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

FRM4DOAS 2.0

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

Deliverable D2.2:

Algorithm Theoretical Basis Document (ATBD) for the cloud classification algorithm

Date: 17/05/2024

Version: 1.1

Contributing authors:

T. Wagner, L. Reischmann, S. Beirle, S. Ziegler (MPIC)

A. Piters (KNMI)

F. Hendrick, M. Friedrich, M. Van Roozendael (IASB-BIRA)

A. Richter and Tim Bösch (IUP-UB)

U. Friess (IUP-HD)

M. Navarro and M. Yela (INTA)

A. Bais and D. Karagkiozidis (AUTH)

Content

Acronyms and abbreviations

Acronyms and abbreviated terms that are specific for this document can be found below.

AMF	Air Mass Factor
ATBD	Algorithm Theoretical Basis Document
DOAS	Differential Optical Absorption Spectroscopy
DOFS	Number of Degrees of Freedom for Signal
FRM_4DOAS	Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations
FWHM	Full Width at Half Maximum
MPIC	Max-Planck Institute for Chemistry
RTM	Radiative transfer model
SCD	Absolute Slant Column Density
SZA	Solar zenith angle
VCD	Vertical Column Density

1 Introduction

The present document is the ATBD (Algorithm Theoretical Basis Document) for the updated version of the cloud classification algorithm developed as part of the FRM$_4$DOAS-2.0 project. The purpose of the ATBD is to provide a detailed description of the algorithm. The ATBD focusses on the processing of quantities obtained from measured MAX-DOAS spectra in a preceding step. These quantities are the so-called color index (CI, the ratio of the measured intensities at the wavelengths of 330 nm and 390 nm) and the O_4 absorption analysed at 360 nm.

The document is structured as follows: the retrieval and calibration of the CI and the O_4 absorption are described in sections 2 and 3. The original algorithm from Wagner et al. (2016) is introduced in section 4. Section 5 presents recent improvements of the algorithm. In section 6, the validation results for the original and the updated algorithm are presented. Section 7 provides a summary.

2 Calibration of the CI

The CI is determined from the measured signals at 330 nm and 390 nm. An interval of ±0.5 nm around the selected wavelengths is used. The exact choice of the wavelength intervals is not critical, because the effects of slightly different wavelength intervals will be compensated by the calibration procedure of the CI as outlined below.

The CI values depend on the specific properties of a MAX-DOAS instrument. Thus the obtained CI values have to be calibrated before they can be used for further processing. The calibration procedure is described in detail in Wagner et al. (2016). A brief summary is given below:

a) First the extracted CI are normalised (divided) by the simulated minimum CI for cloudy conditions for the same SZA. The minimum CI for cloudy conditions is described by the following polynomial:

$$CI_{min} = -5.261 \cdot S^6 + 8.045 \cdot S^5 + 0.621 \cdot S^4 - 6.588 \cdot S^3 + 3.029 \cdot S^2 + 0.09 \cdot S + 0.66 \quad (1)$$

With S the solar zenith angle in degrees divided by 90.
The normalisation procedure is illustrated in Fig. 1.

b) In the next step, frequency distributions of the normalised CI for SZA < 60° are calculated, see Fig. 2. A distinct accumulation point for the minimum CI is found representing cloudy measurements. The normalised CI of the accumulation point is determined by fitting a Gaussian curve to the peak of the frequency distribution (Fig. 2). The measured CI are then divided by the CI value at the peak:

$$CI_{calibrated} = CI_{meas} / CI_{peak} \quad (2)$$

For the example shown in Fig. 2 the CI of the peak is 0.82.

