Mainz, 28 January 2019

Dear Editor,

we are happy to submit a revised version of the manuscript amt-2018-375 on The Mainz Profile Algorithm (MAPA).

We have revised the manuscript in response to the comments of the referees during discussion phase. The main changes are:

- The abstract has been revised, now clarifying what is new about MAPA.

- A new subsection on "Heritage and advancements" has been added to the methods section.

- The results for the external cloud classification have been moved from section 4.5 to section 2.8.5, and Fig. 9 becomes Fig. 4 in the revised manuscript. This was necessary in order to include the cloud classification results in Figures 4-8 (now 5-9) as requested by referee #1.

- In response to the suggestions made by referee #1, Figures 5-9 have been revised:

  - the cloud classification results have been added to (a) and (b)

  - correlation coefficients and linear regressions have been added to (f)

  - the results for the variable scaling factor have been included in Fig. 7.

- A new figure (Fig. 10) has been added, showing the ratio of AODs from MAPA vs. AERONET as function of the height parameter *h* in order to support the discussion of the height flag threshold.

Below we have attached

1. the replies to the reviewer comments (point by point), and

2. a revised version of the manuscript with tracked changes.

Kind regards,

Steffen Beirle

Reply to comments from Referee #1

We would like to thank Referee #1 for the thorough review of our manuscript and several helpful hints for improvements. Below we reply to the raised issues one by one.

General comments

Beirle et al. introduce the Mainz Profile Algorithm (MAPA) on the example of measurements taken during the CINDI-2 campaign. The algorithm is based on parametrization and depends on a pre-calculated LUT. The algorithm itself, its a priori assumptions, a flagging scheme, as well as the still discussed and unsolved issue of an O4 scaling factor (SF) are thoroughly discussed. The manuscript is well structured and the results show good agreement with independent measurements. However, the authors should clarify three major issues:

1. A new version of MAPA is presented but the description of differences to older versions is split up across the complete manuscript (e.g. in Sec. 1, 2.3, 2.5). Please provide one single section with differences to the older versions and relevant improvements.

We have revised the abstract, clarifying which parts of MAPA are actually new. In addition, we have added a new method section 2.1 on “Heritage and advancements”, pointing out the heritage of previous profile inversion algorithms (parameterization, LUT based) versus the new developments within MAPA v0.98 (completely new python implementation, MC approach for determination of best matching profiles and uncertainties, extensive flagging scheme).

Revised abstract:

Abstract. The Mainz profile algorithm MAPA derives vertical profiles of aerosol extinction and trace gas concentrations from MAX-DOAS measurements of slant column densities under multiple elevation angles. This manuscript presents (a) a detailed description of the MAPA algorithm (v0.98), (b) results for the CINDI-2 campaign, and (c) sensitivity studies on the impact of a-priori assumptions such as flag thresholds.

Like previous profile retrieval schemes developed at MPIC, MAPA is based on a profile parameterization combining box profiles, which also might be lifted, and exponential profiles. But in contrast to previous inversion schemes based on least-square fits, MAPA follows a Monte Carlo approach for deriving those profile parameters yielding best match to the MAX-DOAS observations. This is much faster, and directly provides physically meaningful distributions of profile parameters. In addition, MAPA includes an elaborated flagging scheme for the identification of questionable or dubious results.

The AODs derived with MAPA for the CINDI-2 campaign show good agreement to AERONET if a scaling factor of 0.8 is applied for O_4, and the respective NO_2 and HCHO surface mixing ratios match those derived from coincident long-path DOAS measurements. MAPA results are robust **with respect** to modifications of the a-priori MAPA settings within plausible limits.

New section:

2.1 Heritage and advancements

MAPA founds on the parameterized profile inversion approach described in Li et al. (2010) or Wagner et al. (2011). It uses similar profile parameter definitions as Wagner et al. (2011) and forward models linking those parameters to dSCD sequences.

Main advancements of MAPA as compared to Wagner et al. (2011) are:

- **MAPA is completely rewritten from the scratch in Python.**
- **All settings are easily adjustable by separate configuration files.**
- **MAPA provides the option of a variable scaling factor for O_4 (see section 2.7)**
- **MAPA uses a Monte-Carlo approach for the profile inversion (see section 2.6), while Wagner et al. (2011) used a least-squares algorithm.**
  The MC approach is faster and provides physically meaningful uncertainty information.
- **MAPA provides an elaborated flagging scheme for the identification of questionable results (section 2.8).**

In the sections below we provide a full description of the MAPA profile inversion algorithm, including also parts which have been described before (like the profile parameterization) for sake of clarity and completeness.

Furthermore, a brief outlook of features (also new nodes for the LUT) which will be implemented in the near feature should be given. It is interesting for users to know which aerosol settings will be available soon (which SSA and asymmetry factors).

We have updated Appendix A to the current state of available LUTs. In addition, we provide MAPA LUTs at ftp://ftp.mpic.de/MAPA/LUTs, and a general MAPA documentation on ftp://ftp.mpic.de/MAPA/documentation/index.html. Additional or extended LUTs will be included there as soon as available.

2. Figure 6 depicts results for a variable scaling factor. Unfortunately, the corresponding SF are not shown. Since these variable SF are also discussed in Section 4.4, it would be interesting to show the variability of the SF and the dependence on different flags and profiles.

We have added the variable scaling factor to Figure 6 d+e (see the reply to the comment on Fig. 4-8 below and Figure R1-2). In addition, we modified the last paragraph of section 3.1 to

Having the option of a variable (best matching) scaling factor is a new feature of MAPA, to our knowledge not provided by any other MAX-DOAS inversion scheme. However, this additional degree of freedom adds complexity, and different effects (like aerosol properties being different from the RTM a-priori, or cloud effects) might be "tuned" to an acceptable match via the scaling factor. As the variable scaling factor has not yet been tested extensively, we focus on the results for a fixed SF of 0.8 as a more "familiar" and transparent setup below, but plan to systematically investigate the results of best matching SFs for various locations and measurement conditions in the near future.

3. The flagging discussion in Section 4 is questionable as specific flags are changed while keeping the other flags at their default values. As the discussion of flagging is valuable because it hasn't been covered thoroughly in other publications, this analysis should be repeated by applying and changing one flag at a time. How else could you know, if the change in one flag does not mainly affect profiles which were already flagged by other thresholds?

Section 4 provides an extensive sensitivity analysis on the impact of a-priori settings. All relevant configuration settings are modified one by one within a plausible range, and the impact on those modifications on MAPA results is judged based on the agreement to AERONET measurements in comparison to the base run performance. We consider this as a reasonable end-to-end analysis which provides information on the crucial parameters and key sensitivities.

This approach exactly refers to the raised question: if we, for instance, change the height flag threshold, we see the impact on the final result, ignoring those cases already flagged by other flags.

Furthermore, it would be interesting to see the actual (AERONET) values of asymmetry factor and SSA, together with the information of the flagging scheme, to identify inaccuracies based on a wrong aerosol assumption.

Further investigations on the impact of aerosol properties like asymmetry and SSA would indeed be interesting. Unfortunately, the number of AERONET aerosol inversion datasets during CINDI2 is very small, such that a systematic investigation of these effects is not possible within this study.

Specific comments

Table A1: Why are the RAA values chosen that coarse for RAA ≥ 30° ? I would expect that results might change a lot for backward scattering, depending on the aerosol phase function, when changing the RAA results from e.g. 180 to 165.

We agree that the RAA nodes in the AMF LUTs should be better resolved, and we will include additional nodes in future RTM calculations. We have added a respective note to Appendix A.

P4, L7: You wrote that p and T profiles are extrapolated when surface values are provided. How is this extrapolation done? How large would you estimate the uncertainties when doing this extrapolation?

We have extended the respective sentence to

If ground measurements at the station are available only, they are used to construct extrapolated profiles **based on a constant lapse rate up to 12 km, and a constant temperature above (see Wagner et al., 2018, section 4.1.1, for details).**

The uncertainty of the resulting O_4 VCD is less than 4% (as derived from a comparison between extrapolated profiles to ECWMF).

P8, L6: I would add that the agreement might be similar but it is also allowed to be slightly worse based on the definition of $R/R_{bm} < F$.

We modified the sentence to

… yields an ensemble of parameter sets with **$R<F \times R_{bm}$, i.e.** similar **(slightly worse)** agreement between measurement and forward model.

P8, L9: Please add here that the weighting with $1/R^2$ is referred to as weighted mean because the question about the weighting might arise in Line 19.

We changed the sentence to

- weighted mean **(wm)** and standard deviation**, with $1/R^2$ as weights**

P8, L25, Fig 2,3: Thank you for changing the line width during the quick access review. However, now the min and max curves are missing. I was wondering about these min/max curves in the first version of these figures.

The curves represent aerosol scenarios with different AOTs (roughly estimated as 0.74, 1.47 on 15/09). How is it possible that these different AOTs do not lead to larger deviations in the O4 dSCD depicted in the corresponding sub-figures? Same for NO2?

In the initial submission of this manuscript, Figures 2&3 included curves for the absolute minima/maxima extinction calculated independently for each height level. It has to be noted, however, that these curves (as well as the percentiles) do not correspond to any actual profile within the ensemble. Thus, the min/max curves must not be interpreted as aerosol scenario, and their integral does not correspond to the respective min/max AODs.

But still the point made by the reviewer generally holds: MAPA results can in fact reveal a high variation of column parameters (of factor 4 and more). I.e., for some scenarios, the forward model finds quite similar agreement to the measured dSCDs for completely different column parameters. However, these cases are flagged by the consistency flag.

In the revised manuscript, we add the following note to the caption of Fig. 2:

Note that the percentiles of vertical profiles are calculated independently for each height level. I.e. they do not correspond to an actual profile from the ensemble, but indicate the general level of uncertainty of vertical profiles.

Fig 2,3: I would suggest to change the x-axis of the EA/dSCD plots to a numbering of EA instead of the actual values. I this way, the more important details for lower EA are easier to identify when using an equidistant spacing.

As the sequence of EAs might be different for other instruments/campaigns, we have decided not to use the (somehow arbitrary) ordinal as abscissa, but to keep the EA value itself. However, in order to emphasize the details for low EAs, we now use a nonlinear scale for EAs in the revised manuscript:

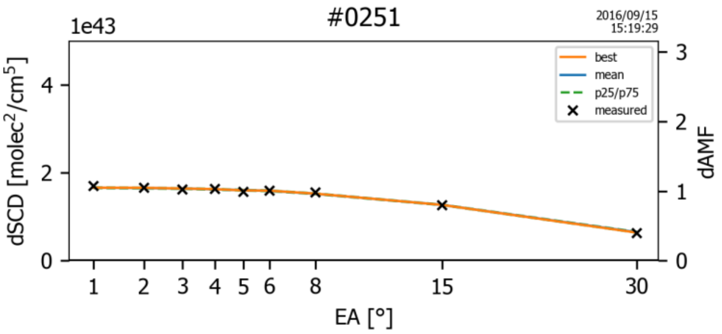

Figure R1-1: Updated layout of the dSCD dependency on EA, using a nonlinear scale for EA.

Tab 4: Since Θ_R scales with S_{err}, please add information about this in Table 4.

S_{err} is the median DOAS fit error, as explained in 2.7.1. For clarity, we now clearly define this quantity already in section 2.2.2 (elevation sequence) and add it to Table 2.

Fig 4-8:

We thank the reviewer for the several suggestions of extensions to Figures 4-8. We have implemented all proposed modifications (including the results for the variable SF as raised in general comment 2), resulting in the following figure for aerosol results for variable SF (corresponding to Fig. 6 of the AMTD manuscript):

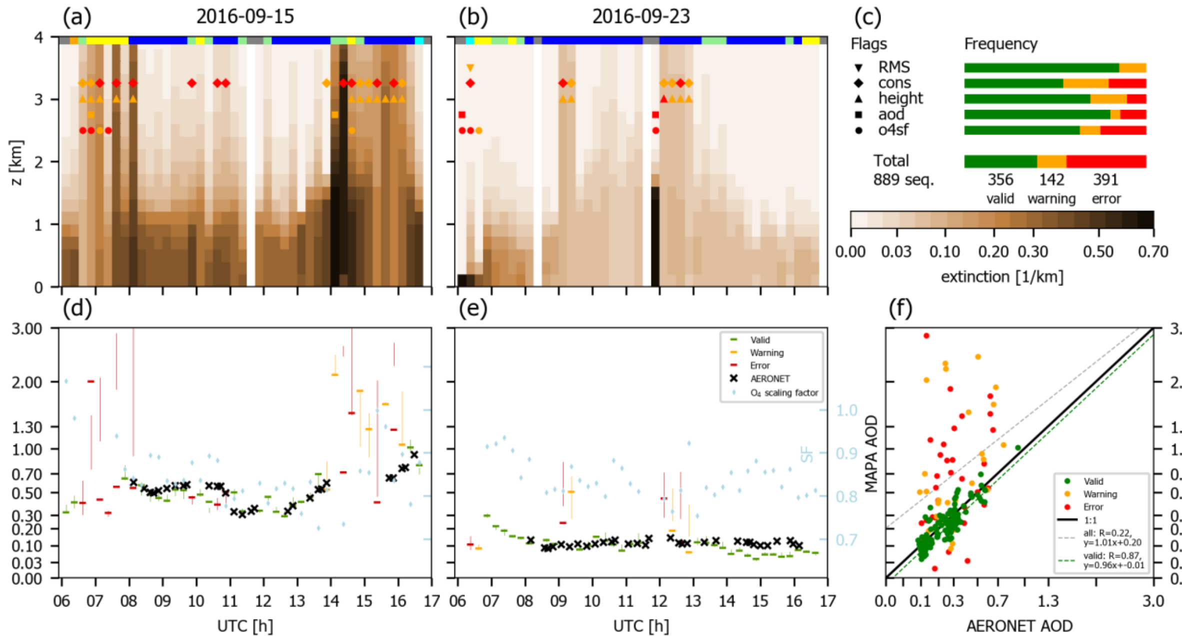

Figure R1-2:.Modified version of Fig. 6, including 1. flagged profiles in (a) and (b),
2. regression lines and correlation coefficients in (f), 3. error bars in (d) and (e),
4. results of the cloud classification in (a) and (b), and 5. the best matching scaling factor in (d) and (e).

We consider this figure to be too overcrowded with information, and to partly distract the reader's attention from the relevant messages. Thus, we only partly follow the reviewer's suggestions for figure modifications in the revised manuscript. Below we reply to the individual proposals, and discuss which changes are adopted in the final figure:

1. Please show all profiles in this kind of plots and use e.g. red rectangles around the flagged profiles to further indicate the discarded scenarios.

If all profiles are included to subplots (a) and (b), the eye is inevitably drawn to the extreme outliers. We consider the flagged profiles to be distractive without providing relevant information. Thus we decided to not include the flagged profiles in the final figures 4-8.

2. It would help if you could add regression lines (and corresponding parameters) and Pearson's correlation coefficient to the figures, for valid and valid+flagged profiles respectively.

We have added regression lines and correlation coefficients to subplot (f) as proposed by the reviewer.

3. Please add error bars to the sub-figures d-f.

Figures d-f display the best value for the fitted column parameter *c* (aerosols) or the lowest layer concentration (trace gases). In Fig. R1-2 we have added the range of ensemble values based on the 25/75 percentiles. These ranges can be quite large for the flagged sequences. For valid sequences, however, these ranges are always small (otherwise, a consistency flag would be raised).

We consider the error bars to be distractive without adding relevant information and have decided not to include them in the revised manuscript.

4. Additional markers for the cloud classification scheme from Section 4.5 should be added to indicate the cloudiness during the corresponding measurement.

We have added cloud classification results to the top of subfigures (a) and (b). In the revised manuscript, we moved Fig. 9 to section 2.7.5. Thus, the legend of cloud classification is now already introduced before Figures 4-8 (becoming 5-9 in the revised manuscript).

Fig 4: I am wondering why MAPA finds nearly all profiles as having issues with the height flag, on 15/09. When considering that the aerosol load was mainly concentrated close to the surface (Fig 5), this indicates an issue with the algorithm or the flagging scheme/threshold. I would not expect a deviation in the profile shape when no SF is applied.

For the investigated MAPA results for CINDI2, this is indeed a clear finding: the resulting profiles are often close to the surface for a scaling factor of 0.8, while the best matching profiles without a SF generally yield higher height parameters, and are thus often discarded by the height flag, often in addition with the consistency flag.

And why is one warning enough to discard the corresponding profile? This appears to be a bit too strict.

