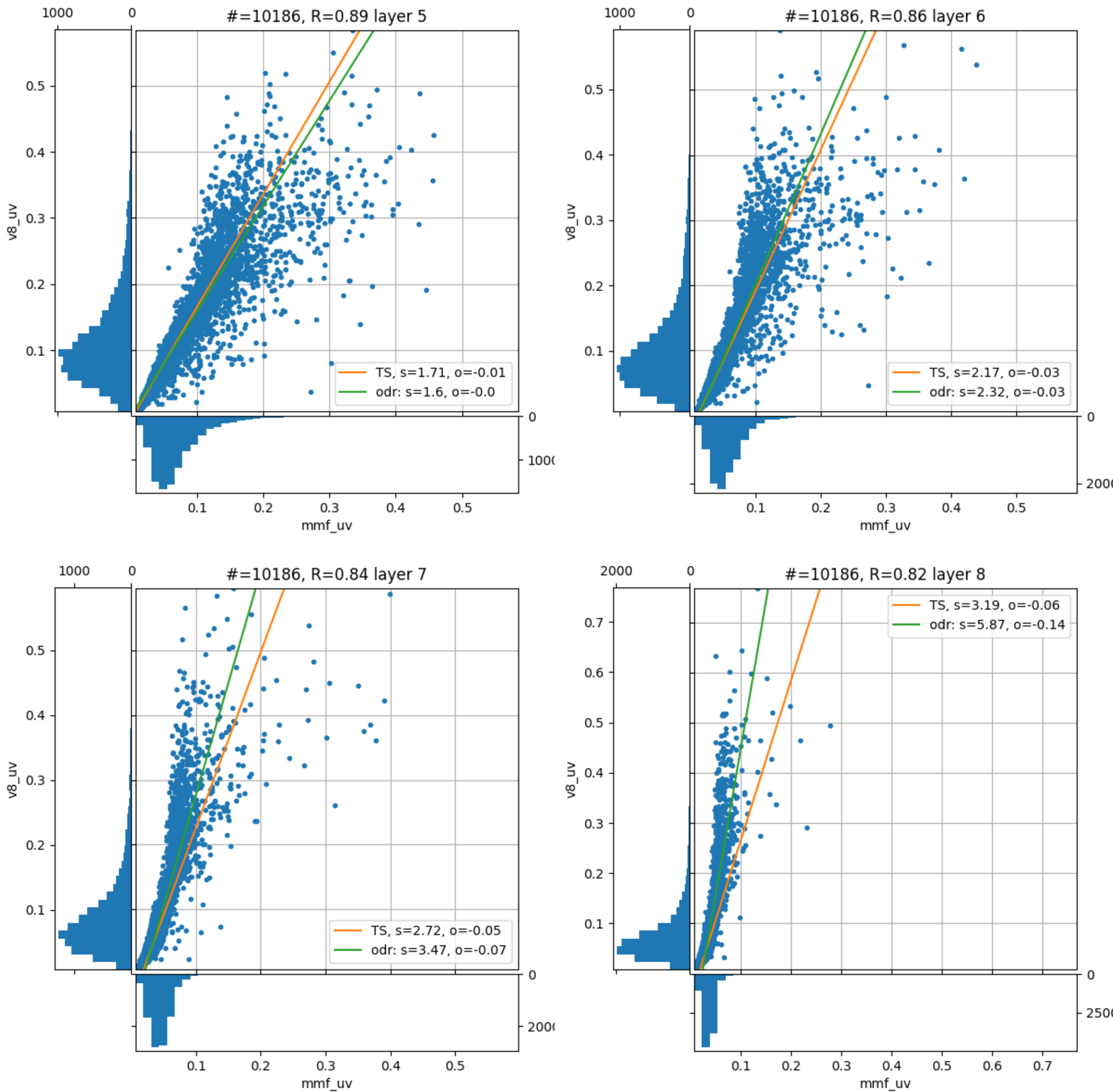


Fig. A2: As Fig. A1, but for layers 5 -- 8. Note the high correlation and the high slope.

A.2 Mainz impact of different MAPA flag components on MMF retrieval correlation with AERONET data

The MAPA flag consists of a number of sub-flags (Beirle et al. 2019). We investigated the effect of the following parts:

(1) Flag on RMS between forward model and measurement
(2) Flag on consistence between best matching profiles
(3) Flag based on the profile shape, whether large parts of the retrieved profile are located at higher altitudes
(4) Flag on the upper limit of the scaling factor
(5) Flag on the lower limit of the scaling factor
(6) Flag on the AOD, retrieved AOD values > 2 are flagged as invalid

In Fig. A3 – A6, we present the possible flag combinations applied to MMF data in descending order of resulting correlations with AERONET data. The blue curve shows the percentage of valid data and is read on the left-hand side axis.