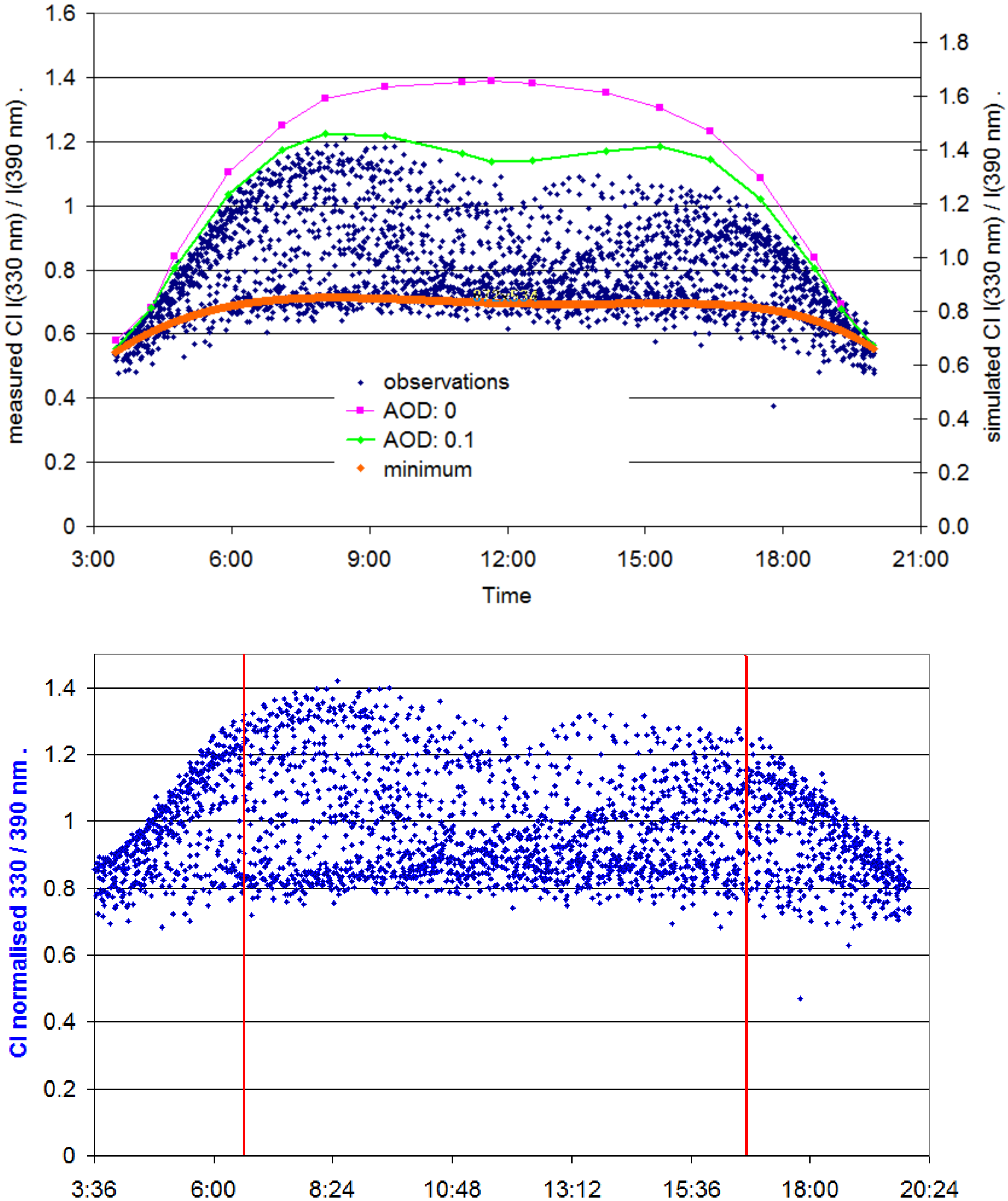

Fig. 1 Top: Comparison of measured (left axis) and simulated (right axis) CI during the CINDI (1) campaign obtained from the MPIC instrument. The thick orange line represents minimum values calculated for cloudy conditions. The green and magenta lines represent CI for cloud-free conditions with low aerosol loads. Bottom: Normalised CI (measured CI divided by the minimum CI). For SZA < 60° (indicated by the red vertical lines) the minima of the normalised CI are almost independent from the SZA.