For MAPA flagging, we indeed follow a quite conservative approach and decided to raise a total warning already when a single warning occurs.

The comparison studies done within FRM4DOAS (e.g., Frieß et al., 2018) also conclude that the current MAPA flagging seems to be too strict. However, by this strict flagging we nearly exclude all outliers, which are still present in the results from other algorithms (compare e.g. Fig. 16 in Frieß et al., 2018).

The comparison to AERONET and LP-DOAS indicates that the aerosol flags are generally plausible, but trace gas flags are indeed too strict. As trace gas flags are dominated by the total aerosol flag, we will check under which circumstances an aerosol warning might be acceptable within the trace gas retrievals in a future study. We have added the following statement to the conclusions:

The MAPA flagging scheme generally succeeds in identifying dubious results, but a considerable fraction of elevation sequences is flagged. For trace gas profiles, the flagging scheme is dominated by the aerosol flag, which seems to be too strict. It has to be checked under which circumstances an aerosol warning might be acceptable within the trace gas retrievals in a future study.

P13, L29: Please add a time series of this variable scaling factor and a brief discussion.

We have added the variable scaling factor to Fig. 6 (becoming Fig. 7 in the revised manuscript), and slightly extended the discussion of the SF as specified in our reply to general comment 2. An in-depth analysis of the variable SF for various conditions will be focus of a future study.

Fig 6,7: Do you have an explanation why your results and LP DOAS data differ mostly in the morning hours (and the late afternoon)? Is this also a problem for the other days of the investigated time period?

We have investigated the ratio of MAPA to LP DOAS as function of time of day:

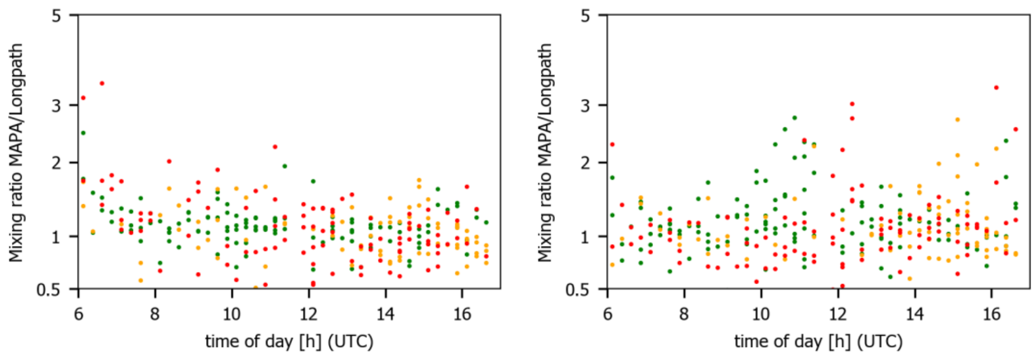

Figure R1-3:.Ratio of lowest layer mixing ratio from MAPA vs. LP-DOAS as function of time of day for HCHO (left) and NO_2 (right).

For formaldehyde, we see no indication for a time-of-day dependency. For NO_2, largest deviations between valid MAPA results and LP-DOAS occur between 10-12, but statistics are rather poor. We have no explanation for this finding, but will keep it in mind and investigate the diurnal cycle for other locations as well.

P16, L6: If a lower F leads to more profiles and the correlation is not deteriorating much, I am wondering why the default is 1.1? Furthermore, when F = 1.3 leads to less profiles due to consistency issues, isn't it possible that the consistency threshold is the problem?

We assume that the reviewer meant to ask *why the default is not 1.1*.

Admittedly, the choice of *F* is somehow arbitrary. Within the MAPA algorithm and flagging scheme, *F* basically fulfils two tasks:

- it allows to retrieve a profile ensemble rather than a single profile, thereby providing uncertainty information

- it allows to check how consistent the ensemble profiles are, thereby providing information about how sensitive the dSCDs are on profile parameters.

It is thus clear that a lower *F* results in more valid sequences, as the consistency flag is deactivated. But we don't see the consistency threshold as "problem", but as a very helpful indicator on which profiles are trustable and which not.

We extended section 4.1. C by the following statement:

For MAPA v0.98 default settings, we stick to the choice of *F*=1.3. But we recommend to also test smaller values for *F* like 1.2 or 1.1, in particular if a large fraction of sequences is flagged by the consistency flag.

P16, L27-28: An increase of the threshold R_n leads to more profiles without a deterioration in r. Could you please test if this is still true for an even larger increase?

We have tested an even higher threshold for R_n. For the investigated CINDI-2 results, this has no effect at all on the resulting AOD and total flag statistics, simply because all effected sequences with R_n>0.1 are already flagged by other criteria (height, consistency, and/or AOD).

P17, L12: "Here we focus of..." → "Here we focus on..."

Fixed.

P17, L14, L29: Here I do not see the point in using 3km v0.96 led to r = 0.826 and Θ h = 4 led to r = 0.783 with 337 and 338 profiles, respectively. One single profile was responsible for this drastic decrease? I would rather say that the individual scenarios at the prevalent site and time led to the conclusion of using 3km instead of 4. This might be completely different for other measurement locations, even though the sensitivity is highest for the lowest altitudes.

As stated in section 4.3, version 0.96 uses slightly different flag definitions (not only thresholds) than version 0.98. Thus, v0.96 and variation d2 correspond to different ensembles and do not differ by just one profile.

The decreasing sensitivity with increasing height is a general restriction of the MAX-DOAS method. We have investigated the dependency of the ratio of MAPA vs. AERONET AOD on the fitted height parameter, and found a clear increase of the error:

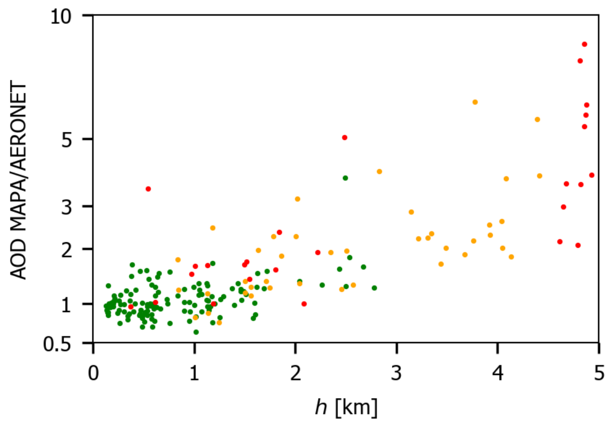

Figure R1-4: Dependency of the AOD from MAPA vs. AERONET as function of the height parameter *h*.

This clearly demonstrates that MAPA results are not at all trustable for high *h* (interestingly, MAPA AOD is always higher than AERONET for these cases). Thus, the height parameter is used for defining the height flag, and to discard all measurements with *h*>3km. From the figure above, this criterion might even be chosen more strictly in the future.

We include this figure to the revised manuscript in order to support the discussion of the threshold for the height flag.

Still, MAPA provides all results also for flagged scenarios, and the user is free to modify the threshold for *h* in the configuration file.

P18, L6: If the variable SF leads to a similar agreement but more profiles remain, why isn't that the preferred option?

As stated above, thorough investigations of the results for variable SF will be performed as next step for various instruments and measurement conditions.

P20, L24: What would be the retrieval response to an exponential scenario? Did the authors make some tests to see if some of the flagged profiles are just exponentially shaped and therefore maybe poorly retrieved by MAPA?

Exponential profiles are included in the comparison study based on synthetic profiles by Frieß et al. (2018). We modified the sentence to

Thus, for synthetic dSCDs based on exponential profiles, the MAPA results try to mimic the exponential shape by a low height parameter and low shape parameter, but performance (in terms of number of valid profiles as well as the agreement of the resulting column parameter) is worse than e.g. for box profiles (see Figures 12 and 16 in Frieß et al., 2018).

With respect to the true profiles during CINDI-2, we cannot check if they are exponentially shaped.

P21, L1: "...profile parameters is derived..." → "...profile parameters are derived..."

Fixed.

Reply to comments from Referee #2

We would like to thank Referee #2 for the thorough review of our manuscript and valuable feedback. Below we reply to the raised issues one by one.

This paper reports a new MAX-DOAS profiling algorithm detailedly. The algorithm is based on a scientific and reasonable method. The results have good correlation with the results from the other instruments. In general the scientific topic is meaningful.

Specific comments:

1, The title of this paper is about a NEW algorithm, so you should highlight what is really NEW and innovative in your algorithm, and what are the advantages comparing to the other MAX-DOAS profiling algorithms. These points should also be included in the Abstracts.

In fact the title does not claim that the paper is about a NEW algorithm. But we understand that it is not fully clear from the current manuscript what is actually new of the described MAPA algorithm. In order to clarify this issue, we extended the abstract and methods respectively:

Revised abstract:

Abstract. The Mainz profile algorithm MAPA derives vertical profiles of aerosol extinction and trace gas concentrations from MAX-DOAS measurements of slant column densities under multiple elevation angles. This manuscript presents (a) a detailed description of the MAPA algorithm (v0.98), (b) results for the CINDI-2 campaign, and (c) sensitivity studies on the impact of a-priori assumptions such as flag thresholds.

Like previous profile retrieval schemes developed at MPIC, MAPA is based on a profile parameterization combining box profiles, which also might be lifted, and exponential profiles. But in contrast to previous inversion schemes based on least-square fits, MAPA follows a Monte Carlo approach for deriving those profile parameters yielding best match to the MAX-DOAS observations. This is much faster, and directly provides physically meaningful distributions of profile parameters. In addition, MAPA includes an elaborated flagging scheme for the identification of questionable or dubious results.

The AODs derived with MAPA for the CINDI-2 campaign show good agreement to AERONET if a scaling factor of 0.8 is applied for O_4, and the respective NO_2 and HCHO surface mixing ratios match those derived from coincident long-path DOAS measurements. MAPA results are robust **with respect** to modifications of the a-priori MAPA settings within plausible limits.

New section:

2.1 Heritage and advancements

MAPA founds on the parameterized profile inversion approach described in Li et al. (2010) or Wagner et al. (2011). It uses similar profile parameter definitions as Wagner et al. (2011) and forward models linking those parameters to dSCD sequences.

Main advancements of MAPA as compared to Wagner et al. (2011) are:

- **MAPA is completely rewritten from the scratch in Python.**
- **All settings are easily adjustable by separate configuration files.**
- **MAPA provides the option of a variable scaling factor for O_4 (see section 2.7)**
- **MAPA uses a Monte-Carlo approach for the profile inversion (see section 2.6), while Wagner et al. (2011) used a least-squares algorithm.**
  The MC approach is faster and provides physically meaningful uncertainty information.
- **MAPA provides an elaborated flagging scheme for the identification of questionable results (section 2.8).**

In the sections below we provide a full description of the MAPA profile inversion algorithm, including also parts which have been described before (like the profile parameterization) for sake of clarity and completeness.

2, In the chapter about CINDI-2 campaign, the results are compared with the results from other instruments. However, it is also important to compare with the MAX-DOAS result from the same instrument but retrieved with the other algorithms.

We fully agree that comparisons with other inversion algorithms is essential. However, within this study, we focus on the description of the MAPA algorithm itself and selected results.

Within the ESA FRM4DOAS project (http://frm4doas.aeronomie.be/), extensive comparisons of different inversion schemes (both OE and parameter based) have been performed for both synthetic as well as measured dSCD sequences. The results of these studies are or will be published in near future:

- Frieß et al., Intercomparison of MAX-DOAS Vertical Profile Retrieval Algorithms: Studies using Synthetic Data, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-423, in review, 2018.
- Tirpitz et al., MAX-DOAS profiles for CINDI–2, in preparation.
- Richter et al., FRM4DOAS verification report, in preparation.

We have added the following sentence to the conclusions:

Within the FRM4DOAS project, different parameter-based as well as OE-based profile inversion algorithms have been compared extensively for synthetic dSCDs (Frieß et al., 2018) as well as real measurements (Tirpitz et al., in prep.; Richter et al., in prep.).

3, In the description of the algorithm, it is better to use the symbols that are commonly used in the related papers. For example, in Equation (1), it is better to use "AMF" instead of "A", "SCD" instead of "S", and "VCD" instead of "V". In other equations, they have the same problem.

We understand that abbreviations like *AMF* and *VCD* would be easier to digest. Still, we prefer single letters as symbols for variables in all equations (as recommended by NIST: https://www.nist.gov/pml/nist-guide-si-chapter-10-more-printing-and-using-symbols-and-numbers-scientific-and-technical), whereas *AMF* within an equation might be read as *A*x*M*x*F*.

Table 2 helps the reader to quickly understand the meaning of symbols/variables used in the equations throughout the document.

4, How accurate is the retrieval results when the distribution of aerosol and trace gases is high (i.e. 3km).

The profile parameterization used in MAPA includes the height parameter *h*. We have investigated the dependency of the ratio of AOD from MAPA versus AERONET on *h*:

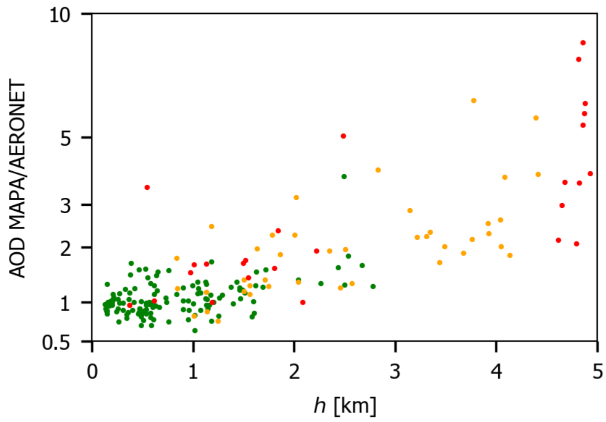

Figure R2-1: Dependency of the AOD from MAPA vs. AERONET as function of the height parameter h.

This clearly demonstrates that MAPA results are not at all trustable for high *h* (interestingly, MAPA AOD is always higher than AERONET for these cases). Thus, the height parameter is used for defining the height flag, and to discard all measurements with *h*>3km. From the figure above, this criterion might even be chosen more strictly in the future.

We have included this figure to the revised manuscript in order to support the discussion of the threshold for the height flag.

In addition, it will be better if the aerosol and trace gases profiles retrieved using MAPA are validated by corresponding profiles measured using other instruments (i.e. air balloon).

We fully agree that accurate independent profile measurements are desirable for validation of MAX-DOAS inversion schemes. Within CINDI-2, some NO_2 sonde measurements have been performed by KNMI, generally revealing polluted boundary layers of about 500m altitude, in agreement with the MAPA profiles. These sonde measurements are included in the extensive CINDI-2 profiling intercomparison by Tirpitz et al. (in preparation).

Minor comments:

1, In Figure7 and 8, "mixing ratio [ppb]" => "Mixing ratio [ppb]"

Done.

2, page4 line1, "to be retrieved first as perquisite for trace gas inversions" => "to be retrieved first as a prerequisite for trace gas inversions"

Fixed.

3, page5 line6, "increase from ground to h" => "increase from the ground to h"

Fixed.

4, Page 5 Line21, "aerosol profiles, and trace gases", "comma" and "and" can't be used together. Delete comma.

We have modified the sentence to

Below, the forward models will be described for both O_4 **(**which is the basis for retrieving aerosol profiles**)** and trace gases.

5, Page 18 Line19, "cloud, and no sequence", "comma" and "and" can't be used together. Delete comma. Correct this mistake throughout your manuscript

5 We are not aware of a general rule that prohibits the usage of "and" after a comma. On the contrary, according to https://www.grammarly.com/blog/comma-before-and/, the usage of a comma before "and" is needed when joining two independent clauses.

We will ask the Copernicus copyeditor for guidance for the respective sentence.

6, Page 7 Line25, "if lowest R is" to "if the lowest R is"

We have modified the sentence to

"if lowest **RMS values** are always found for …".

7, Page 15 Line12, "we focus of variations of" to "we focus on variations of"

Fixed.

8, Page 18 Line30, "cloud scenes still remains" to "cloud scenes still remain"

Fixed.

9, Page 19 Line21, "Currently, an MAX-DOAS" to "Currently, a MAX-DOAS"

Fixed.