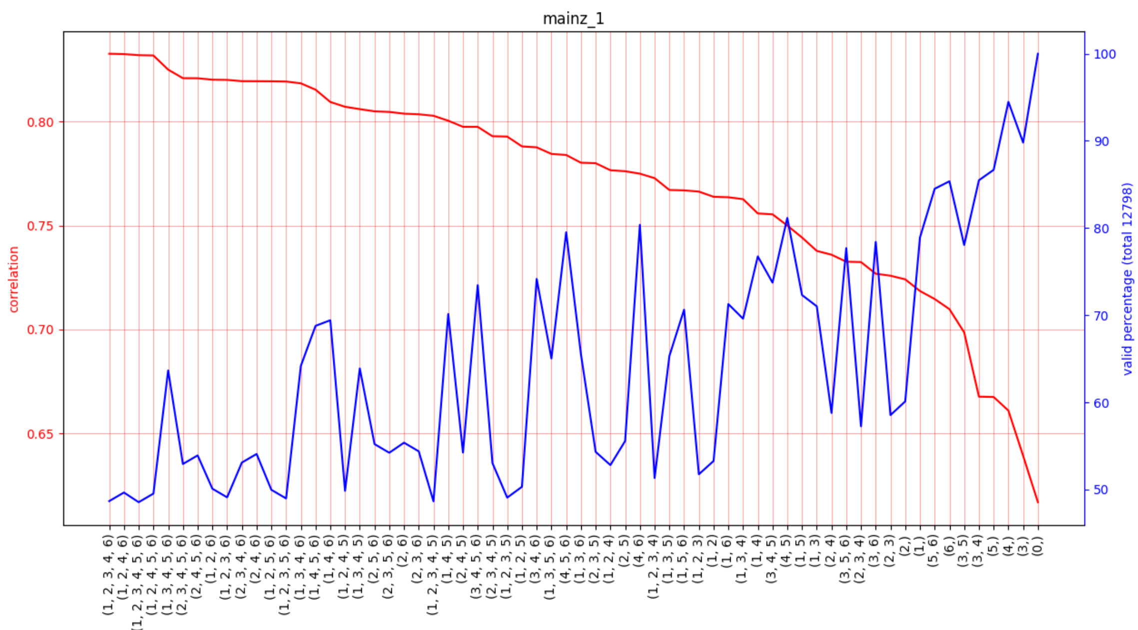


Fig. A.3: Correlation of MMF original AOD with AERONET AOD applying different combinations of MAPA flagging for Mainz direction 1 (north west).

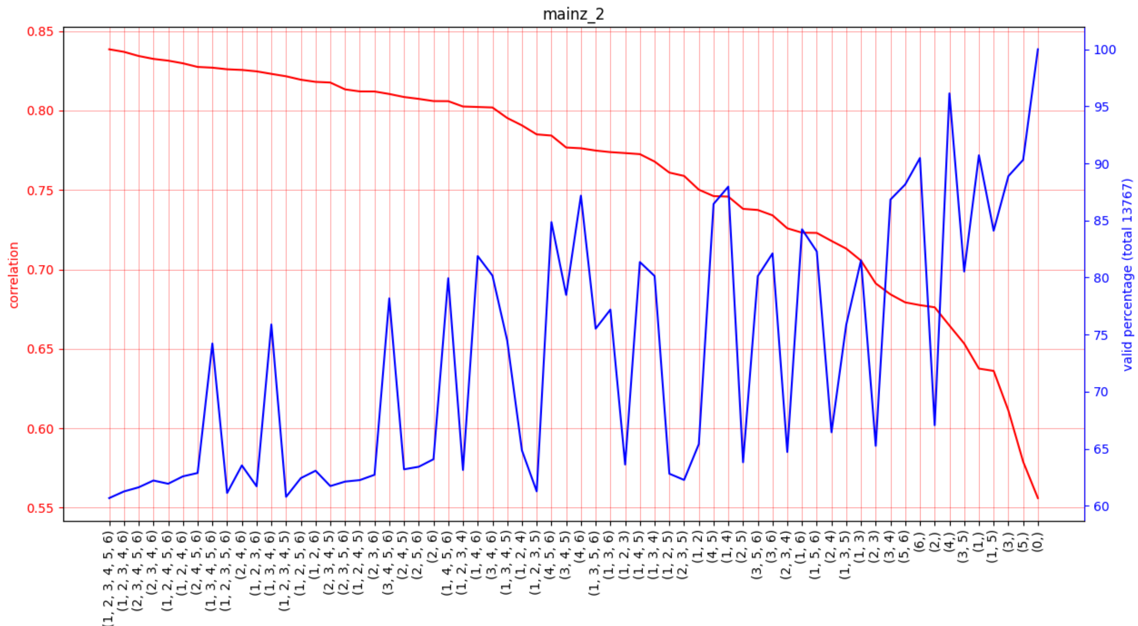


Fig. A.4: As Fig. A.3 but for Mainz direction 2 (north-east)