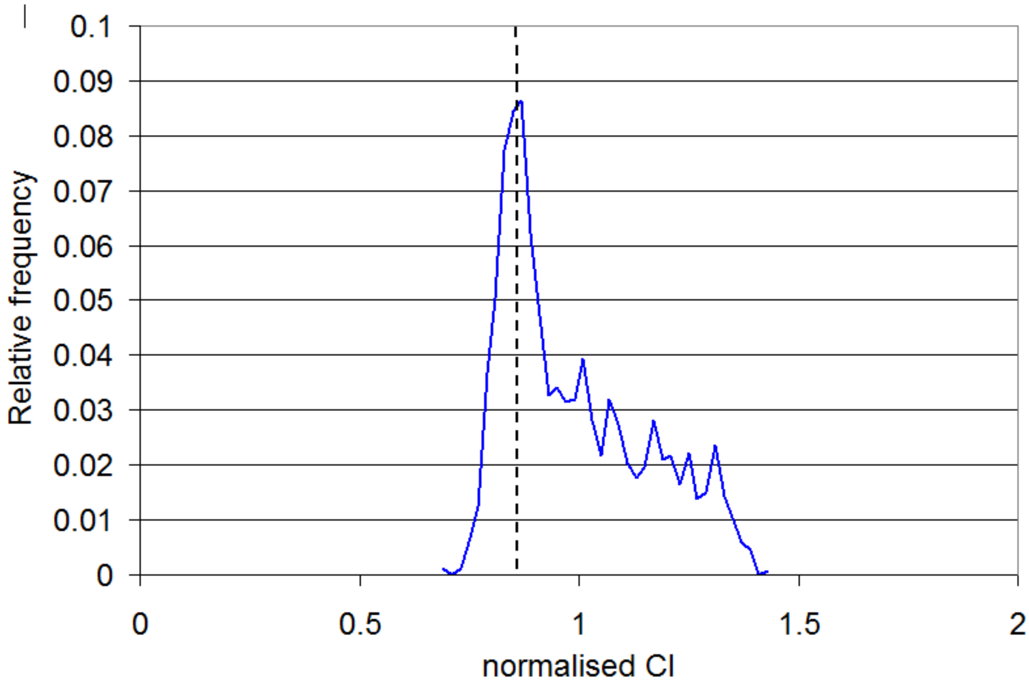

Fig. 2 Frequency distribution of the normalised CI during the CINDI (1) campaign for SZA<60° (for bins of 0.02). The CI at the peak is 0.82.

3 Calibration of the O4 dSCDs

For the analysis of the O_4 dSCD the standard settings provided in Wagner et al. (2019) are used (and recommended). They are summarised in table 1.

Table 1 Standard settings for the DOAS analysis of O_4.

Parameter	**Value, remark, references**
Spectral range	352 – 387 nm
Degree of DOAS polynomial	5
Degree of intensity offset polynomial	2
Ring spectrum 1*	Normal Ring spectrum calculated from DOASIS
Ring spectrum 2*	Ring spectrum 1 multiplied by λ^{-4}
O_3 cross section	223 K, Bogumil et al. (2003)
NO_2 cross section	294 K, Vandaele et al. (1997)**
BrO cross section	223 K, Fleischmann et al. (2004)
O_4 cross section	293 K, Thalman and Volkamer (2013)

*The O_4 analysis in this study used one Ring spectrum calculated from a high resolution solar spectrum. The effect on the O_4 results is very small.

** The O_4 analysis in this study used a second NO_2 cross section for 220 K. Again, the effect on the O_4 results is very small.

Since the O_4 absorption in the Fraunhofer reference spectrum is not known, the O_4 dSCDs derived from the DOAS analysis have to be calibrated as well. The aim of the calibration process is to determine the O_4 absorption of the Fraunhofer reference spectrum (expressed as O_4 AMF, see below). Again, for that purpose the method presented in Wagner et al. (2016) is applied. The different steps are briefly summarised below:

a) the differential O_4 slant column density (dSCD) derived from the spectral analysis is divided by the O_4 vertical column density (VCD) to obtain the corresponding O_4 dAMF:

$$\text{dAMF} = \text{dSCD} / \text{VCD} \quad (4)$$

The VCD can be calculated from vertical profiles of temperature and pressure. Alternatively, but with slightly higher uncertainty, it can also be calculated from the surface values of temperature and pressure (Wagner et al., 2019; Beirle et al., 2022). For the measurements during the CINDI campaign a O_4 VCD of $1.41 \cdot 10^{43}$ molecules²/cm^5 was calculated (Wagner et al., 2014).

b) From the O_4 dAMFs the simulated O_4 AMFs for an AOD of 0.2 are subtracted:

$$\text{dAMF}_{corr} = \text{dAMF} - \text{AMF}_{ref} \quad (3)$$

The O_4 AMFs for a AOD of 0.2 are described by the following polynomial:

$$\text{dAMF}_{ref} = = -81.975 \cdot S^6 + 197.773 \cdot S^5 - 172.649 \cdot S^4 + 64.482 \cdot S^3 - 7.832 \cdot S^2 + 0.964 \cdot S + 1.265 \quad (4)$$