The Mainz Profile Algorithm (MAPA)

Steffen Beirle[1], Steffen Dörner[1], Sebastian Donner[1], Julia Remmers[1], Yang Wang[1], and Thomas Wagner[1]
[1]Max-Planck-Institut für Chemie (MPI-C), Mainz, Germany
Correspondence to: Steffen Beirle (steffen.beirle@mpic.de)

Abstract. The Mainz profile algorithm MAPA derives vertical profiles of aerosol extinction and trace gas concentrations from MAX-DOAS measurements of slant column densities under multiple elevation angles. This manuscript presents (a) a detailed description of the MAPA algorithm (~~v0.98), including the flagging scheme for the identification of questionable or dubious results,~~ v0.98), (b) results for the CINDI-2 campaign, and (c) sensitivity studies on the impact of a-priori assumptions such as flag thresholds.

Like previous profile retrieval schemes developed at MPIC, MAPA is based on a profile parameterization combining box profiles, which also might be lifted, and exponential profiles. ~~The~~ But in contrast to previous inversion schemes based on least-square fits, MAPA follows a Monte Carlo approach for deriving those profile parameters yielding best match to the MAX-DOAS observations~~are derived by a Monte Carlo approach, making MAPA much fasterthan previous parameter-based inversion schemes~~. This is much faster, and directly ~~providing~~ provides physically meaningful distributions of profile parameters. In addition, MAPA includes an elaborated flagging scheme for the identification of questionable or dubious results.

The AODs derived with MAPA for the CINDI-2 campaign show good agreement to AERONET if a scaling factor of 0.8 is applied for ~~O_4~~O_4, and the respective NO_2 and HCHO surface mixing ratios match those derived from coincident long-path DOAS measurements. MAPA results are robust with respect to modifications of the a-priori MAPA settings within plausible limits.

1 Introduction

Multi AXis Differential Optical Absorption Spectroscopy (MAX-DOAS), i.e. spectral measurements of scattered sunlight under different viewing elevation angles, have become a useful tool for the determination of vertical profiles of aerosols and various trace gases within the lower troposphere (e.g., Hönninger and Platt, 2002; Hönninger et al., 2004; Wagner et al., 2004; Wittrock et al., 2004; Clémer et al., 2010; Frieß et al., 2006), which is a key for the validation of trace gas columns derived from satellite measurements.

MAX-DOAS is based on the elevation angle dependency of spectral absorption, i.e. the differential slant column density (dSCD) determined by DOAS (Platt and Stutz, 2008). The profile retrieval is performed in two steps: first, aerosol extinction profiles are derived based on dSCDs of the oxygen dimer O_4. In a second step, the concentration profiles of various trace gases detectable in the UV/vis range (such as nitrogen dioxide, NO_2, and formaldehyde, HCHO) can be determined.

For given aerosol/trace gas profiles, dSCDs of O_4 and atmospheric trace gases can be modeled by radiative transfer models (RTMs) for a sequence of elevation angles. The "profile inversion" consists of inverting this forward model, i.e. finding the extinction/concentration profile where forward modeled and measured dSCDs elevation sequences agree.

Profile inversion can be done based on a regularized matrix inversion method denoted as optimal estimation (Rodgers, 2000). It provides an elaborated mathematical framework yielding the best extinction/concentration profile estimate and the corresponding averaging kernels for a given measurement and a-priori error (e.g., Frieß et al., 2006; Clémer et al., 2010). However, results depend on the a-priori settings, in particular the a-priori profile and its uncertainty, which are generally not known.

An alternative approach involves parameterized profiles (Irie et al., 2008; Li et al., 2010; Wagner et al., 2011; Vlemmix et al., 2011, 2015). The basic idea is to represent vertical profiles by few parameters, typically representing total column, height and shape. The profile inversion then corresponds to finding the best matching parameters. Due to the limited number of parameters, a regularization as used in optimal estimation is not required, and the method makes no a-priori assumptions on the actual profile (except that its shape can be represented by the chosen parameterization).

So far, parameter-based inversion was using non-linear least squares algorithms like Levenberg-Marquardt (LM). This is an established method; however, it has some drawbacks: First, LM is based on local linearisation, while the forward model is typically highly nonlinear in the parameters. As a consequence, the confidence intervals (CI) resulting from LM are symmetric by definition and often result in unphysical values of the fitted parameter $\pm$ CI, like a negative layer height. Second, the profile parameters are often strongly correlated, i.e. different parameter combinations can result in similar profile shapes. This implies the existence of local minima in the minimization task, making LM challenging and slowing down the inversion.

Here we present an alternative parameter-based inversion method using a Monte Carlo (MC) approach: The (finite) space of parameter combinations is covered by random numbers, and those best matching the measurement are kept. This approach directly yields distributions rather than single estimates for each parameter, thereby accounting for the correlation of parameters. In addition, the distributions do not contain unphysical parameters (as occur for LM best estimates $\pm$ CI).

The MC approach used in MAPA v0.98 is much faster than the previous LM implementation. In addition, the information on a distribution of the best matching parameters allows for a straightforward determination of the vertical concentration profiles and their uncertainties. The algorithm can also be easily adopted to additional or different profile parameterizations.

MAPA is included as representative of parameter-based algorithms in the processing chain of Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations (FRM4DOAS), a 2-year ESA project which started in July 2016 (http://frm4doas.aeronomie.be/).

In this paper, the MAPA algorithm v0.98 is described in section 2. Exemplary results for the CINDI-2 campaign are shown in section 3. The dependency of MAPA results on a-priori settings as well as clouds is investigated in section 4. The limitations of profile inversions from MAX-DOAS measurements in general and MAPA in particular are discussed in section 5, followed by conclusions.

Table 1 lists the abbreviations and used within this study. A list of mathematical symbols used for variables and parameters is provided in Table 2.

Table 1 about here.

Table 2 about here.

2 Method

In this section, we describe the MAPA profile inversion algorithm. First, the similarities and differences to existing parameter-based inversion schemes are outlined in 2.1. The measurement principle is shortly described in section 2.2. In section 2.3, the required input to the MAPA algorithm is specified. Section 2.4 describes the profile parameterization. In section 2.5, the forward model, linking profile parameters to elevation sequences of dSCDs, is provided. The profile inversion algorithm is described in section 2.6. Section 2.7 deals with the O_4 scaling factor. Finally, the flagging procedure, in order to identify questionable results and outliers, is explained in section 2.8.

2.1 Heritage and advancements

MAPA founds on the parameterized profile inversion approach described in Li et al. (2010) or Wagner et al. (2011). It uses similar profile parameter definitions as Wagner et al. (2011) and forward models linking those parameters to dSCD sequences.

Main advancements of MAPA as compared to Wagner et al. (2011) are:

- MAPA is completely rewritten from the scratch in Python.
- All settings are easily adjustable by separate configuration files.
- MAPA provides the option of a variable scaling factor for O_4 (see section 2.7).
- MAPA uses a Monte-Carlo approach for the profile inversion (see section 2.6.2), while Wagner et al. (2011) used a least-squares algorithm. The MC approach is faster and provides physically meaningful uncertainty information.
- MAPA provides an elaborated flagging scheme for the identification of questionable results (section 2.8).

In the sections below we provide a full description of the MAPA profile inversion algorithm, including also parts which have been described before (like the profile parameterization) for sake of clarity and completeness.

2.2 MAX-DOAS

With DOAS, slant column densities, i.e. integrated columns along the effective light path, can be determined from spectral measurements of scattered sunlight for molecules with absorption structures in the UV/Vis spectral range (Platt and Stutz, 2008). They can be converted into vertical column densities (VCDs), i.e. vertically integrated columns, by division with the so-called air mass factor.

MAX-DOAS measurements are performed from ground based spectrometers with different elevation angles (EA) α, including zenith sky measurements, in order to derive profile information from the EA dependency of slant column densities.

By using the zenith measurements before and/or after a sequence of different EAs as reference spectrum within the DOAS analysis, so-called differential slant column densities (dSCDs) S, representing the SCD excess compared to zenith viewing geometry, are derived. Analogously, differential air mass factors (dAMFs) A relate the dSCDs S to the VCD V:

$$V = S/A \qquad (1)$$

Note that the DOAS spectral analysis is not part of MAPA, but has to be done beforehand.

2.3 Input

Here we list the basic quantities needed as input for MAPA. A detailed description of the MAPA input file format is provided in the supplement.

2.3.1 Viewing and solar angles

The geometry has to be specified in the MAPA input data, defined by the EA α, the solar zenith angle (SZA) ϑ, and the relative azimuth angle (RAA) φ between viewing direction and direction of the sun. Absolute (solar and viewing) azimuth angles are not needed.

2.3.2 Elevation sequence

A sequence of $i = 1..M$ EAs with corresponding dSCDs $S_i = S(\alpha_i)$ is required for one profile to be retrieved. Below, a dSCD sequence is noted as vector $\boldsymbol{S}$, where the i^{th} component corresponds to α_i. In addition, the corresponding sequence of the DOAS fit error $_{\mathrm{err}}$ is required. Note that the dependency on α is implicit in all vectors below and not written explicitly any more.

In addition, the corresponding sequence of the DOAS fit error $\boldsymbol{S}_{\mathrm{err}}$ is required. We define the typical dSCD error S_{err} as the sequence median DOAS fit errors.

As aerosol profiles have to be retrieved first as a prerequisite for trace gas inversions, each MAPA input file must contain at least one dataset of O_4 dSCDs. In addition, trace gas dSCD sequences can be included as needed.

2.3.3 O_4 VCD

For the MAPA aerosol retrieval, an a-priori O_4 VCD V_{O4} is required for each sequence in order to relate the measured O_4 dSCDs to O_4 dAMFs (see equation 1 and section 2.5). V_{O4} can be provided explicitly in the input data. If missing, it is calculated from temperature and pressure profiles. If full profile measurements are provided in the input, they are used. If ground measurements at the station are available only, they are used to construct extrapolated profiles based on a constant lapse rate up to 12 km, and a constant temperature above (see Wagner et al., 2018, section 4.1.1, for details). If no temperature and pressure information is provided in the MAPA input, ERA-interim data (Dee et al., 2011) from the European Centre for Medium-Range Weather Forecasts (ECMWF) is used for the calculation of V_{O4}.

2.4 Profile parameterization

Within MAPA, vertical profiles $p(z)$ of aerosol extinction and trace gases concentration are parameterized by 3 parameters, similar as in Wagner et al. (2011):

1. the integrated column c (i.e. AOD for aerosols, VCD for trace gases),
2. the layer height h, and
3. the shape parameter $s \in]0,2[$.

A shape parameter of $s = 1$ represents a simple box profile:

$$p(z)_{c,h,s=1} = \begin{cases} c/h & \text{for } z \leq h \\ 0 & \text{for } z > h \end{cases} \tag{2}$$

For a shape parameter of $0 < s < 1$, the fraction s of the total column c is placed within a box. The remaining fraction $(1-s)$ is exponentially declining with altitude:

$$p(z)_{c,h,s<1} = \begin{cases} s \times c/h & \text{for } z \leq h \\ s \times c/h \times \exp(-\frac{z-h}{h} \times \frac{s}{1-s}) & \text{for } z > h \end{cases} \tag{3}$$

A shape parameter of $2 > s > 1$ represents an elevated layer from h_1 to h of thickness h_2:

$$p(z)_{c,h,s>1} = \begin{cases} 0 & \text{for } z < h_1 \\ c/h_2 & \text{for } h_1 < z \leq h \\ 0 & \text{for } z > h \end{cases} \tag{4}$$

with

$$\begin{aligned} h_1 &= (s-1)h \\ h_2 &= (2-s)h \\ h_1 + h_2 &= h \end{aligned} \tag{5}$$

Equations 3 and 4 converge to a box profile for $s \to 1$, thus equations 2 to 4 describe a set of parameterized profiles which are continuous in s. Figure 1 exemplarily displays extinction profiles for c=1 and different heights h and shape parameters s.

Figure 1 about here.

Alternative parameterizations (like a linear increase from the ground to h (compare Wagner et al., 2011), or even completely different profile shapes) might be used instead or in addition in future MAPA versions. This would require the calculation of corresponding look-up tables (LUTs) for dAMFs (see below).

2.5 Forward model

In this section the forward model (fm) is specified which connects the profile parameters c, h, and s, with dSCDs for the given solar and viewing geometry specified by ϑ, φ, and α.

Essentially, the forward model is given by eq. 1: $S = V \times A$, where the dAMF depends on profile parameters and solar and viewing geometry. Within MAPA, dAMFs have been calculated offline with the radiative transfer model McArtim (Deutschmann et al., 2011) for fixed nodes for each parameter, and stored as look-up-table (LUT). Within MAPA profile inversion, these multi-dimensional LUTs are interpolated linearly for the given parameter values. For details on the dAMF LUT properties see Appendix A.

Note that the profile parameterization (sec. 2.4) is the same for aerosols and trace gases. The forward models for aerosols and trace gases, however, are similar (and the profile retrieval is based on the same code as far as possible), but not identical. This is due to the fact that the column parameters c_{aer} and c_{tg} have different meanings in the context of $\boldsymbol{S}$ and V: For aerosols, c equals the AOD τ, which is completely independent from the O_4 VCD. For trace gases, c equals the VCD V_{tg}.

Below, the forward models will be described for both O_4, (which is the basis for retrieving aerosol profiles,) and trace gases.

2.5.1 Forward model for aerosols

For aerosols, the O_4 dAMF is a direct function of the profile parameters $c_{\text{aer}}(\equiv \tau), h_{\text{aer}}, s_{\text{aer}}$ and viewing geometry ϑ, φ:

$$\boldsymbol{A}^{\text{O4}} = f(c_{\text{aer}}, h_{\text{aer}}, s_{\text{aer}})|_{\vartheta,\varphi} \tag{6}$$

The corresponding dSCD is:

$$\boldsymbol{S}^{\text{O4}}_{\text{fm}} = V^{\text{O4}}_{\text{apriori}} \times \boldsymbol{A}^{\text{O4}} \tag{7}$$

The respective VCD of O_4 (or vertical profiles of pressure and temperature, which allow for the calculation of V_{O4}) has to be provided in the MAPA input or is calculated from ECMWF profiles.

2.5.2 Forward model for trace gases

For trace gases, the dAMFs also depend on the aerosol profile parameters as determined from the analysis of O_4 dSCDs[1], but not on the trace gas VCD c_{tg}, as long as optical depths are low (which is a prerequisite for DOAS analysis):

$$\boldsymbol{A}^{\text{tg}} = f(h_{\text{tg}}, s_{\text{tg}})|_{\vartheta,\varphi,c_{\text{aer}},h_{\text{aer}},s_{\text{aer}}} \tag{8}$$

The corresponding dSCD is:

$$\boldsymbol{S}^{\text{tg}}_{\text{fm}} = V^{\text{tg}} \times \boldsymbol{A}^{\text{tg}} = c_{\text{tg}} \times \boldsymbol{A}^{\text{tg}} \tag{9}$$

The trace gas VCD V^{tg} is identical to the column parameter c_{tg}.

[1] Note that it is not possible to directly use an a-priori vertical aerosol extinction profile within MAPA trace gas inversion.

2.6 Profile inversion

The forward model as defined above translates the aerosol and trace gas profile parameters c, h and s into dSCD sequences $\boldsymbol{S}_{\text{fm}}$. Within profile inversion, the task is now to find those model parameters yielding the "best match" (bm) between $\boldsymbol{S}_{\text{fm}}$ and the measured dSCD sequence $\boldsymbol{S}_{\text{ms}}$. Typically, "best match" is defined in terms of least-squares of the residue, i.e. the root-mean-square (RMS)

$$R = \sqrt{\frac{(\boldsymbol{S}_{\text{fm}} - \boldsymbol{S}_{\text{ms}})^2}{M}} \tag{10}$$

is minimized, with M being the number of EAs (i.e. the length of $\boldsymbol{S}$).

In previous parameter-based inversion schemes, the best matching parameters have been determined by non-linear least squares algorithms like Levenberg-Marquardt (Li et al., 2010; Wagner et al., 2011; Vlemmix et al., 2015). This approach, however, has some drawbacks, in particular

- as the parameters are highly correlated and local minima can exist, high computational effort, i.e. multiple minimization calls with different initial values, is needed in order to soundly determine the absolute minimum.

- as the least-squares algorithms are based on local linearisation, the resulting parameter uncertainties are per construction symmetric. The resulting parameter range spanned by the fitted parameter $\pm$ CI is often unphysical (e.g. $h < 0$ or $s > 2$) and thus meaningless.

Within MAPA (from v0.6 onwards), thus a different, Monte-Carlo (MC) based approach is chosen. The idea is to (a) generate multiple random sets of profile parameters, (b) calculate the respective dSCD and RMS, and (c) keep those yielding the best agreement. This approach results in a best matching set of parameters, plus an ensemble of parameter sets with similar low R, which reflects the uncertainty range of the estimated profile parameters, which per construction only contains physically valid values.