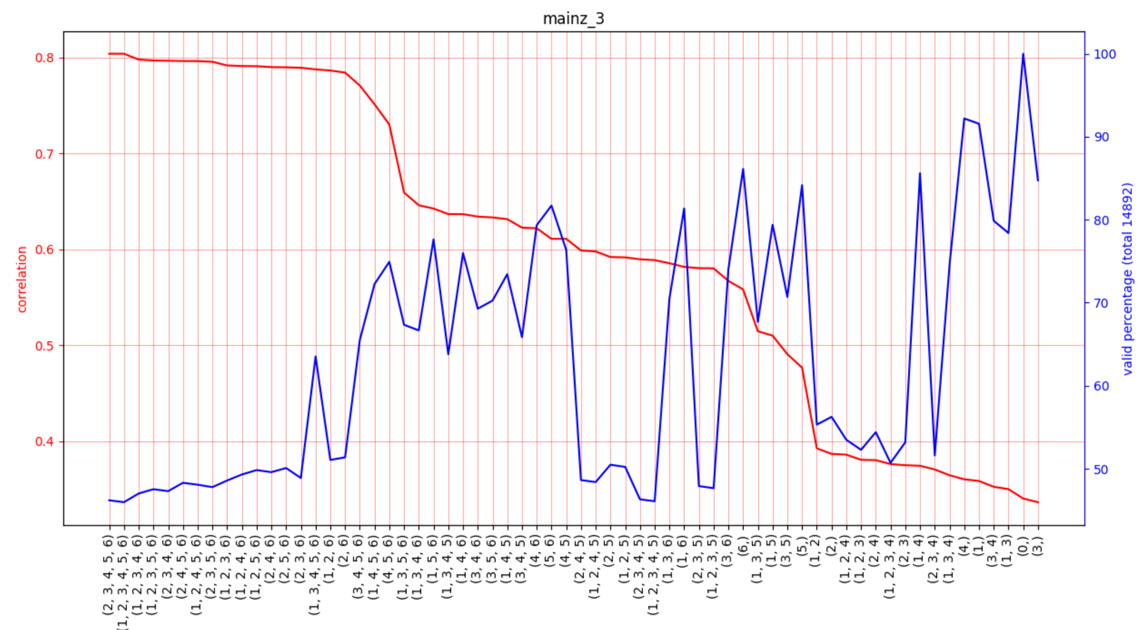


Fig. A.5: As Fig. A.3 but for Mainz direction 3 (south-east)

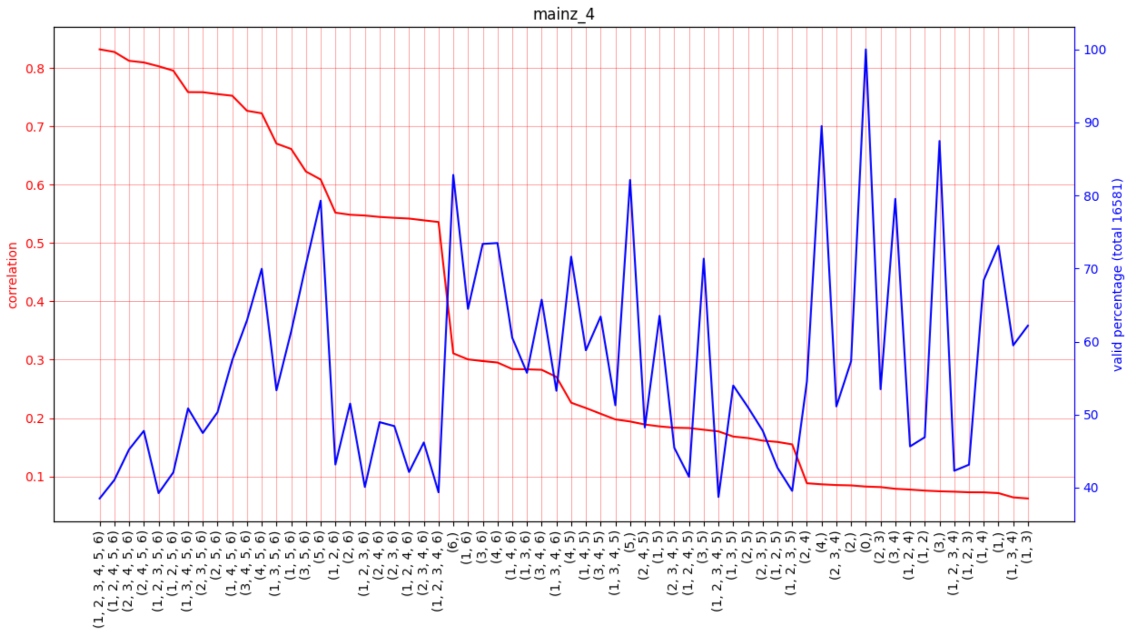


Fig. A.6: As Fig. A.3 but for Mainz direction 4 (south-west).

A.3 Mainz: Tests considering the smallest and biggest RAA of each set in the different directions

At PM 11, tests were presented to investigate whether a closer match in viewing direction between the sun photometer and the MAX-DOAS instruments results in an enhanced correlation. At each time, the direction (the measurements in Mainz in the four different directions happen simultaneously) the valid measurement in the direction with the lowest relative azimuth angle was selected. Once it was required that all four directions be valid, once that at least one direction is valid. As a common flag, MAPA qa-flag = 0 was applied to both retrieval codes.

Considering only data when all four directions are valid, resulted in a mean relative azimuth angle (RAA) of 28 degrees if selecting the measurement with the smallest RAA. If selecting the measurement with the largest RAA, the mean RAA is 152 degrees. Note that the RAA is with 0 and 180 degrees.

The results for MMF are presented in Fig. A.7, the results for MAPA are presented in Fig. A.8. For both retrieval codes, the lowest correlation and slope are observed in the low RAA selection of measurements. However, this difference is much more pronounced for MAPA than for MMF.