With S the solar zenith angle in degrees divided by 90.
The corrected O_4 dAMF for the MPIC measurements during CINDI (1) are shown in Fig. 3.

c) For the determination of the O_4 AMF of the reference spectrum, clear sky measurements for SZA between 30° and 50° are chosen. The selection of clear sky measurements is performed using the measured CI after they were calibrated as outlined above. Only measurements with CI above a SZA-dependent threshold value are classified as clear. The threshold value is (Wagner et al., 2016):

$$\text{CI-threshold} = 8.399 \cdot S^6 - 24.253 \cdot S^5 + 29.143 \cdot S^4 - 21.056 \cdot S^3 + 7.673 \cdot S^2 - 0.197 \cdot S + 0.964 \quad (5)$$

With S the solar zenith angle in degrees divided by 90.
Fig. 4 shows the frequency distribution of the corrected O_4 dAMFs (for SZA between 30 and 50°) for either all measurements (blue) or clear sky measurements (red). For both sets of measurements the peak is found for a value of –1.78. The negative value of the peak of the frequency distribution represents the O_4 absorption of the Fraunhofer reference spectrum. It has to be added to the measured O_4 dSCDs to obtain the total O_4 AMF.

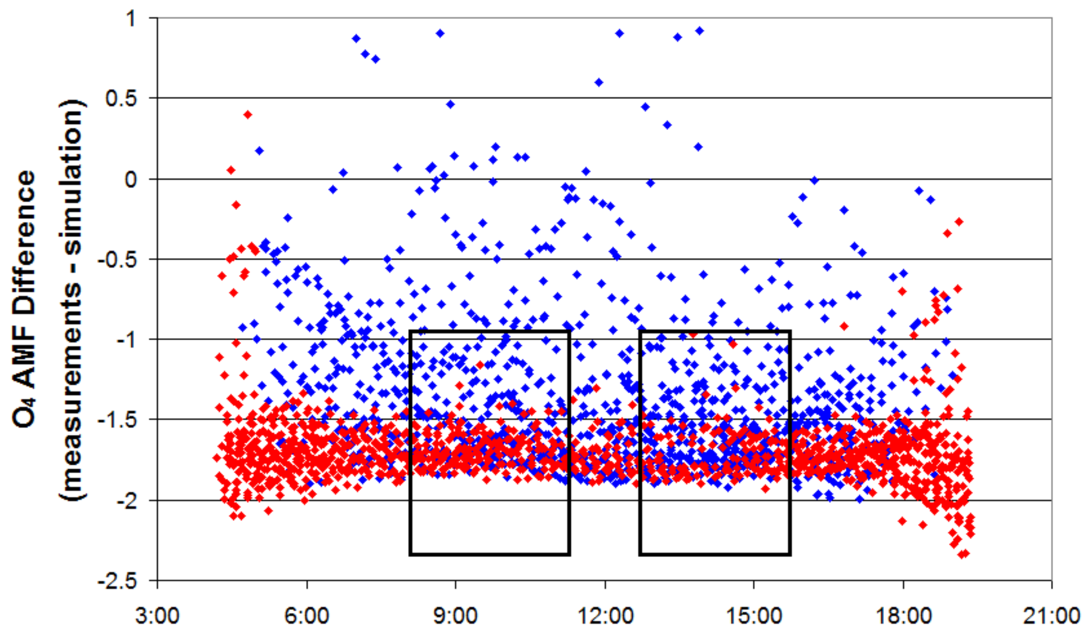

Fig. 3 Normalised O_4 DAMF for all measurements (blue) or measurements under clear sky conditions (red). The normalisation is performed by subtracting the simulated O_4 AMFs for AOD of 0.2. The rectangles indicate the SZA ranges used for the calibration of the O_4 DAMF. The measurements are from the MPIC instrument during the CINDI (1) campaign from 12 June to 15 July 2009.