Section 2.6.2 describes the details of the MC inversion approach, which is used for the determination of h, s, and c_{aer}. Before that, in section 2.6.1 the determination of c_{tg} is described which is implemented differently by a simple linear fit.

2.6.1 VCD: linear fit

The dSCD forward model is highly non-linear in h and s and also in AOD c_{aer}. These parameters are derived by MC as described in detail the next section.

The trace gas VCD c_{tg}, on the other hand, is just a scaling factor of $\boldsymbol{A}$ (eq. 9). Thus, for a given set of profile parameters, and a given sequence of measured dSCDs, the best matching trace gas VCD $c_{\text{tg}} = V_{\text{bm}}$ can just be determined by a linear fit (forced through origin) of V:

$$V_{\text{bm}} = \frac{\boldsymbol{S}_{\text{ms}} \cdot \boldsymbol{A}}{\boldsymbol{A} \cdot \boldsymbol{A}} \tag{11}$$

(Note that $\boldsymbol{S}$ and $\boldsymbol{A}$ are vectors, and the multiplications are scalar products).

In other words, the best matching V equals the mean of V_i for individual elevation angles, weighted by the respective dAMF (i.e., sensitivity). This is different from Wagner et al. (2004), where V was calculated as simple mean of V_i for the individual elevation angles without weighting.

The same formalism is used to define a VCD uncertainty V_{err} as the weighted mean of dSCD errors (from DOAS analysis) for individual EAs. V_{err} is used as column error proxy within the flagging algorithm in order to decide if the found variability of column parameters is within expectation or not (see section 2.8 for details).

2.6.2 Other profile parameters: Monte-Carlo

Within MAPA, profile parameters are determined by just covering the parameter space by random numbers[2] and keeping the matches. In detail, the following steps are performed:

1. limits are defined for each parameter[3],
2. n_{tot} sets of random parameters are drawn[4,5],
3. the RMS R is calculated for each random parameter set,
4. the lowest RMS is identified as "best match" (bm) R_{bm}, and
5. an ensemble of up to n_{sel} parameter sets with $R/R_{\text{bm}} < F$ is kept.

Table 3 lists the default values for parameter limits, number of randoms, and thresholds for MAPA v0.98. The impact of variations of these settings is discussed in section 4.1.

The steps listed above are iterated 3 times, where the resulting ensemble is used to narrow down the parameter limits for the next iteration. I.e., if lowest ~~R is~~ RMS values are always found for low s, the limits for s will be narrowed for the next iteration. As the total number of randoms stays the same, this procedure results in increasingly finer spacing of random numbers.

The procedure results in a best matching parameter set, plus an ensemble of acceptable parameter sets. For each parameter set, also the corresponding VCD V_{bm} is determined by eq. 11.

Table 3 about here.

2.6.3 Best match and ensemble statistics

MAPA yields the best matching parameter combination. The corresponding vertical profile is given by equations 2-4. In addition, MAPA yields an ensemble of parameter sets with ~~similar~~ $R < F \times R_{\text{bm}}$, i.e. similar (slightly worse) agreement between

---

[2]MAPA also provides the option to fix each of the parameters to a predefined value.

[3]This approach (as well as the implementation the dAMF as a LUT) is only possible since the (physical or plausible) parameter ranges are limited.

[4]By default the random number generator is initialized with a seed β in order to generate reproducible results

[5]Parameter combinations yielding thin elevated layers (less than 50 m thick), which correspond to high s and low h, are excluded, as the respective profiles might not be vertically resolved within the RTM calculation of the dAMF LUT.

measurement and forward model. From this ensemble, the following statistics are derived for both the profile parameters as well as the corresponding vertical profiles:

– ~~mean~~ weighted mean (wm) and standard deviation~~(weighted by~~, with $1/R^2$ ~~)~~as weights,

– 25 and 75 percentiles, and

– absolute minimum and maximum.

The mean profiles are often smeared out; in particular strong vertical gradients (occurring for $s \geq 1$) are smeared. The degree of smearing depends on the variability of parameters within the ensemble, which is determined by R_{bm} and the a-priori threshold for accepted RMS values F.

Note that mean±standard deviation might exceed pyhsical limits for parameters and profiles, similar to LM fit results $\pm$ CI. The 25/75 percentiles avoid this. Only for c_{tg}, which is not determined by MC but by a linear fit, unphysical (negative) VCDs and concentrations can occur. These can be understood as noise for quasi-zero VCDs, and must not be set to 0 or skipped in order to keep unbiased means.

Below we mainly focus on the best match ~~(bm)~~ and weighted mean ~~(wm)~~ of parameters and profiles.

Within trace gas retrievals, aerosol profile parameters are required for accessing the dAMF LUT. For this, the best matching parameters are taken. Due to nonlinearities (the mean of ensemble profiles does not equal the profile corresponding to the mean parameters), it is not possible to take mean parameters for this. If one is interested in the actual aerosol profile and its uncertainty, however, the mean profile and the percentiles might still yield valuable information.

Figure 2 exemplarily displays O_4 dSCDs (top) and the retrieved aerosol extinction profiles (bottom) for an afternoon sequence on 15 (left) and 23 (right) September 2016. Best match, weighted mean, 25/75 percentiles and min/max are shown. For these examples, a scaling factor of 0.8 has been applied for O_4 (see next section). This choice will be justified in section 3.

Figure 3 displays the respective dSCDs and profiles for NO_2.

Figure 2 about here.

Figure 3 about here.

2.7 Scaling of O_4 dSCDs

Some previous studies have reported on a significant mismatch between modeled and measured dSCDs of O_4, which is usually accounted for by applying an empirical scaling factor (SF) f of about 0.8 to the O_4 dSCDs, while other studies (e.g. Ortega et al., 2016) do not see a need for a SF, for reasons still not understood. An in-depth discussion of the O_4 SF is provided in Wagner et al. (2018).

MAPA provides the option for defining a fixed a-priori scaling factor f of e.g. 0.8. Note that within MAPA, the measured dSCD is unchanged (in order to have the same measured dSCD in plots and result files for comparison), but the modeled dSCD is divided by f instead.

Another option arises from the profile inversion procedure: the linear fit of the best matching VCD (eq. 11), used for the determination of c_{tg}, can likewise be used to determine the best matching VCD of O_4. This defines the best matching SF as

$$f_{\mathrm{bm}} = V_{\mathrm{apriori}} / V_{\mathrm{bm}} \tag{12}$$

Note that extreme deviations of f from 1 are flagged later (see section 2.8).

As the issue of the O_4 SF is still not understood and its value or even its need is highly debated within the community, it was decided to always run MAPA with 3 different settings for f within the FRM4DOAS project:

1. no scaling of O_4 dSCDs, i.e. $f \equiv 1$,

2. a SF of $f = 0.8$,

3. a variable (best matching) SF f_{bm}.

This setup has also been adopted as default in MAPA v0.98. The comparison of the MAPA results for the different settings for f for different campaigns, instruments, and conditions hopefully will help to clarify the SF issue in the future.

2.8 Flags

The profile inversion scheme as described in section 2.6 just searches for the parameter combinations yielding best agreement in terms of lowest R. Thus, it will always result in a "best match", even if the agreement between measured and modeled dSCDs is actually poor, or the resulting parameter ensembles are inconsistent. Therefore, additional information is needed in order to evaluate whether the resulting profile is trustable or not.

Within MAPA, flags raising warnings or errors are provided based on the performance of the profile inversion. Note that output is generated for each elevation sequence, also for those flagged by an error, and the final decision on which profiles are considered as meaningful is in the users hand. Nevertheless, we strongly recommend to consider the raised warnings and errors; error flags should generally lead to a rejection of the affected profiles.

In this section we describe the warning and error flag criteria and thresholds for MAPA v0.98. The thresholds, denoted by Θ below, are defined in the flag configuration file and can easily be modified. However, any change should only be made for good reasons and has to be tested carefully.

Within the FRM4DOAS processing chain, MAPA has to provide reasonable output for a wide variety of instruments and measurement conditions, which could not all be tested beforehand. Thus, the general strategy is to have low thresholds for warnings (conservative approach), and higher thresholds for errors, indicating cases which do not make sense at all.

The flags defined in MAPA v0.98 can be grouped in 4 categories:

1. Flags based on the agreement between forward-modelled and measured $\boldsymbol{S}$,

2. Flags based on consistency of the ensemble of derived MC parameters,

3. Flags based on the profile shape, and

4. Miscellaneous.

Below the different flag criteria are explained in detail. The default warning and error thresholds for MAPA v0.98 are listed in table 4.

Table 4 about here.

2.8.1 RMS

The RMS R as defined in eq. 10 reflects the agreement between measured and best matching $\boldsymbol{S}$. Thus R might directly be used for flagging, as high RMS values generally indicate that the forward model is not capable of reproducing the measurement. In order to account for the instrument dependent uncertainty of the measured dSCDs, the flag threshold Θ_R is given in units of the typical (sequence median) DOAS fit error S_{err}.

Since S scales with the actual VCD V and the dAMF $\boldsymbol{A}$, R is generally large for high trace gas columns and/or high dAMFs. The first corresponds to polluted episodes, while the second represents conditions under which the MAX-DOAS technique is particularly sensitive. Both cases are of particular interest, but would often be flagged if just a threshold for R based on typical values is defined.

Thus we also consider the RMS normalized by the maximum dSCD S_{max}:

$$R_{\mathrm{n}} = R/S_{\mathrm{max}} \tag{13}$$

Due to the normalization, R_{n} removes the scaling of R with V and $\boldsymbol{A}$. However, for very low V or $\boldsymbol{A}$, i.e. dSCDs about 0, R_{n} can become quite large and the intrinsic noise of the dAMF LUT (if calculated by MC RTM as McArtim) matters.

Warning and errors are thus only risen if the values for R and R_{n} both exceed the thresholds given in table 4.

2.8.2 Consistency

In addition to the best matching parameters, MAPA derives an ensemble of parameter sets yielding similar agreement in terms of R. But this does not mean that the ensemble parameters are consistent. While different height and shape parameters might be acceptable (and just result in a larger profile uncertainty), the column parameter is an important integrated property of the profile. Thus a consistency flag is defined based on the spread of the column parameter within the ensemble.

In order to evaluate if the spread is acceptable or not, we define ε as proxy of the column uncertainty. For aerosols, ε is defined in absolute terms in the MAPA flag configuration (default: 0.05). For trace gases, ε is set to V_{err}, which is derived from the SCD error $\boldsymbol{S}_{\mathrm{err}}$ provided in the input data according to eq. 11.

Based on ε, we define the tolerated deviation for c as

$$c_{\mathrm{tol}} = \Theta_{\mathrm{abs}} \times \varepsilon + \Theta_{\mathrm{rel}} \times c_{\mathrm{bm}}, \tag{14}$$

consisting of an absolute and a relative term. I.e., for low columns, the tolerance is dominated by ε scaled with the absolute threshold defined in the flag settings, whereas for high columns, the relative term $\Theta_{\mathrm{rel}} \times c_{\mathrm{bm}}$ dominates.

Flags are raised if the ensemble standard deviation of c or the difference between c_{bm} and c_{wm} exceed the column tolerance.

The consistency flag indicates that the observations have been reproduced with comparable RMS by parameter sets with considerably different column parameters. I.e., the dSCD sequence shows no strong dependency on c, and MAXDOAS measurements are thus not sensitive for c under these conditions.

2.8.3 Profile shape

MAXDOAS measurements are sensitive to the lower troposphere up to about 2-3 km (Frieß et al., 2006). Profiles reaching up in the free troposphere thus have to be treated with care. Within MAPA v0.98, these cases are identified and flagged based on two quantities:

- the fitted height parameter h, and
- the integrated profile within the lower troposphere c_{LT} (default: below 4km).

A flag is raised if $h > \Theta_h$ or $c_{\mathrm{LT}}/c_{\mathrm{bm}} < \Theta_{\mathrm{LT}}$, but only if also the column c_{bm} exceeds the column detection limit

$$c_{\mathrm{DL}} = \Theta_{\mathrm{DL}} \times \varepsilon, \tag{15}$$

as for very low columns, the profile shape can not be specified anyhow. Note that per default Θ_{abs} equals Θ_{DL}, thus c_{DL} is the same as the absolute tolerance term in equation 14, but MAPA also allows to have different thresholds for both.

2.8.4 Miscellaneous

In addition, the following flags are defined:

- Missing elevation angles:
  In case of incomplete elevation sequences, an error is raised during the MAPA preprocessing. As profile inversion determines 2–3 parameters for about 2–4 degrees of freedom (Frieß et al., 2006), the number M of available EAs must not be too small, otherwise (default: M <5) an error is raised.

  Note that for the results for CINDI-2 shown in the following sections, all incomplete sequences are removed first, as this is related to missing input data, not to the MAPA performance.

- NaNs:
  Best match, mean and std of c are checked for NaNs. These might occur in case of NaNs present in the input data. NaN values automatically raise an error.

- AOD:
  High AOD likely indicates the presence of clouds. But even in case of cloud free conditions, high AOD indicate complex ratiative transfer conditions. Thus flags are raised if $c_{\mathrm{aer}} \equiv \tau > \Theta_\tau$.

– RAA:

If the relative azimuth angle is too low ($\varphi < \Theta_\varphi$), i.e. the instrument is directed towards the sun, and the AOD is high enough ($c_{\mathrm{aer}} \equiv \tau > \Theta_{\varphi,\tau}$), a warning flag is raised, as for this scenario the forward peak of aerosol scattering matters, which is only roughly captured by the Henyey-Greenstein parameterization used in RTM.

– O_4 scaling factor:

MAPA provides the option to derive a best matching SF for O_4 (see section 2.7). Large deviations of the SF from 1 are flagged according to the thresholds defined in table 4.

2.8.5 Cloud flag

Several studies have characterized cloud conditions based on MAX-DOAS elevation sequences, making use of radiance and color index and their (inter- and intra sequence) variability (Gielen et al., 2014; Wagner et al., 2014, 2016; Wang et al., 2015). While dedicated algorithms have been optimized for specific instruments, it is difficult to automatize these algorithms as MAX-DOAS instruments are usually not radiometrically calibrated. I.e. the thresholds for cloud classification have to be adjusted for each instrument.

Therefore, no automatized cloud flagging algorithm is included within MAPA so far. However, MAPA provides the option to add external cloud flags to the MAPA input. A-priori flags in input data are treated like the other flags during MAPA processing, included in the calculation of the total flag (see below), and written to the MAPA output.

Similarly, also other external flags (like an "instrument failure flag" etc.) can easily be added to the MAPA flagging scheme. In section 4.5, we investigate the impact of clouds on MAPA results We have derived a cloud classification based on the scheme described in Wagner et al. (2016), with thresholds adjusted for CINDI-2. Note that cloud information is missing for some elevation sequences due to missing O_4 dSCDs for single elevation angles. Fig. 4 displays the classification of clouds during CINDI-2 for all elevation sequences as well as for those sequences where AERONET AOD measurements are available. During the campaign, 33% of the sequence are categorized as cloud free. If only sequences with coincident AERONET measurements are considered, 72% are cloud free, and the remaining cases are to equal parts cloud hole conditions or missing cloud information. Only 2% are characterized as broken cloud, and no sequence as continous cloud. Thus, a comparison of MAPA results to AERONET to large extent implies a cloud filtering even if no dedicated cloud flag is available.

Figure 4 about here.

In this study, we do not include the cloud classification to the MAPA flagging scheme, as it is not part of the MAPA algorithm. Instead, we use the external cloud classification in order to investigate how far MAPA flags and results for aerosol retrieval depend on cloud conditions, and how far the current default settings for MAPA flagging succeed in identifying cloudy scenesMAPA flags are able to catch clouded conditions in section 4.5.

2.8.6 Total flag

As final step of the flagging procedure, a total warning or error flag is raised if any of the flags defined above indicate a warning or an error, respectively.

3 Results

In this section we present MAPA results exemplarily for dSCD sequences of O_4, NO_2 and HCHO measured during the Second Cabauw Intercomparison campaign of Nitrogen Dioxide Measuring Instruments (CINDI-2) during September 2016 (Kreher et al., in prep.). We focus on two days, September 15 and September 23, which are mostly cloud free and have also been selected as reference days within CINDI–2 intercomparisons (Tirpitz et al., in prep.). The required O_4 VCD is derived from ECMWF interim temperature and pressure profiles, interpolated in space and time.