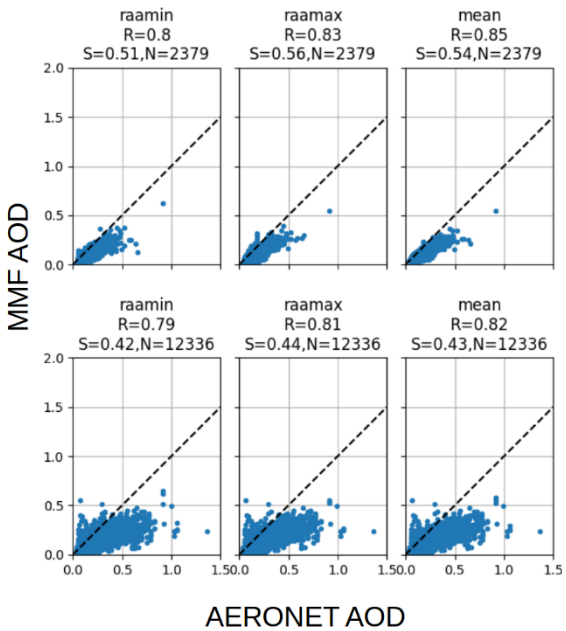


Fig. A.7: Scatter plot of MMF vs AERONET AOD selecting always the smallest (right panels) or highest RAA (middle panel) and considering the mean (right panel) of the four measurement directions in Mainz. The upper panels consider only times at which all 4 directions are valid (MAPA qa-flag = 0). The lower panels consider all times at which at least one direction is valid.

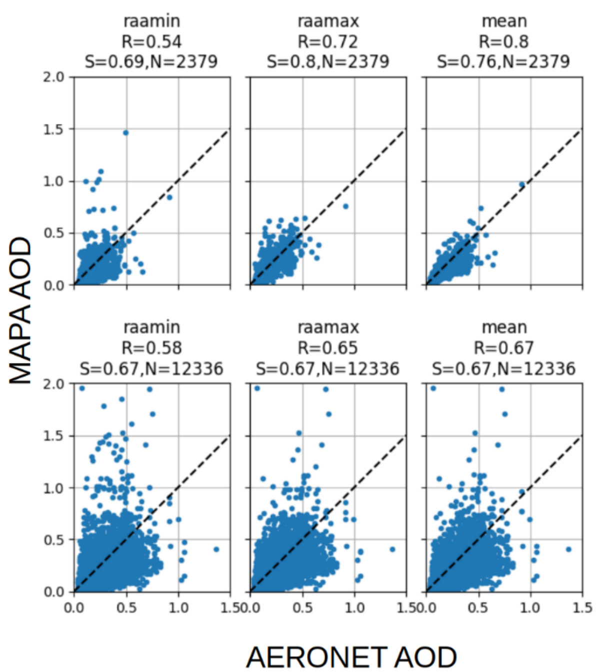


Fig. A.8: As Figure A.7 but for MAPA.

A.4 Unified AOD scatter plots

Additionally, to the difference plots presented in Sect.6, we present here all the corresponding scatter plots of the AODs.