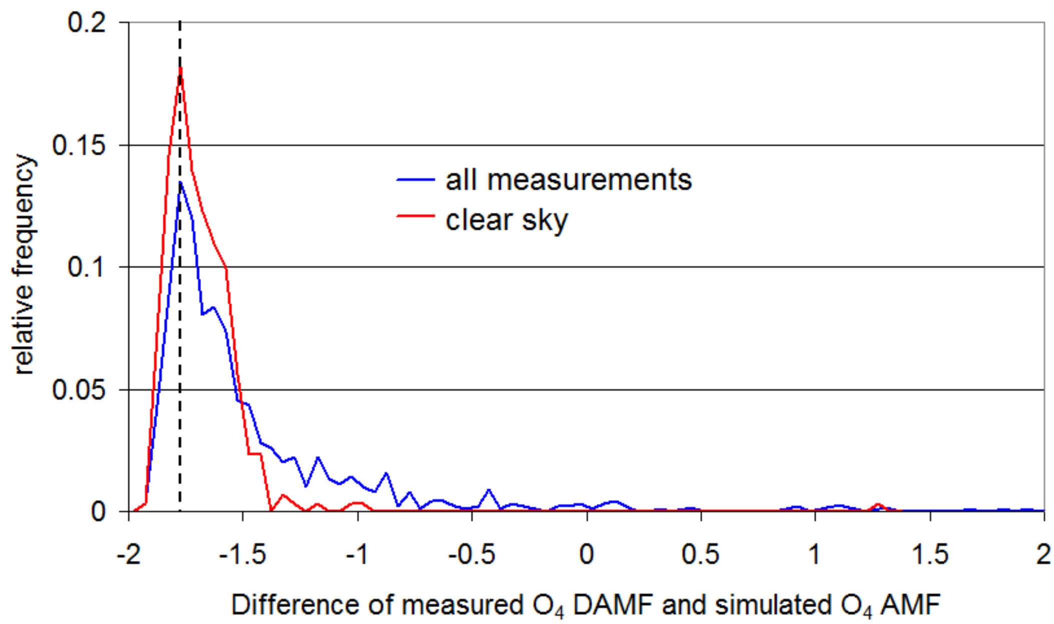

Fig. 4 Frequency distribution (for bins of 0.05) of the normalised O_4 DAMF for SZA between 30° and 50°. The blue and red curves represent observations for all sky conditions (896 measurements) and clear sky conditions (302 measurements), respectively. For both selections the frequency maximum is found for –1.78.

4 The original algorithm from Wagner et al. (2016)

After the preparatory steps described in sections 2 and 4 are carried out, the cloud classification algorithm can be applied. The flowchart of the original algorithm from Wagner et al. (2016) is presented in Fig. 5. The corresponding threshold values are summarised in table 1. The cloud classification algorithm was recommended for measurements with SZA below 75°.
A cloud classification for individual (non-zenith) measurements was so far not implemented (grey symbols). Suitable thresholds for the identification of cloud contaminations based on the temporal smoothness indicator (TSI) should be determined in future studies (for more details, see also RD 2).