For details on the MPIC MAX-DOAS instrument and DOAS fit settings see the supplementary material provided by Kreher et al. (in prep.).

3.1 Aerosols

O_4 dSCDs have been analyzed according to the DOAS settings specified in table A3 in Kreher et al. (in prep.), but with sequential instead of noon reference spectra. Fig. 5 displays the MAPA results based on the original O_4 dSCD sequences. In subplots (a) and (b), the valid vertical extinction profiles are displayed for the two selected days. The invalid sequences are marked by the respective flags (symbols as in (c)). In (d) and (e), the respective timeseries of AOD are shown and compared to AERONET measurements (Dubovik and King, 2000)[6]. In (c), flag statistics are provided for all available measurements during the campaign, covering the period from 9 September to 2 October 2016. Panel (f) displays a scatterplot of MAPA AOD compared to 15 minute AERONET means where available for the full campaign. Note that the scales are not linear in order to cover the different order of magnitude in AOD for the two selected days.

Figure 5 about here.

A large fraction of sequences is flagged (overall, less than 1/4 of all sequences are valid). On 23 September, not a single valid sequence was found from 9:00 to 14:00. Even worse, the remaining AODs do not match AERONET (e.g. afternoon of 23 September).

This poor performance is related to a general mismatch between modeled and measured dSCDs, as has been also found for other campaigns in the past (see Wagner et al., 2018, and references therein). We thus perform another MAPA retrieval with an O_4 SF of $f = 0.8$ (Fig. 6).

Figure 6 about here.

[6]The original level 2 AERONET AOD determined at 440 nm has been transferred to 360 nm by assuming an Ångström exponent of 1

The application of a SF largely improves MAPA performance and the agreement to AERONET. A far higher number of sequences is now categorized as valid. The temporal pattern of AOD generally matches well between MAPA and AERONET. Correlation to AERONET AOD (15 minute averages) is as good as r =0.874 with a mean deviation of 0.012±0.067.

Fig. 6 displays MAPA results based on a variable SF. They are overall similar to the results for a fixed SF of 0.8. For the complete campaign, mean and std of the best matching SF in variable mode are 0.85±0.08.

Figure 6 about here.

Below we focus on the results for $f = 0.8$ as this is clearly defined, whereas the free scaling factor increases the degrees of freedomHaving the option of a variable (best matching) scaling factor is a new feature of MAPA, to our knowledge not provided by any other MAX-DOAS inversion scheme. However, this additional degree of freedom adds complexity, and different effects might affect the best matching SF(like aerosol properties being different from the RTM a-priori, or cloud effects) might be "tuned" to an acceptable match via the scaling factor. As the variable scaling factor has not yet been tested extensively, we focus on the results for a fixed SF of 0.8 as a more "familiar" and transparent setup below, but plan to systematically investigate the results of best matching SFs for various locations and measurement conditions in the near future.

3.2 Nitrogen Dioxide (NO_2)

The MPIC DOAS retrieval for NO_2 has been performed in a fit window slightly different from that of O_4, i.e. 352 to 387 nm. Figure 8 displays MAPA results for NO_2. The bottom row now displays the mixing ratio in the lowest 200 m layer instead of the total column. For comparison, mixing ratios derived from long path (LP) DOAS measurements are shown. The LP measurements have been provided by Stefan Schmitt (IUP Heidelberg). Details on LP instruments and retrieval are given in Pöhler et al. (2010) and Eger et al. (in prep.).

NO_2 profiles are generally far closer to the ground compared to aerosol profiles, which is expected, as sources are located at the ground and the NO_x lifetime of some hours is far shorter than that of aerosols.

Comparison of the NO_2 mixing ratio in the lowest 200m layer to LP measurements yields a correlation of $r = 0.887$. The mean difference between MAPA and LP mixing ratios for valid sequences is 0.84±2.26 ppb.

The flagging is strongly dominated by the aerosol flag inherited from the aerosol analysis.

Figure 8 about here.

3.3 Formaldehyde (HCHO)

HCHO dSCDs have been analyzed according to the DOAS settings specified in table A4 in Kreher et al. (in prep.), but with a sequential instead of a noon reference spectrum.

Figure 9 displays MAPA results for HCHO. Profiles reach up higher than for NO_2 as expected due to HCHO being secondary product in VOC oxidation.

As for NO_2, the flagging is dominated by the aerosol flag. But in addition, several more sequences are flagged, with contributions from all RMS, consistency and profile shape flags.

Comparison of the HCHO mixing ratio in the lowest 200 m layer to LP measurements yields a correlation of $r = 0.937$. The mean difference between MAPA and LP mixing ratios for valid sequences is 0.35±0.56 ppb.

Figure 9 about here.

4 Sensitivity studies

The MAPA profile inversion and flagging algorithms are controlled by a-priori parameters. These have been defined by plausible assumptions. In this section we investigate how sensitive the MAPA results are for different a-priori settings, based on the aerosol retrieval for CINDI-2 applying a fixed SF of 0.8, and its comparison to AERONET.

In section 4.1, the sensitivity on MC settings is investigated. The impact of flagging thresholds is analyzed in section 4.2. Note that flag settings can easily be modified a-posteriori, while different MC settings require a complete reanalysis. Table 5 lists the investigated variations for both MC and flag settings, and the impact on the number of valid sequences and the resulting AOD, as compared to AERONET. It also includes results for a previous MAPA version as well as for different O_4 SF, as discussed in sections 4.3 and 4.4.

Finally, section 4.5 investigates the dependency of MAPA flag statistics on cloud conditions.

Table 5 about here.

4.1 MC settings

In this section, the MC settings as defined in the MAPA MC configuration file are modified one by one.

A Random seed

The random generator can be initialized by the seed β provided in MAPA MC configuration. This allows to generate reproducible results even though the method is based on MC. We have tested two alternative seed values just to check how strong the impact of usage of random numbers is. The number of valid sequences and the results for AOD only change slightly for different random sets.

B Number of randoms

As default, each profile parameter is sampled by a=50 values per variable. I.e. for the height parameter, which is within 0.02 and 5 km, the average spacing of the raster in h dimension is about 0.01 km (note that the average spacing gets smaller in the second and third iteration of the narrowed parameter intervals, see section 2.6.2). The total number of random parameter sets n_{tot} is a to the power of MC variables, i.e. 50^3=125000 for aerosols. This corresponds to a duration of about 3 seconds per elevation sequence on a normal PC.

If a is lowered to 20 (n_{tot}=8000), the profile inversion is much faster. But only 269 instead of 324 sequences are identified as valid. However, the remaining profiles show good agreement to AERONET. If a number of a=100 (n_{tot}=10^6) is chosen, about

20 more sequences are labeled as valid compared to the baseline. But the agreement to AERONET gets slightly worse, and the required time is more than 10 fold.

The impact of a on the number of valid sequences can be understood as for higher a, the parameter space is sampled on finer resolution. Thus the RMS of the best match, R_{bm}, generally becomes lower. Consequently, the parameter ensemble defined by $R < F \times R_{\text{bm}}$ is more homogeneous, and less sequences are flagged as inconsistent.

We found a=50 as good compromise between computation time and the number of valid sequences.

C Ensemble threshold for RMS

MAPA determines the best matching parameter combination by the lowest RMS R. In addition, an ensemble of parameter sets is kept with $R < F \times R_{\text{min}}$. The resulting ensemble allows to estimate the uncertainty of the derived parameters and profiles. Per default, F is set to 1.3. We have tested smaller and higher values for F in scenarios C1 and C2.

For a low value of $F = 1.1$, a far higher number of sequences is characterised as valid. This is due to the variety of parameters in the ensemble is being lowered, and consequently the consistency thresholds are less often exceeded. Another side effect is that also the profile uncertainty estimate, which is derived from the variability of profile parameters, is lowered. For the extreme scenario $F_R \to 1$, only the best matching parameter set would be left, which would be close to the result from LM if the number of randoms is high enough. Interestingly, the agreement to AERONET is slightly worse for a low F.

Contrary, a higher value for F results in less valid sequences (as more sequences are characterized as inconsistent), but the remaining ones show better agreement to AERONET.

For MAPA v0.98 default settings, we stick to the choice of F=1.3. But we recommend to also test smaller values for F like 1.2 or 1.1, in particular if a large fraction of sequences is flagged by the consistency flag.

D Shape parameter limits

The shape parameter s determines the profile shape according to sec. 2.4. Modifying the allowed parameter range thus changes the basic population of possible profile shapes within the random ensemble.

As default, the shape parameter almost covers the nodes of the dAMF LUT, except for s_{min} which is set to 0.2. Changing this to 0.1 means allowing for boxes with long exponential tails, which are likely flagged later by the profile shape flag due to the LT criterium. Setting s_{min}=0.1 worsens the performance (less valid sequences as expected, slightly poorer agreement to AERONET), while a value of 0.5 improves the difference, but not the correlation to AERONET.

Setting s_{max} to 1.5 (i.e. removing very thin elevated layers from the basic population) has almost no effect on the CINDI-2 aerosol results.

4.2 Flag settings

Here we modify the flag settings and thresholds as defined in the MAPA flag configuration file one by one. Except for the thresholds for height parameter and AOD, the default values are halved and doubled.

a RMS

We have changed the RMS thresholds for R and R_n in both directions. A change of the threshold of R has hardly any effect in the case of our CINDI-2 results. This might of course be different for other instruments or measurement conditions.

Lowering the threshold for R_n has a tremendous effect: 86 more sequences would be flagged compared to the default. The remaining sequences show a better correlation, but slightly worse agreement to AERONET AOD. Increasing Θ_{R_n} has only a small effect, as most sequences with high R_n are already flagged by one of the other criteria.

b Column uncertainty proxy

For trace gases, $\varepsilon_{\mathrm{tg}}$ can be determined from the dSCD sequence (see sect. 2.6.1). This is not possible for the aerosol retrieval. Instead, ε_τ has to be defined by the user.

Per default, ε_τ is set to 0.05. A lower/higher value for ε_τ slightly decreases/increases the number of valid sequence, but the agreement to AERONET does hardly change.

c Consistency

The variations of the thresholds related to the consistency flag can be summarized as follows: More strict criteria (c1&c3) result in less valid sequences, but a slightly better agreement to AERONET. Vice versa, less strict criteria (c2&c4) result in more valid sequences with poorer agreement to AERONET. We consider the current default settings as plausible and a good compromise.

d Profile shape

Here we focus ~~of~~ on variations of Θ_h. The impact of modifications of Θ_{LT} (not shown) is similar.

If h_{max} is set to 4 km, which was the default value in previous MAPA versions (compare section 4.3), more sequences are labeled as valid, but the agreement to AERONET gets worse. For instance, for the measurements around 16:00 on 15 September, where MAPA AOD is far higher than AERONET, a warning was raised by the height parameter (see Fig. 6 (a) and (d)). For h_{max}=4 km, these sequences are labeled as valid.

If the threshold for h_{max} is lowered to 2 km, less valid sequences remain, but those show significantly better agreement to AERONET, both for correlation and difference. This reflects that MAX-DOAS measurements are mainly sensitive for profiles close to the ground (Frieß et al., 2006). Consequently, inversion results for profiles reaching up to higher altitudes have higher uncertainties.

This is also illustrated in Fig. 10, showing the agreement between MAPA and AERONET AOD as function of the height parameter h.

Figure 10 about here.

e AOD

Modifications of the AOD threshold have almost no effect. This might however be different for measurements under higher aerosol load.

4.3 MAPA version 0.96

In table 5, also the results for previous MAPA version 0.96 are included. This version was used for the FRM4DOAS verification study (Richter and Tirpitz, in prep.).

Versions 0.96 was based on the same MC algorithm with the same MC settings as v0.98. However, the flag definitions and thresholds differ slightly. The main difference is that the height threshold for the profile shape flag was set to 4 km in v0.96. Consequently, v0.96 results in more valid sequences, but with slightly poorer agreement to AERONET AOD, similar as for variation d2.

4.4 Different scaling factors

The results presented above are based on an O_4 SF of 0.8. If instead no scaling factor would be applied, a far higher number of sequences would be flagged, and only 218 sequences remain. These show a good correlation to AERONET, but a systematic bias of -0.115 (compare Fig. 5). The ratio of the mean AOD from MAPA vs. AERONET is 0.53, i.e. MAPA results are too low by a factor of 2 on average if no SF is applied.

If the SF is considered as variable, about 30 more sequences are valid, with similar agreement to AERONET as for a fixed SF of 0.8.

4.5 Clouds

~~As MAX-DOAS measurements are usually not radiometrically calibrated, a cloud classification cannot easily be automatized. Thus, so far no dedicated cloud flag is included in MAPA default settings. In this section we investigate how far MAPA flags and results for aerosol retrieval depend on cloud conditions, and how far the current MAPA flags are able to catch clouded conditions.~~ Figure 11 displays the MAPA flag statistics in dependency of cloud conditions (section 2.8.5) during CINDI–2.

~~We have derived a cloud classification based on the scheme described in Wagner et al. (2016), with thresholds adjusted for CINDI-2. Note that cloud information is missing for some elevation sequences due to missing O_4 dSCDs for single elevation angles.~~

~~Fig. 4 displays the classification of clouds during CINDI-2 for all elevation sequences as well as for those sequences where AERONET AOD measurements are available. During the campaign, 33% of the sequence are categorized as cloud free. If only sequences with coincident AERONET measurements are considered, 72% are cloud free, and the remaining cases are to equal parts cloud hole conditions or missing cloud information. Only 2% are characterized as broken cloud, and no sequence as continous cloud. Thus, a comparison of MAPA results to AERONET to large extent implies a cloud filtering even if no dedicated cloud flag is available.~~

Figure 11 about here.

We have investigated the MAPA flag statistics for different cloud conditions in Fig. 11. For the full campaign, 36% of all sequences are valid. If only cloud free scenes with low aerosol are considered, 68% are valid, while for clouded scenes (broken+continuous clouds), only 13% are valid. Note that the flags for RMS, consistency, height and AOD all contribute significantly to the flagging of clouded scenes.

For the selection of sequences where AERONET is available, 65% sequences are valid.

As demonstrated for For CINDI–2, most clouded cases are successfully flagged in MAPA. But a significant number of cloud hole/broken cloud scenes still remainsremain. We thus recommend that the user applies an additional cloud classification according to e.g. Wagner et al. (2016), and to flag cloud holes with a warning, and continuous and broken cloud scenes with an error.

5 Limitations

In this section we discuss challenges and limitations of MAX-DOAS profile inversion, which have to be kept in mind when interpreting the results and comparing them to other datasets. We start with issues generally affecting MAX-DOAS inversions, followed by MAPA specific issues.

5.1 General limitations of MAX-DOAS profile inversions

In this section we discuss general MAX-DOAS limitations, which also account for optimal estimation algorithms. Still, the issues are discussed from a MAPA perspective.

5.1.1 RTM assumptions

Within forward models, RTM calculations are required which need a-priori information on e.g. aerosol properties like single scattering albedo. If this information is not available and wrong assumptions are made, resulting profiles are biased.

For MAPA, the dAMF LUT used in the forward model has been calculated based on a-priori assumptions as specified in Appendix A. Currently, additional LUTs for different a-priori settings are calculated which might be used alternatively in future and allow to quantify the impact of a-priori RTM assumptions on MAPA results.

5.1.2 Horizontal gradients

Current MAX-DOAS inversion schemes are based on the assumption of horizontally homogeneous layering. In reality, however, aerosol and trace gas distributions reveal horizontal gradients, as can be clearly demonstrated by comparing the results for different azimuthal viewing directions (e.g. Wagner et al., 2011).

It is very challenging to account for horizontal gradients in trace gas inversion algorithms, as (a) the degrees of freedom are numerous (and have to be limited by some simplifications), and (b) fully 3D radiative transfer modelling has to be performed, which is only supported by few RTMs (e.g. McArtim), and far more time consuming.

Currently, ~~an~~ a MAX-DOAS inversion scheme accounting for horizontal gradients is developed at MPIC (Remmers et al., in prep.) based on simultaneous measurements in four azimuth directions. For MAPA, horizontal gradients are so far ignored, but corrections might possibly be added in future versions based on the lessons learned in Remmers et al. (in prep.).