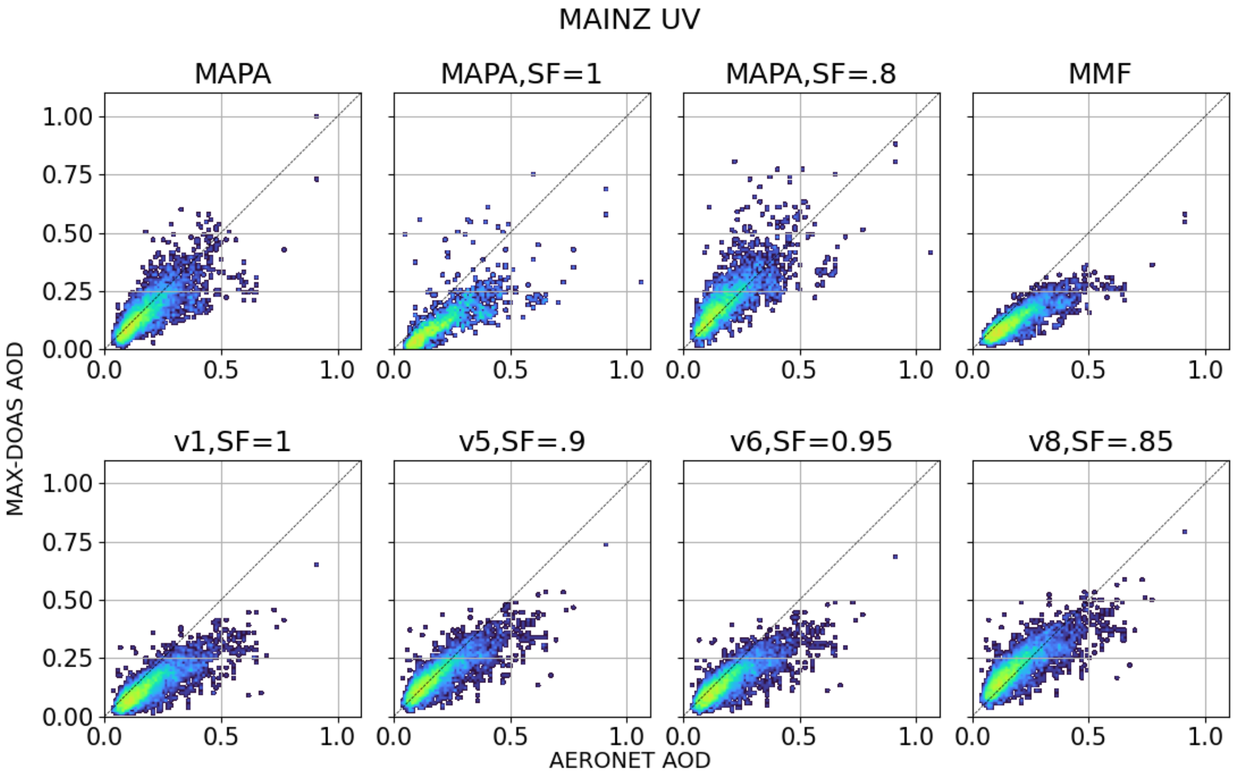


Fig.A.9: Unified scatter density plots (see Sect.6) for MAINZ mean direction. Each panel compares a different MAX-DOAS retrieval code and setting (on the y-axis) to AERONET AOD on the x-axis. The total number of points is not identical since each code (with each setting) uses their own validity flag. Consult the overview plots in Sect. 6 for information on the number of coincidences.

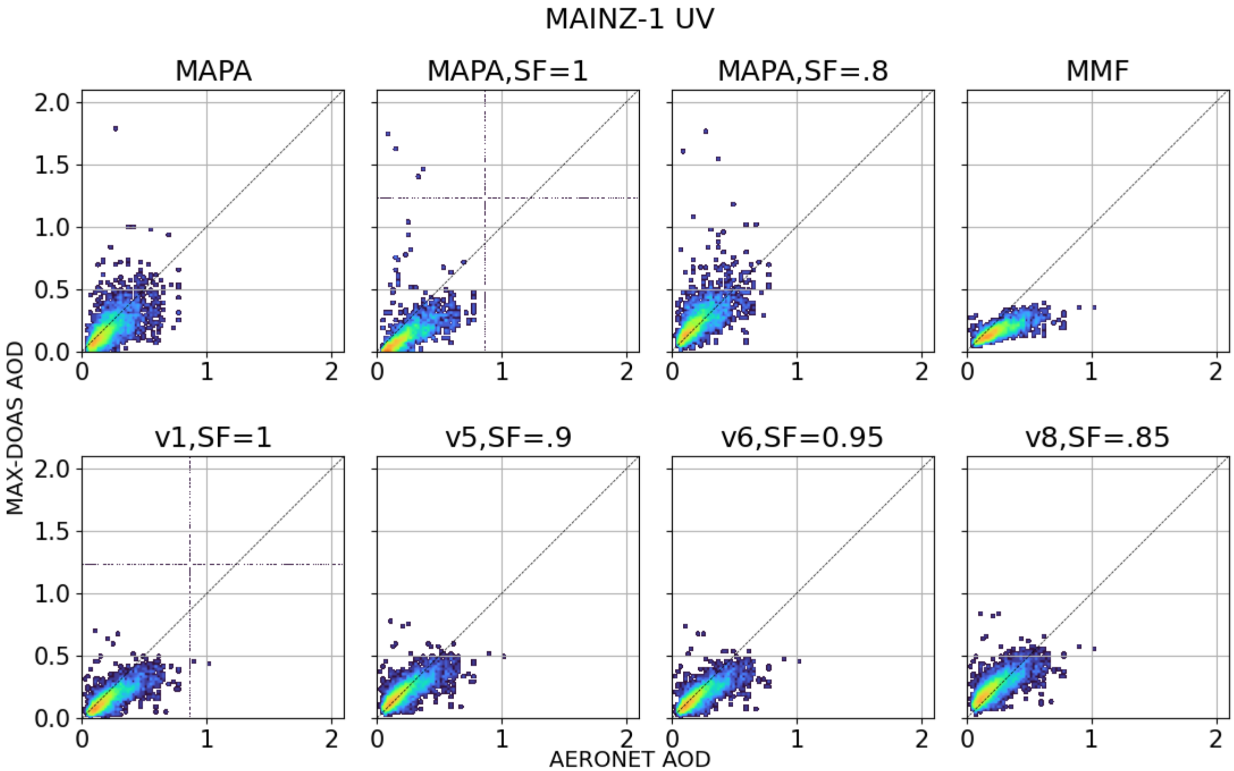


Fig.A.10: As Fig.A.9 but for MAINZ direction 1.