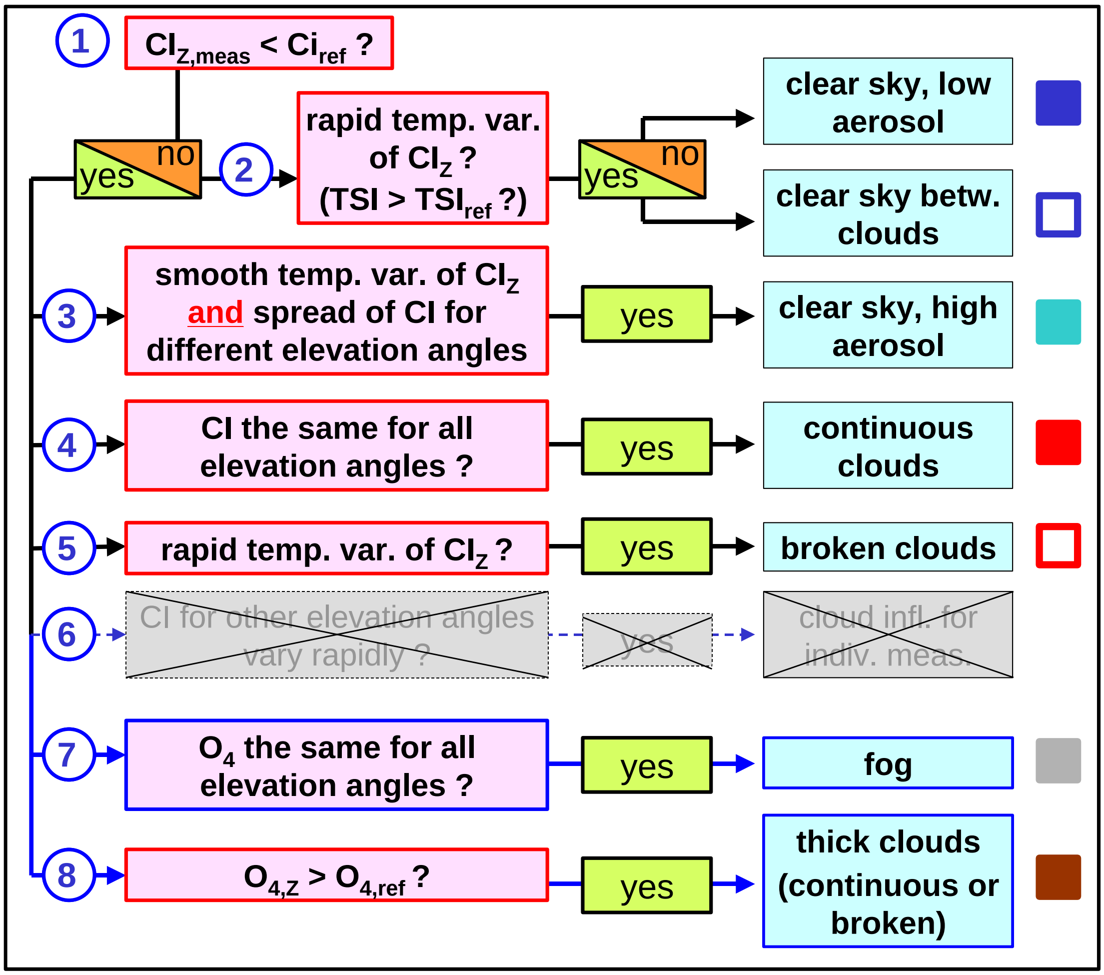

Fig. 5 Flow chart of the cloud classification scheme from Wagner et al. (2016). Decision #6 is so far not implemented. The spread is defined as the difference between the maximum and minimum value of all non-zenith viewing angles of an elevation sequence. For the definition of the TSI, see section 5.1.

Table 2 Thresholds used in the original cloud classification scheme (Wagner et al., 2016).

Quantity	Decision	Threshold (S is the SZA in degrees divided by 90)
CI	1	$= -0.654 \cdot S^6 + 0.367 \cdot S^5 + 2.647 \cdot S^4 - 6.006 \cdot S^3 + 3.576 \cdot S^2 - 0.094 \cdot S + 0.779$
TSI	2,3,5	$= 0.06 \cdot (13.66 \cdot S^6 - 32.298 \cdot S^5 + 28.522 \cdot S^4 - 14.468 \cdot S^3 + 4.644 \cdot S^2 - 0.288 \cdot S + 0.304)$
spread of the CI	3, 4	0.14
Spread of the O_4 dAMF	7	0.37
O_4 AMF	8	$= -81.975 \cdot S^6 + 197.773 \cdot S^5 - 172.649 \cdot S^4 + 64.482 \cdot S^3 - 7.832 \cdot S^2 + 0.964 \cdot S + 2.11$

5 Applied updates

In the updated algorithm, the basic structure of the cloud classification scheme from Wagner et al. (2016) is kept. The updates developed in this project include:

-the definition and threshold of the temporal smoothness indicator (TSI) were changed

-a second threshold was implemented for the TSI for additional sub-categories

-the definition and threshold of the spread of the CI were changed

-a new so-called change flag was introduced

These updates are described in detail in RD-2. A brief summary is given in the next section.

5.1 Changes to the TSI

In the updated version the TSI is calculated in a symmetric way using the zenith measurements before and after a chosen elevation sequence:

$$TSI_{n,new} = \left| CI_{n,start} - CI_{n,end} \right| \quad (6)$$

Here $CI_{n,start}$ and $CI_{n,end}$ indicate the CI of the zenith observation before and after the selected elevation sequence. This new definition was introduced to avoid the asymmetry of the original definition.

Another new feature is the normalization of the TSI by a polynomial as function of the SZA. The normalisation is performed by dividing the TSI obtained by equation 6 by the following polynomial:

$$CI_{AOD=0.2}(S) = 13.144 \cdot S^6 - 39.445 \cdot S^5 + 47.422 \cdot S^4 - 31.335 \cdot S^3 + 10.383 \cdot S^2 - 0.5107 \cdot S + 1.0026 \quad (7)$$

Here S indicates the SZA in degrees divided by 90.
Finally, the existing threshold (see table 2) is multiplied by a factor of 2.5. The outcome of the validation (see RD-2) clearly indicated that the existing threshold was too strict. Increasing the existing threshold by a factor of 2.5 strongly improved the validation results.
Nevertheless, in order to identify measurements with particular stable sky conditions, the old threshold is still applied in addition to the new threshold. Measurements, which fulfil also the old threshold are now assigned to the new sub-categories 'constantly clear' and 'constantly cloudy' (see Fig. 6).