5.1.3 Clouds

Clouds are usually ignored in MAX-DOAS inversion. Thus, elevation sequences affected by clouds have to be flagged. Several algorithms have been proposed for the classification of cloud conditions from MAX-DOAS meaurements (Gielen et al., 2014; Wagner et al., 2014, 2016; Wang et al., 2015), using the zenith values as well as EA dependency of radiances and color indices. However, as MAX-DOAS radiances are usually not calibrated, it is not straightforward to define a universal standardized cloud classification for all kind of instruments. Instead, thresholds have to be adjusted for each instrument.

For CINDI–2, the MAPA flagging scheme raises a warning or error in 87% of all clouded scenes.

5.1.4 O_4 scaling factor

The issue of the O_4 scaling factor is still an unresolved conundrum. MAPA results strongly depend on the choice of the SF. For CINDI–2, a SF of about 0.8 results in much better agreement to AERONET, while the unscaled O_4 dSCDs result in low biased AODs by a factor of 2, and a far higher number of sequences are flagged.

Thus the SF is a general limitation of MAX-DOAS analysis. As shown in Wagner et al. (2018), the discrepancies between modeled and measured $\boldsymbol{S}$ can in some cases not be explained by the involved uncertainties of e.g. temperature and pressure profiles, O_4 cross section uncertainty, etc.

The MAPA option of determining the best matching SF (see section 2.7), allowing to analyse the dependency of the SF on various observation conditions, might help to investigate and hopefully clarify this issue in the future.

5.1.5 Flags

Profile inversions yield a best estimate for aerosol and trace gas profiles, but no direct clue on whether this profile is realistic or not. Thus, within MAPA flags have been defined based on plausibility criteria and basic uncertainty information such as the RMS of the forward model and the DOAS fit error of input dSCDs. The thresholds have been defined carefully and the sensitivity of the a-priori has been investigated in the previous section. But still, it cannot be ruled out that "good" profiles are flagged, as well as that "bad" profiles are not yet flagged.

So far, flags have been investigated based on CINDI-2 measurements and synthetic dSCDs (see Frieß et al., 2018). Further investigations for different instruments and measurement conditions will be made possible by the automatized processing within the FRM4DOAS project. Further extensive validation is desirable, preferably to actual profile measurements from e.g. sondes or drones.

5.2 Specific limitations of MAPA

5.2.1 Profile parameterization

The simple profile parameterization can only represent a limited set of profile shapes. In particular, multi-layer profiles (like a surface-near pollution plus an elevated layer) are not covered by the parameterization.

But also pure exponential profile shapes, which are often assumed in synthetic data and might be considered as "simple" cases, are not directly included in the current MAPA parameterization. They would result from the limit of $h \rightarrow 0$ and $s \rightarrow 0$, but this limit is not covered by the dAMF LUT. ~~Thus~~

Thus, for synthetic dSCDs based on exponential profiles, the MAPA results ~~show less good agreement for exponential profiles compared to~~ try to mimic the exponential shape by a low height parameter and low shape parameter, but performance (in terms of number of valid profiles as well as the agreement of the resulting column parameter) is worse than e.g. ~~box profiles for synthetic data (Frieß et al., 2018)~~for box profiles (see Figures 12 and 16 in Frieß et al., 2018).

5.2.2 dAMF LUT

The dAMF LUT has been calculated with the MC RTM McArtim (Deutschmann et al., 2011). Thus the calculated dAMFs are affected by MC noise. This might become relevant in case of low dAMFs which occur for low VCDs.

In addition, the dAMFs for given geometry and profile parameters ~~is~~ are derived from the multi-dimensional dAMF LUT by linear interpolation, though the dependencies are generally nonlinear.

Based on the MAPA results for synthetic dSCDs (Frieß et al., 2018), both effects can be considered as noncritical.

5.2.3 Averaging Kernels

Averaging kernels are not provided by MAPA. But still, the information on the sensitivity of MAPA for different vertical layers is inwoven in the dAMF LUTs. Further investigations will be made in the future how far the dAMF LUTs used for aerosol and trace gas inversion by MAPA might be used to construct an averaging kernel proxy.

6 Conclusions

The MAinz Profile Algorithm MAPA retrieves lower tropospheric profiles of aerosol extinction and trace gas concentrations from dSCD sequences derived from MAX-DOAS measurements. MAPA is based on a simple profile parameterization. In contrast to previous parameter-based profile inversion schemes, MAPA uses a MC approach to derive a distribution of best matching parameter sets (and associated profiles) rather than just one best solution. This is much faster, can deal with correlation of parameters and multiple minima, and allows to also derive an estimate of profile uncertainties. In addition, a two-stage scheme is provided for flagging probably dubious and errorneous results by warning and error, respectively, based on several criteria.

MAPA aerosol results during CINDI–2 agree well to AERONET AOD only if a scaling factor of 0.8 is applied for O_4, for reasons still not understood. In this context, the option of having a variable SF in MAPA might help to solve this issue in the future. Trace gas results for NO_2 and HCHO agree well to LP measurements. The results are robust with respect to the a-priori settings for MC and flagging.

MAPA flagging removes a large fraction, but not all scenes affected by clouds. It is thus recommended to generally apply an additional cloud flagging. The MAPA flagging scheme generally succeeds in identifying dubious results, but a considerable fraction of elevation sequences is flagged. For trace gas profiles, the flagging scheme is dominated by the aerosol flag, which seems to be too strict. It has to be checked under which circumstances an aerosol warning might be acceptable within the trace gas retrievals in a future study.

MAPA performance is affected by general MAX-DOAS limitations like a-priori assumptions in RTM like aerosol scattering properties, or the usually made assumption of horizontal homogeneity, clouds, and the uncertainty caused by the basic lack of understanding of the O_4 SF.

In addition, complex profiles like multiple layers, which are not adequately reflected by the chosen parameterization, cannot be retrieved.

~~MAPA is~~ Within the FRM4DOAS project, different parameter-based as well as OE-based profile inversion algorithms have been compared extensively for synthetic dSCDs (Frieß et al., 2018) as well as real measurements (Tirpitz et al., in prep.; Richter and Tirpitz . MAPA has been included in the ~~operational processing within the FRM4DOAS project~~FRM4DOAS operational processing chain. This will allow for extensive comparisons to profiles from Optimal Estimation inversion, as well as detailed studies on the O_4 SF, for a variety of instruments and measurement conditions in the future.

Competing interests. None.

Acknowledgements. This study has received funding from the FRM4DOAS project under the ESA contract n°4000118181/16/I-EF.

We would like to thank the CINDI–2 and FRM4DOAS communities for valuable discussions and feedback, in particular Martina Friedrich from BIRA (Brussels, Belgium), Jan-Lukas Tirpitz and Udo Frieß from IUP (Heidelberg, Germany), and Andreas Richter from IUP (Bremen, Germany).

Christian Borger from MPIC (Mainz, Germany) is acknowledged for valuable comments on coding and support in bug fixing.

We thank Stefan Schmitt from IUP Heidelberg for providing Long Path DOAS measurements for NO_2 and HCHO.

We thank Bas Henzing for his effort in establishing and maintaining AERONET measurements at Cabauw.

References

Clémer, K., Van Roozendael, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Spurr, R., Wang, P., and De Mazière, M.: Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, Atmos. Meas. Tech., 3, 863–878, doi:10.5194/amt-3-863-2010, 2010.

Dee, D. P., et al.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553–597, doi: 10.1002/qj.828, 2011.

Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., Platt, U., Prados-Roman, C., Pukite, J., Wagner, T., Werner, B. and Pfeilsticker, K.: The Monte Carlo atmospheric radiative transfer model McArtim: Introduction and validation of Jacobians and 3D features, Journal of Quantitative Spectroscopy and Radiative Transfer, 112(6), 1119-1137, doi: 10.1016/j.jqsrt.2010.12.009, 2011.

Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, J. Geophys. Res., 105, 20673-20696, 2000.

Eger, P., Nasse, J. M., Pöhler, D., Schmitt, S., Platt, U., Improving long-path DOAS measurement accuracy by applying a modified optical setup and use of Laser Driven Light Source, Atmospheric Measurement Techniques Discussions, in preparation.

Frieß, U., Monks, P., Remedios, J., Rozanov, A., Sinreich, R., Wagner, T., and Platt, U.: MAX-DOAS O4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies, J. Geophys. Res., 111, D14203, doi:10.1029/2005JD006618, 2006.

Frieß, U., Klein Baltink, H., Beirle, S., Clémer, K., Hendrick, F., Henzing, B., Irie, H., de Leeuw, G., Li, A., Moerman, M. M., van Roozendael, M., Shaiganfar, R., Wagner, T., Wang, Y., Xie, P., Yilmaz, S. and Zieger, P.: Intercomparison of aerosol extinction profiles retrieved from MAX-DOAS measurements, Atmos. Meas. Tech., 9(7), 3205-3222, doi:10.5194/amt-9-3205-2016, 2016.

Frieß, U., ~~Steffen Beirle, Leonardo~~ Beirle, S., Alvarado Bonilla, ~~Tim~~ L., Bösch, ~~Martina M.Friedrich, François Hendrick, Ankie Piters, Andreas Richter, Michel~~ T., Friedrich, M. M., Hendrick, F., Piters, A., Richter, A., van Roozendael, ~~Elena SpineiLind, Jan-Lukas Tirpitz, Tim Vlemmix, Thomas Wagner, and Yang Wang~~M., Rozanov, V. V., Spinei, E., Tirpitz, J.-L., Vlemmix, T., Wagner, T., and Wang, Y.: Intercomparison of MAX-DOAS Vertical Profile Retrieval Algorithms: ~~Synthetic Studies . In preparation.~~ Studies using Synthetic Data, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-423, in review, 2018.

Gielen, C., Van Roozendael, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., Gillotay, D., and Wang, P.: A simple and versatile cloud-screening method for MAX-DOAS retrievals, Atmos. Meas. Tech., 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Hönninger, G. and Platt, U.: Observations of BrO and its vertical distribution during surface ozone depletion at Alert, Atmos. Environ., 36, 2481–2489, 2002.

Hönninger, G., von Friedeburg, C., and Platt, U.: Multi axis differential optical absorption spectroscopy (MAX-DOAS), Atmos. Chem. Phys., 4, 231-254, doi:10.5194/acp-4-231-2004, 2004.

Irie, H., Kanaya, Y., Akimoto, H., Iwabuchi, H., Shimizu, A., and Aoki, K.: First retrieval of tropospheric aerosol profiles using MAX-DOAS and comparison with lidar and sky radiometer measurements, Atmos. Chem. Phys., 8, 341-350, doi.org/10.5194/acp-8-341-2008, 2008.

Kreher, K., M. Van Roozendael, et al.: Intercomparison of NO2, O4, O3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV-Visible spectrometers, in preparation.

Li, X., Brauers, T., Shao, M., Garland, R. M., Wagner, T., Deutschmann, T., and Wahner, A.: MAX-DOAS measurements in southern China: retrieval of aerosol extinctions and validation using ground-based in-situ data, Atmos. Chem. Phys., 10, 2079-2089, doi:10.5194/acp-10-2079-2010, 2010.

Ortega, I., Berg, L. K., Ferrare, R. A., Hair, J. W., Hostetler, C. A., and Volkamer, R.: Elevated aerosol layers modify the O2-O2 absorption measured by ground-based MAX-DOAS, J. Quant. Spectrosc. Ra., 176, 34–49, doi:10.1016/j.jqsrt.2016.02.021, 2016.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy, Springer-Verlag Berlin Heidelberg, Germany, 2008.

Pöhler, D., Vogel, L., Frieß, U., and Platt, U., Observation of halogen species in the Amundsen Gulf, Arctic, by active long-path differential optical absorption spectroscopy, Proceedings of the National Academy of Sciences, 107 (15), 6582-6587, doi:10.1073/pnas.0912231107, 2010.

Richter, A., and Tirpitz, L.: FRM4DOAS verification report, in preparation.

Remmers, J., et al.: MAX-DOAS inversion for horizontal gradients, in preparation.

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific, 2000.

Tirpitz et al.: MAXDOAS profiles for CINDI–2, in preparation.

Vlemmix, T., Piters, A. J. M., Berkhout, A. J. C., Gast, L. F. L., Wang, P., and Levelt, P. F.: Ability of the MAX-DOAS method to derive profile information for NO2: can the boundary layer and free troposphere be separated?, Atmos. Meas. Tech., 4, 2659-2684, doi: 10.5194/amt-4-2659-2011, 2011.

Vlemmix, T., Hendrick, F., Pinardi, G., De Smedt, I., Fayt, C., Hermans, C., Piters, A., Wang, P., Levelt, P. and Van Roozendael, M.: MAX-DOAS observations of aerosols, formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches, Atmos. Meas. Tech., 8(2), 941-963, doi:10.5194/amt-8-941-2015, 2015.

Wagner, T., Dix, B., Friedeburg, v. C., Frieß, U., Sanghavi, S., Sinreich, R., and Platt, U.: MAX-DOAS O_4 measurements – a new technique to derive information on atmospheric aerosols, (I) Principles and information content, J. Geophys. Res., 109, D22205, doi:10.1029/2004JD004904, 2004.

Wagner, T., Beirle, S., Brauers, T., Deutschmann, T., Frieß, U., Hak, C., Halla, J. D., Heue, K. P., Junkermann, W., Li, X., Platt, U. and Pundt-Gruber, I.: Inversion of tropospheric profiles of aerosol extinction and HCHO and NO2 mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets, Atmospheric Measurement Techniques, 4, 2685-2715, doi:10.5194/amt-4-2685-2011, 2011.

Wagner, T., Apituley, A., Beirle, S., Dörner, S., Frieß, U., Remmers, J., and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, Atmos. Meas. Tech., 7, 1289-1320, https://doi.org/10.5194/amt-7-1289-2014, 2014.

Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R., and Wang, Y.: Absolute calibration of the colour index and O4 absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm, Atmos. Meas. Tech., 9, 4803-4823, https://doi.org/10.5194/amt-9-4803-2016, 2016.

Wagner, T., Beirle, S., Benavent, N., Bösch, T., Chan, K. L., Donner, S., Dörner, S., Fayt, C., Frieß, U., García-Nieto, D., Gielen, C., González-Bartolome, D., Gomez, L., Hendrick, F., Henzing, B., Jin, J. L., Lampel, J., Ma, J., Mies, K., Navarro, M., Peters, E., Pinardi, G., Puentedura, O., Pukite, J., Remmers, J., Richter, A., Saiz-Lopez, A., Shaiganfar, R., Sihler, H., Van Roozendael, M., Wang, Y., and Yela, M.: Is a scaling factor required to obtain closure between measured and modelled atmospheric O4 absorptions? – A case study for two days during the MADCAT campaign, Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-238, in review, 2018.

Wang, Y., Penning de Vries, M., Xie, P. H., Beirle, S., Dörner, S., Remmers, J., Li, A., and Wagner, T.: Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, Atmos. Meas. Tech., 8, 5133-5156, https://doi.org/10.5194/amt-8-5133-2015, 2015.

Wittrock, F., Oetjen, H., Richter, A., Fietkau, S., Medeke, T., Rozanov, A., and Burrows, J. P.: MAX-DOAS measurements of atmospheric trace gases in Ny-Ålesund - Radiative transfer studies and their application, Atmos. Chem. Phys., 4, 955-966, doi:10.5194/acp-4-955-2004, 2004.

Table 1. Abbreviations used in text and for indexing, sorted alphabetically.

Abbreviation	Meaning
aer	Aerosol
AOD	Aerosol optical depth
bm	Best match
CI	confidence interval
CINDI	Cabauw Intercomparison of Nitrogen Dioxide Measuring Instruments
dAMF	differential Air mass factor
DL	detection limit
DOAS	Differential Optical Absorption Spectroscopy
dSCD	differential Slant column density
EA	Elevation angle
ECMWF	European Centre for Medium-Range Weather Forecasts
err	error
fm	Forward model
FRM4DOAS	Fiducial Reference Measurements for DOAS
LM	Levenberg-Marquardt
LT	Lower troposphere
LUT	Look up table
MAX-DOAS	Multi AXis DOAS
MC	Monte Carlo
ms	measured
RAA	Relative azimuth angle
RMS	root mean squared
RTM	radiative transfer model
sel	selected
SF	Scaling factor (for O_4)
SZA	Solar zenith angle
tg	Trace gas
tol	tolerance
tot	total
VCD	Vertical column density
wm	weighted mean

Table 2. Symbols used in this study, sorted chronologically.