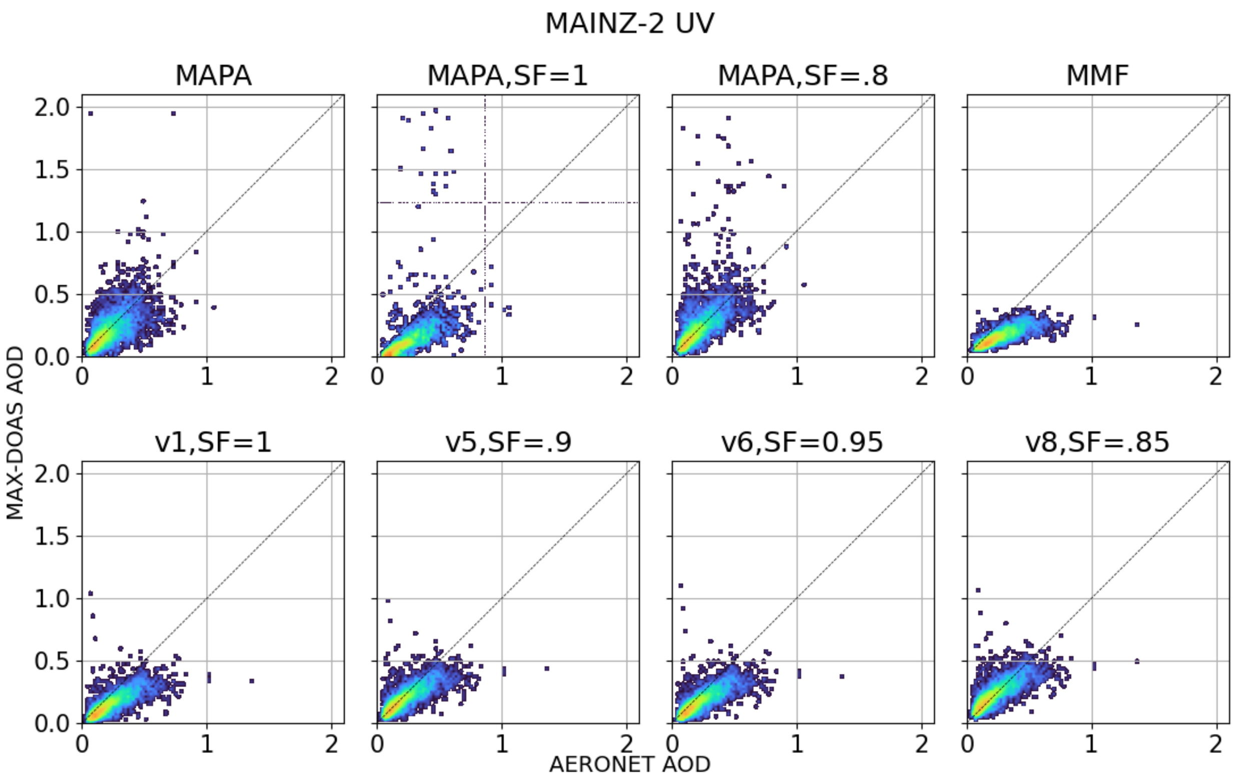


Fig.A.11: As Fig.A.9 but for MAINZ direction 2.

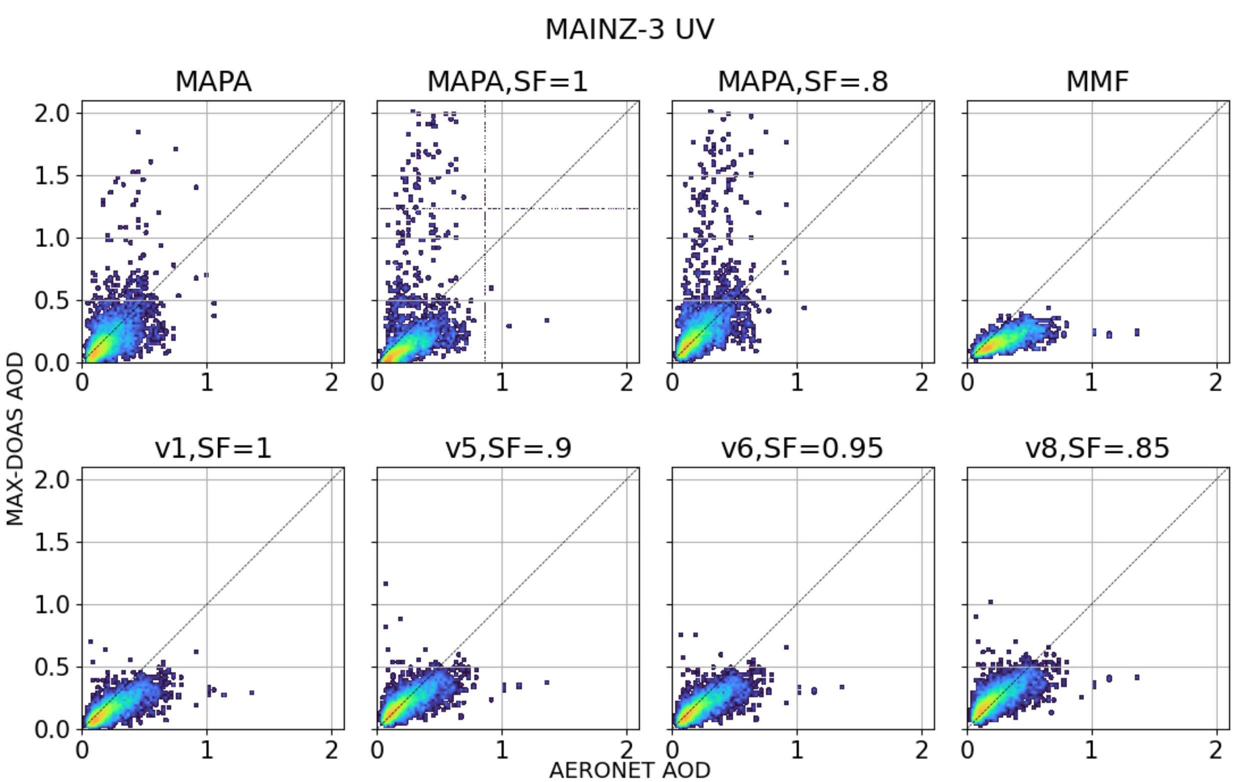


Fig.A.12: As Fig.A.9 but for MAINZ direction 3.