Section	Symbol	Meaning
2.2	α	EA
	S	dSCD
	V	VCD
	A	dAMF
2.3	ϑ	SZA
	φ	RAA
	M	number of EAs
	$\boldsymbol{S}$	sequence of dSCDs
	$\boldsymbol{S}_{\mathrm{err}}$	sequence dSCD errors (from DOAS fit)
	S_{err}	median of $\boldsymbol{S}_{\mathrm{err}}$
2.4	z	altitude coordinate
	$p(z)$	vertical profile
	c	column parameter
	$c_{\mathrm{aer}} \equiv \tau$	AOD
	$c_{\mathrm{tg}} \equiv V_{\mathrm{tg}}$	VCD
	h	height parameter
	s	shape parameter
2.5	$\boldsymbol{A}$	dAMF sequence
2.6	R	RMS
	β	Seed of random generator
	d	number of MC variables
	a	sampling per MC variable
	n	number of random parameter sets
	F	tolerance for R compared to minimum
2.7	f	O_4 SF
2.8	ε	column uncertainty proxy
	Θ	Flag threshold

Table 3. Default values for the Monte-Carlo based inversion algorithm for MAPA v0.98.

Variable	Default
β	1
a	50
d	3 (aer)
	2 (tg)
$n_{\text{tot}} = a^d$	125000 (aer)
	2500 (tg)
n_{sel}	100
F	1.3
c_{aer} range	[0.0, 5.0]
h range	[0.02, 5.0] km
s range	[0.2, 1.8]

Table 4. Warning and error threshold default values for MAPA v0.98. The meaning of the thresholds is explained in the text. The default column uncertainty ε is 0.05 for aerosols and V_{err} for trace gases.

Symbol	Description	Warning	Error
Θ_R	Upper threshold for R in units of S_{err}	1	3
Θ_{R_n}	Upper threshold R_{norm}	0.05	0.3
Θ_{rel}	Relative column tolerance	0.2	0.5
Θ_{abs}	Absolute column tolerance in untis of ε	1	4
Θ_{DL}	column detection limit in untis of ε	1	4
Θ_τ	Upper threshold for AOD	2	3
Θ_h	Upper threshold for h	3 km	4.5 km
Θ_{LT}	Lower threshold for LT fraction of total column	0.8	0.5
Θ_φ	Lower threshold for RAA	15	nan
$\Theta_{\varphi,\tau}$	Lower threshold for AOD in order to raise RAA flag	0.5	3
Θ_f	O_4 SF threshold interval (Only affects variable SF mode)	$[0.6, 1.2]$	$[0.4, 1.4]$

Table 5. Variations of a-priori settings (compared to the default) and their impact on the MAPA aerosol retrieval, quantified by the number of valid sequences and the AOD comparison between MAPA and AERONET (correlation coefficient r and difference $\Delta\tau$). The default settings of MAPA v0.98 with a SF of $f = 0.8$ are considered as baseline. Variations A-D refer to settings of the MC algorithm (sect. 4.1). Variations a-e refer to flag thresholds (sect. 4.2). Results for a previous MAPA release, and results for different SF are included as well. For details and discussion see text.

Setup	Variation (Default)	#Valid	r	$\Delta\tau$
f=0.8	-	324	0.874	0.012 $\pm$ 0.067
A1	β=2 (1)	320	0.882	0.014 $\pm$ 0.070
A2	β=1000 (1)	329	0.876	0.014 $\pm$ 0.069
B1	a=20 (50)	269	0.882	0.014 $\pm$ 0.076
B2	a=100 (50)	342	0.860	0.026 $\pm$ 0.088
C1	F=1.1 (1.3)	389	0.872	0.026 $\pm$ 0.072
C2	F=1.5 (1.3)	279	0.908	0.006 $\pm$ 0.058
D1	s_{min}=0.1 (0.2)	311	0.875	0.019 $\pm$ 0.071
D2	s_{min}=0.5 (0.2)	348	0.848	0.004 $\pm$ 0.073
D3	s_{max}=1.5 (1.8)	330	0.887	0.018 $\pm$ 0.067
a1	Θ_R=0.5 (1)	324	0.874	0.012 $\pm$ 0.067
a2	Θ_R=2 (1)	325	0.874	0.012 $\pm$ 0.067
a3	Θ_{R_n}=0.025 (0.05)	238	0.911	0.022 $\pm$ 0.064
a4	Θ_{R_n}=0.1 (0.05)	338	0.874	0.012 $\pm$ 0.067
b1	ε_τ=0.025 (0.05)	311	0.877	0.011 $\pm$ 0.067
b2	ε_τ=0.1 (0.05)	334	0.876	0.014 $\pm$ 0.068
c1	Θ_{rel}=0.1 (0.2)	299	0.894	0.006 $\pm$ 0.054
c2	Θ_{rel}=0.4 (0.2)	340	0.787	0.022 $\pm$ 0.094
c3	Θ_{abs}=0.5 (1)	311	0.877	0.011 $\pm$ 0.067
c4	Θ_{abs}=2 (1)	334	0.876	0.014 $\pm$ 0.068
d1	Θ_h=2 (3) km	307	0.916	0.003 $\pm$ 0.055
d2	Θ_h=4 (3) km	338	0.783	0.032 $\pm$ 0.124
e1	Θ_τ=1 (2)	323	0.874	0.012 $\pm$ 0.067
e2	Θ_τ=3 (2)	327	0.874	0.012 $\pm$ 0.067
v0.96		337	0.826	0.037 $\pm$ 0.126
f=1.0	-	218	0.905	-0.115 $\pm$ 0.043
variable f	-	356	0.873	-0.018 $\pm$ 0.069

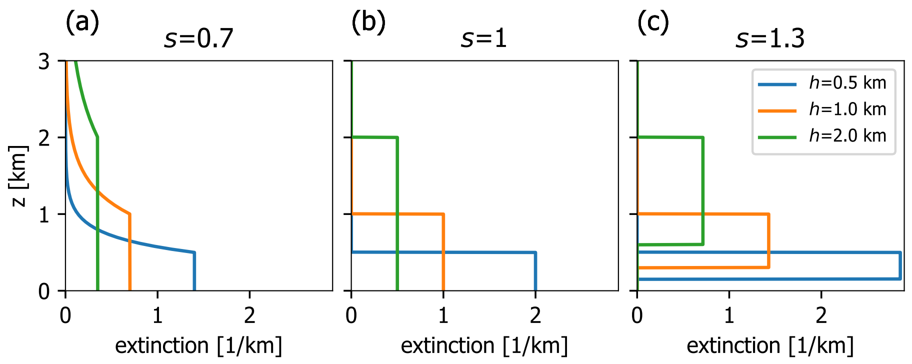

Figure 1. Illustration of the profile parameterization. Aerosol extinction profiles are shown for $c_{\text{aer}} \equiv \tau$=1, different heights h (color coded), and shape parameters $s = 0.7$ (a), 1.0 (b), and 1.3 (c).

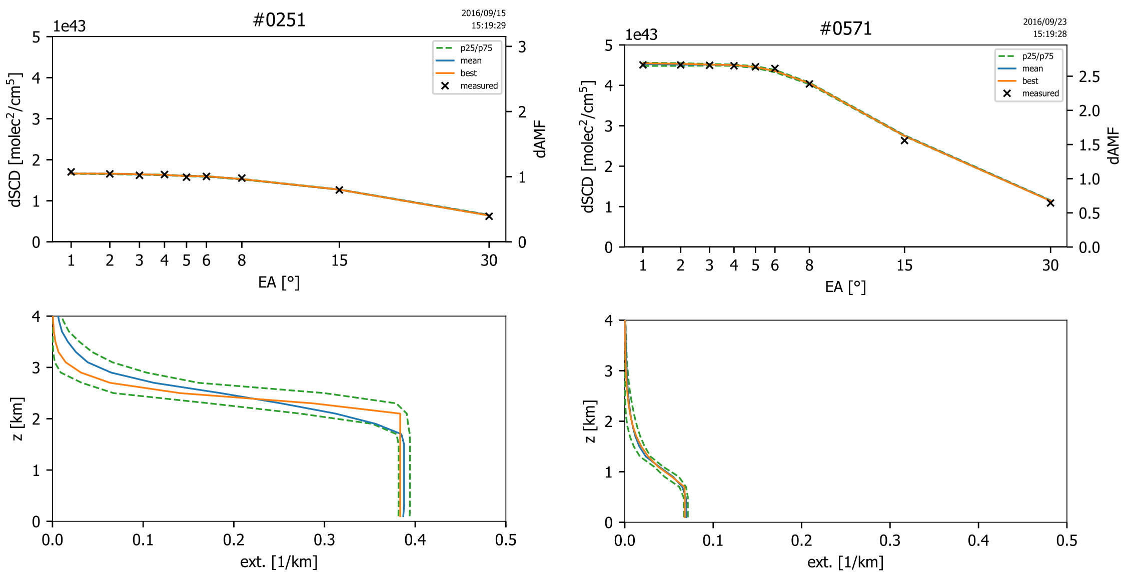

Figure 2. Illustration of the profile inversion for dSCD sequences of O_4 from 15 September (left) and 23 September (right) 2016. A scaling factor of 0.8 has been applied (see section 2.7). Top: measured and modeled dSCDs. The parameter ensembles are represented by statistical key quantities. Bottom: Corresponding vertical profiles. Note that the percentiles of vertical profiles are calculated independently for each height level. I.e. they do not correspond to an actual profile from the ensemble, but indicate the general level of uncertainty of vertical profiles.

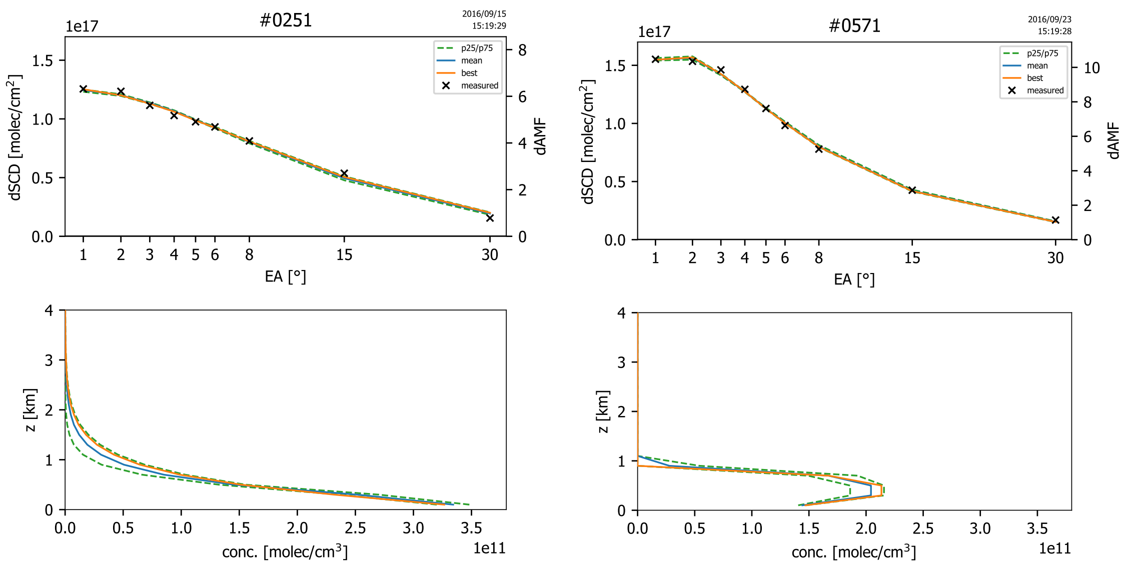

Figure 3. Illustration of the profile inversion for dSCD sequences of NO_2 from 15 September (left) and 23 September (right) 2016, based on the aerosol retrievals shown in Fig. 2. Top: measured and modeled dSCDs. The parameter ensembles are represented by statistical key quantities. Bottom: Corresponding vertical profiles.

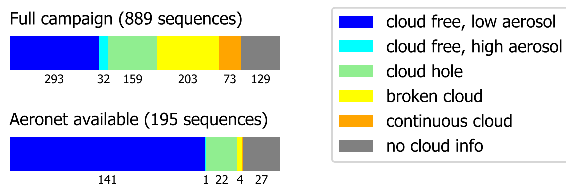

Figure 4. ~~MAPA results for aerosols during CINDI-2. (a) Vertical extinction profile on 15 September. Gaps are flagged~~ Frequence of cloud conditions as ~~warning (orange) or error (red), indicated by different symbols for~~ classified based on the ~~different flag criteria. (b) as (a)~~ procedure described in Wagner et al. (2016) with adjusted thresholds for ~~23 September. (c) Flag statistics for the whole~~ CINDI-2~~campaign~~. ~~(d) AOD from MAPA compared~~ Missing cloud information is related to ~~AERONET~~ missing O_4 dSCDs for ~~15 September~~single elevation angles. ~~(e) as (d) for 23 September~~Top: All available elevation sequences. ~~(f) MAPA AOD compared to~~ Bottom: Only sequences where AERONET ~~for the whole CINDI-2 campaign~~measurements are available.

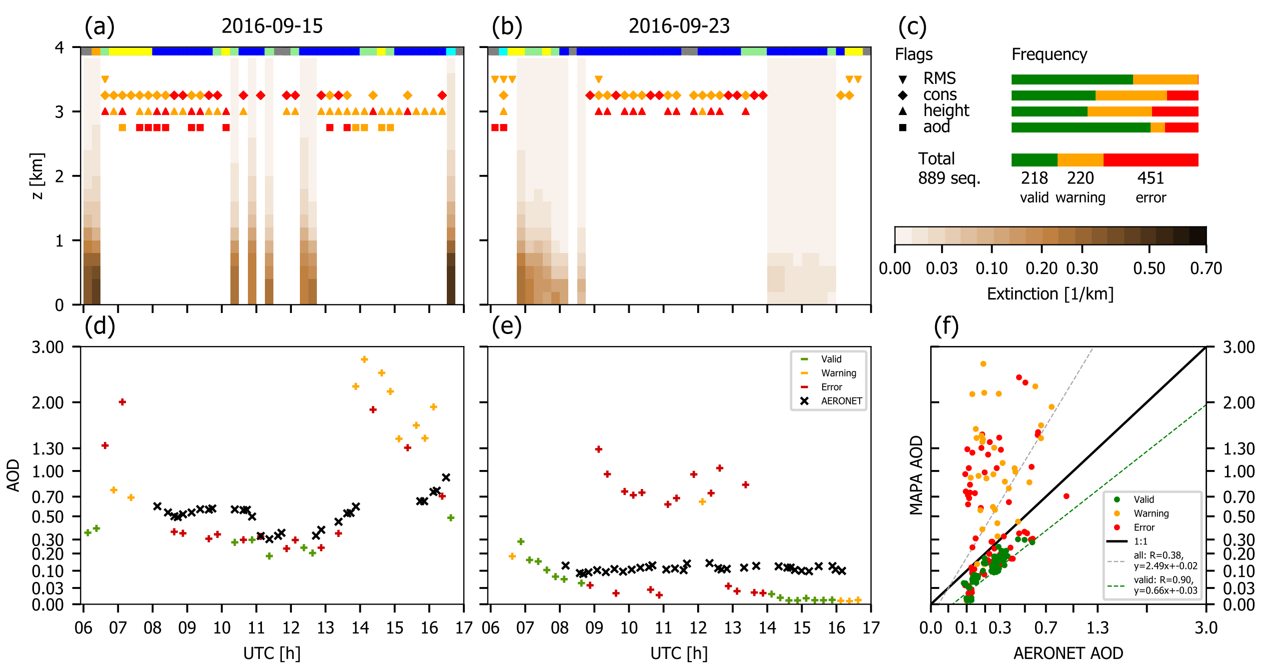

Figure 5. MAPA results for aerosols during CINDI-2. (a) Vertical extinction profile on 15 September. Gaps are flagged as warning (orange) or error (red), indicated by different symbols for the different flag criteria. Results of the cloud classification are provided at the top (for details see section 2.8.5; colors as in Fig. 4) (b) as (a) for 23 September. (c) Flag statistics for the whole CINDI-2 campaign. (d) AOD from MAPA compared to AERONET for 15 September. (e) as (d) for 23 September. (f) MAPA AOD compared to AERONET for the whole CINDI-2 campaign.