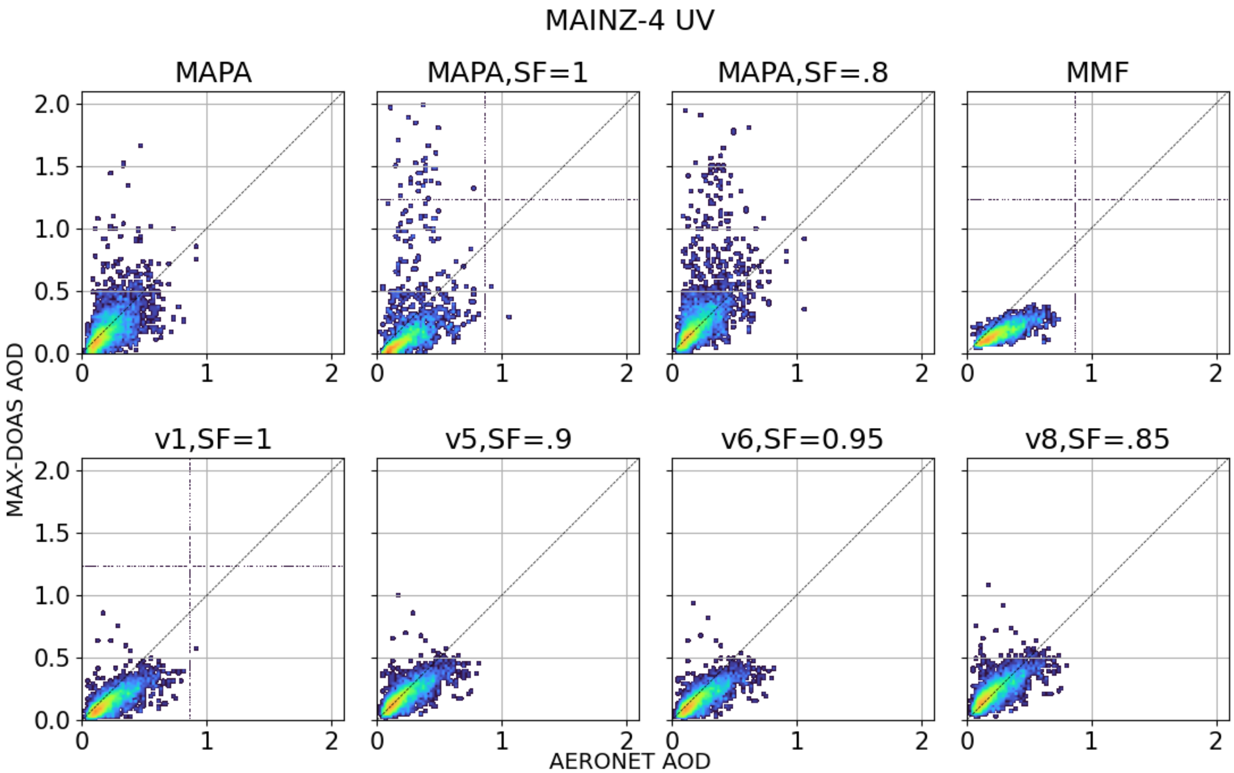


Fig.A.13: As Fig.A.9 but for MAINZ direction 4.

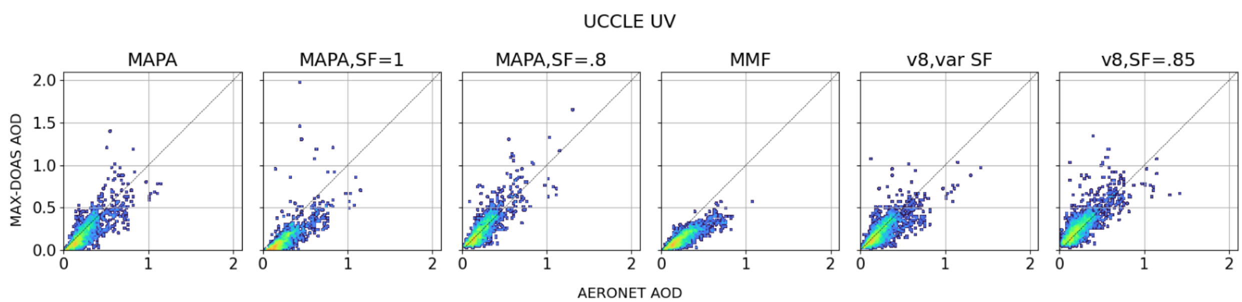


Fig.A.14: As Fig.A.9 but for UCCLE UV.

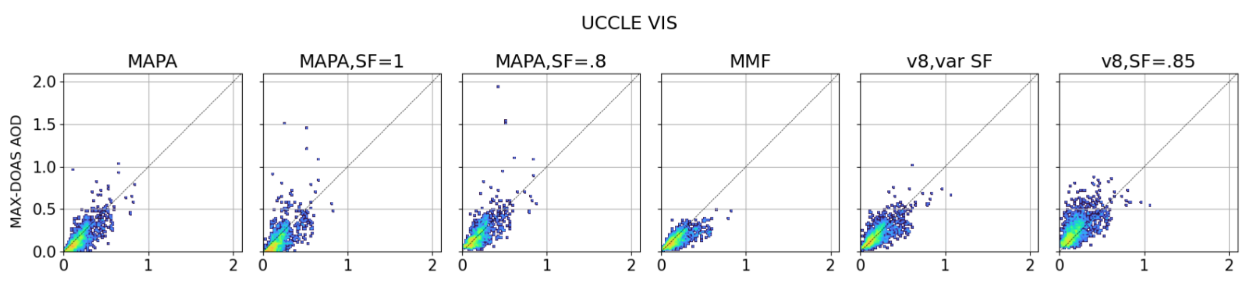


Fig.A.15: As Fig.A.9 but for UCCLE VIS.

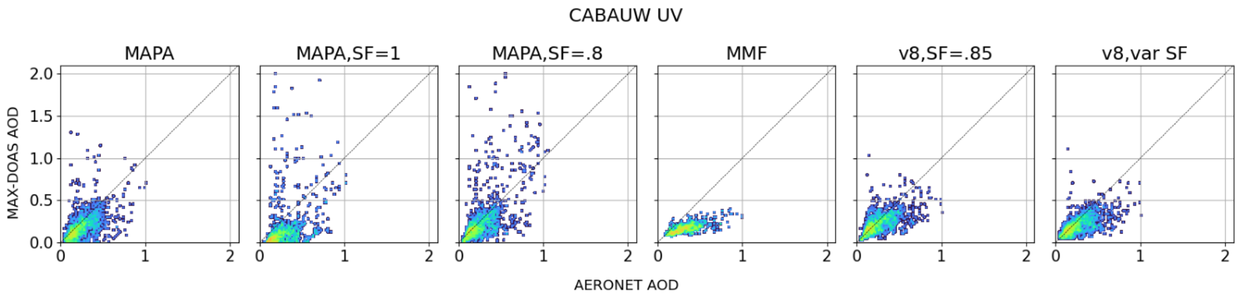


Fig.A.16: As Fig.A.9 but for CABAUW UV.

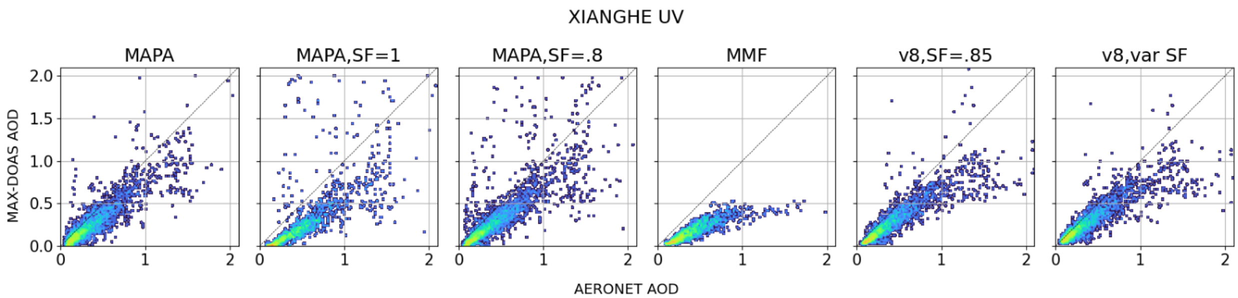


Fig.A.17: As Fig.A.9 but for XIANGHE UV.

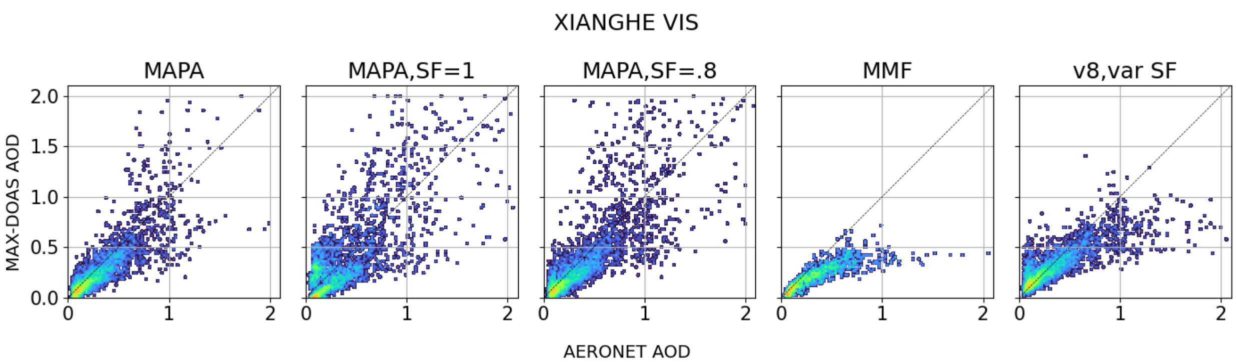


Fig.A.18: As Fig.A.9 but for XIANGHE VIS.