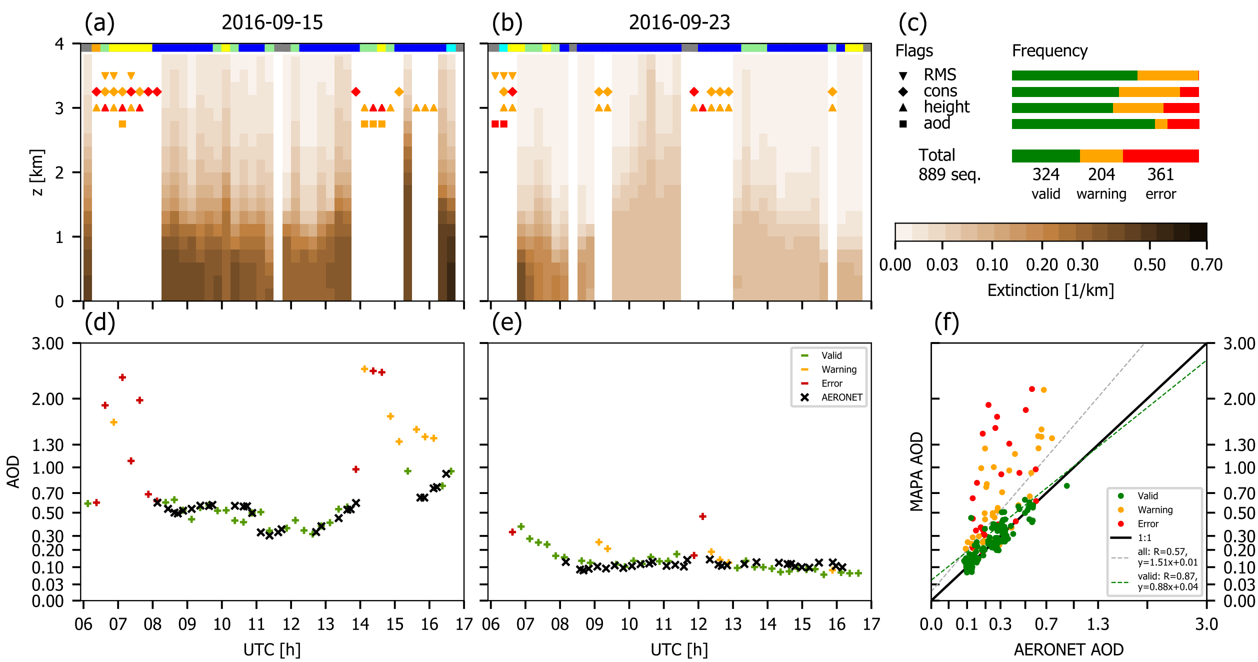

Figure 6. As fig. 5 but for a SF of 0.8.

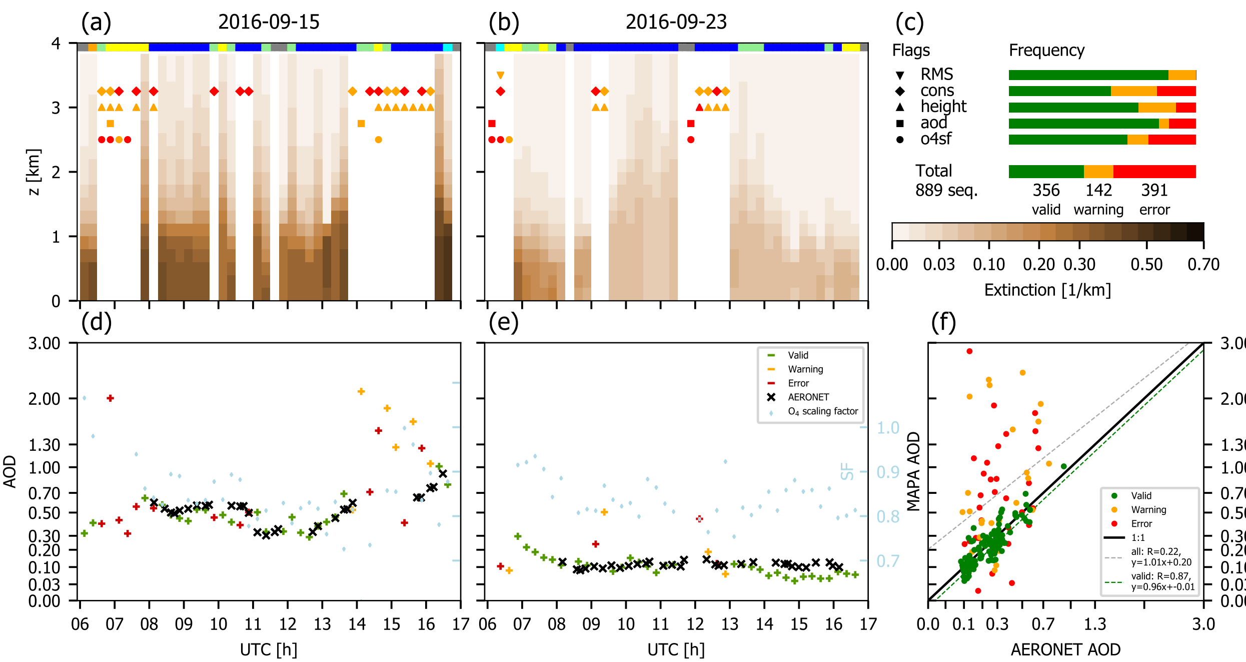

Figure 7. As fig. 5 but for a variable (best matching) SF. The resulting SFs are shown in light blue in subplots (d) and (e) (for scale see right axis of (e)).

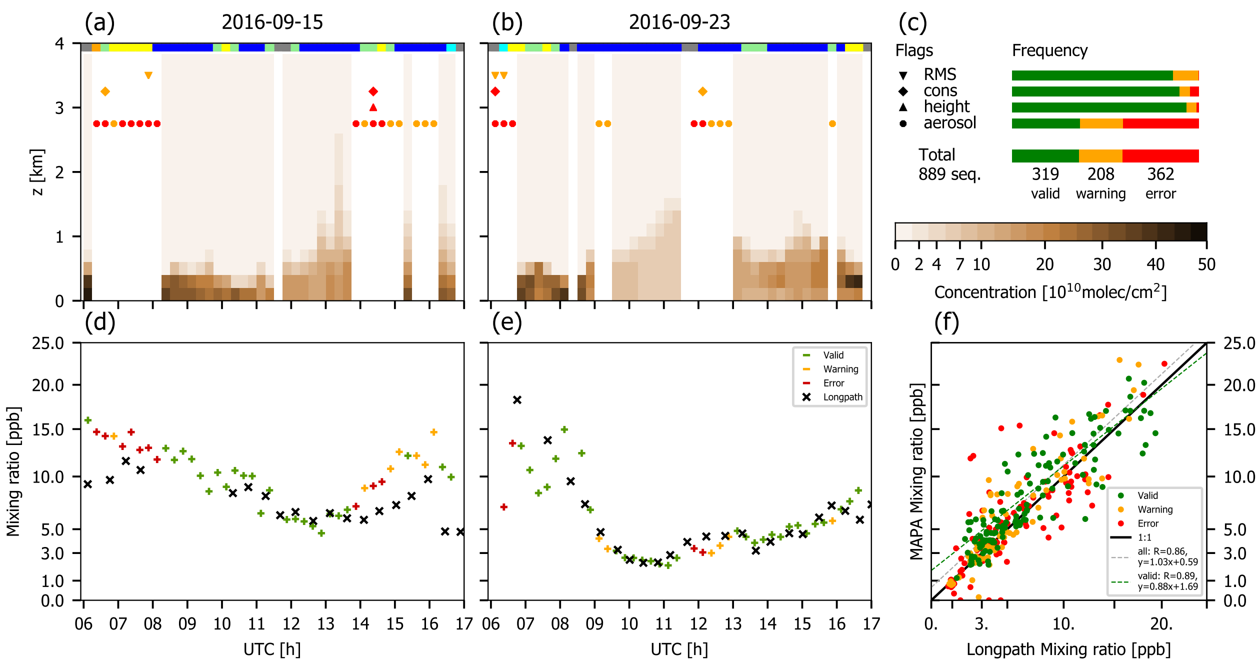

Figure 8. MAPA results for NO_2 during CINDI-2, based on aerosol profiles retrieved with a SF of 0.8. (a) Vertical extinction profile on 15 September. (b) as (a) for 23 September. (c) Flag statistics for the whole CINDI-2 campaign. (d) Mixing ratio in lowest layer (0-200m above ground) from MAPA compared to Long Path (LP) DOAS results for 15 September. (e) as (d) for 23 September. (f) MAPA lowest layer mixing ratio compared to LP for the whole CINDI-2 campaign.

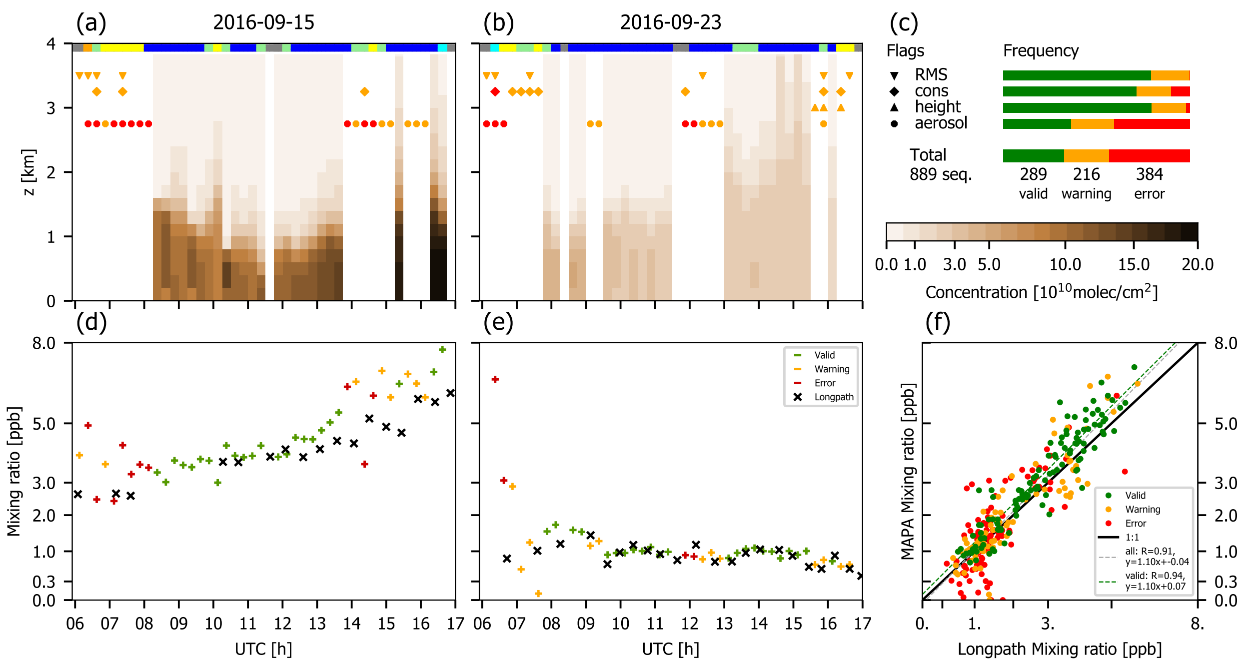

Figure 9. As fig. 8 but for HCHO.

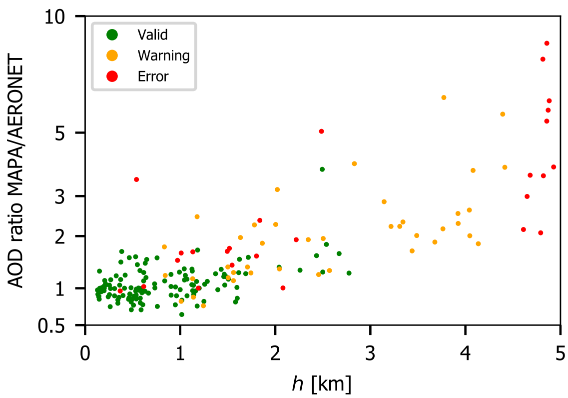

Figure 10. Frequence Dependency of cloud conditions as classified based on the procedure described in Wagner et al. (2016) with adjusted thresholds for CINDI-2ratio of AOD from MAPA vs.Missing cloud information is related to missing O_4 dSCDs for single elevation angles. Top: All available elevation sequences. Bottom: Only sequences where AERONET measurements are availableas function of the height parameter h. Color indicates MAPA flags.

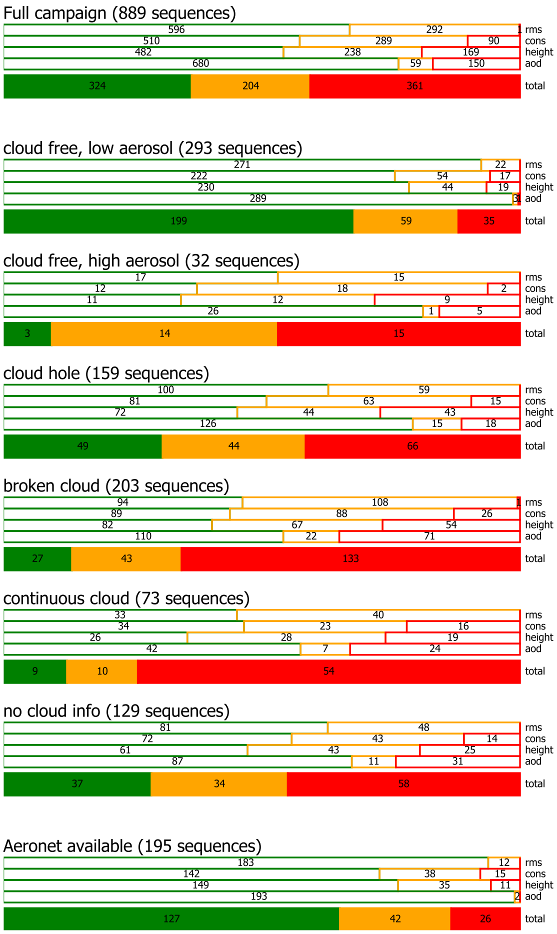

Figure 11. Statistics of MAPA flags for different cloud conditions.

Appendix A: LUTs for dAMFs

dAMFs for O_4 and trace gases are derived from RTM calculations using McArtim (Deutschmann et al., 2011) for a set of viewing geometries and profile parameters. The results are stored in a multi-dimensional LUT in netCDF format, which is interpolated linearly within the MAPA forward model. Table A1 lists the nodes of the parameters in the LUT. Table A2 provides additional settings and a-priori assumptions made for the RTM calculation. Currently, additional LUTs with other settings are calculated (starting with elevated ground altitude which will be automatically be used for elevated stations in future MAPA versions). Future LUT calculations will also provide additional nodes, like φ=170° or $s = 1.1$.

Note that the LUT approach used within MAPA allows for any combination of SZA and RAA, while parameter based profile retrievals shown in previous studies (Wagner et al., 2011; Frieß et al., 2016) were based on LUTs calculated only for the actual SZA/RAA combinations matching the time and place of the measurements.

So far, LUTs are calculated for a set of wavelengths covering the UV and blue spectral range. For a given MAXDOAS-retrieval, MAPA v0.98 just takes the LUT with closest match in wavelength (per default: center of DOAS fit window, can be modified in configuration). In future interpolation in wavelength will also be possible.

Table A1. Nodes of the LUT for dAMFs. Note that other variables like wavelength, detector altitude, or aerosol settings are not included as nodes, but one LUT is determined for each combination of these additional parameters. Compare table A2.

Variable	Symbol	unit	nodes
EA	α	°	1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 30, 45, 90
SZA	ϑ	°	10, 20, 30, 40, 50, 60, 70, 80, 85
RAA	φ	°	0, 5, 10, 20, 30, 60, 90, 120, 150, 180
AOD	$c_{\text{aer}} \equiv \tau$	-	0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 3.0
height	h_{aer}, h_{tg}	km	0.02, 0.1, 0.2, 0.3, 0.5, 0.7, 1.0, 1.2, 1.5, 1.75, 2.0, 2.5, 3.0, 5.0
shape	s_{aer}, s_{tg}	-	0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 1.0, 1.2, 1.5, 1.8

Table A2. RTM settings for LUT calculation. Every combination (so far: different wavelengths) is stored as separate LUT. Further LUTs for other wavelengths, ground altitudes, and aerosol settings are currently calculated and will be provided when ready.

Variable	unit	value(s)
wavelength	nm	315, 325, 343, 360, 410, 430, 477
Single scattering albedo	-	0.95
Henye-Greenstein asymmetry parameter	-	0.68
Ground altitude (above sea level)	m	0
Detector altitude (above ground)	m	0
Ground albedo		0.05