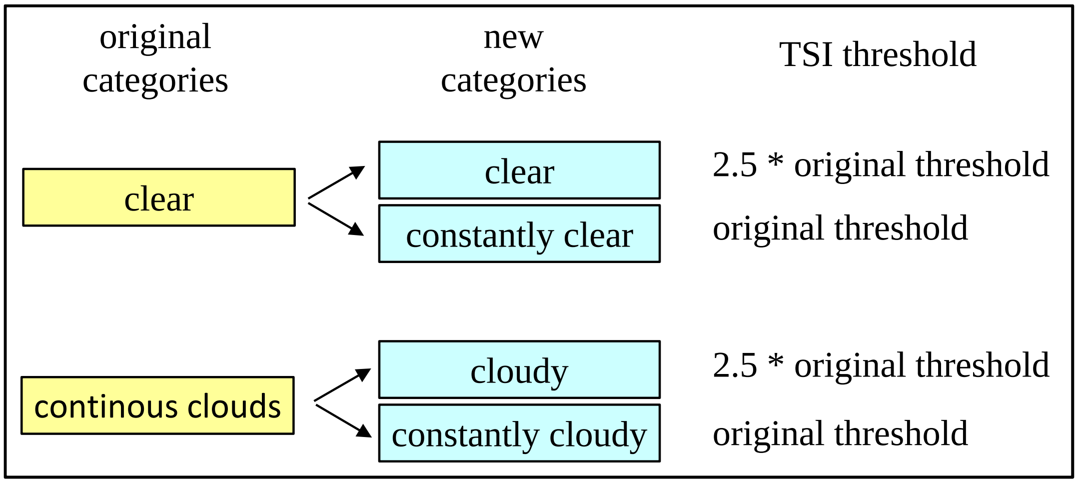

Fig. 6 Introduction of new categories based on different threshold values for the temporal smoothness indicator. The categories 'constantly clear' and 'constantly cloudy' are new sub categories of 'clear' and 'cloudy'.

5.2 Changes to the spread of the CI

To reduce the SZA-dependence of the spread (see RD-2) a normalisation is also applied to the spread of the CI in the updated version. The original spread (difference between the maximum and mimimum CI of the non-zenith directions of an elevation sequence) is divided by the following polynomial (for more details, see RD-2):

$$\text{spread}_{AOD\ 0.1,\ 0.2} = -0.38 \cdot S^6 + 3.4836 \cdot S^5 - 9.186 \cdot S^4 + 6.2471 \cdot S^3 + 0.0513 \cdot S^2 - 0.318 \cdot S + 0.3716 \quad (8)$$

With S being the SZA in degrees divided by 90.

It should be noted that, as discussed in RD-2, for SZA < 40°, the spread is not a useful quantity to discriminate between the categories of ‚continuous clouds' and ‚clear sky conditions with high aerosol load'.
Because of the normalisation of the spread of the CI, a new threshold was determined. By comparison with AOD observations from AERONET a threshold value of 0.5 for the normalised spread was found to be a good choice to discriminate cases of clear sky conditions with an AOD either below or above about 0.5 (at 360 nm).

5.3 Introduction of the change flag

The so-called change flag was introduced as a new flag. It indicates elevation sequences, which have different classification results as their preceding and following sequences. Discarding these sequences largely improves the validation results (from about 61% to 89% correct assignments, see RD-2). However, at the same time, the number of cases is reduced by more than a factor of 3. Fortunately, the application of the change flag mainly removes measurements under broken cloud conditions which are the most unfavorable conditions for MAX-DOAS profile inversions.

5.4 Updated algorithm

Fig. 7 and table 3 provide an overview about the updated algorithm. The basic structure of the algorithm remains unchanged, but two new sub-categories are added. The changes of the definitions and threshold values of the TSI and spread of the CI are marked in table 3.In addition to these changes, also the change flag as defined in section 4.3 is added to the updated cloud classification algorithm.

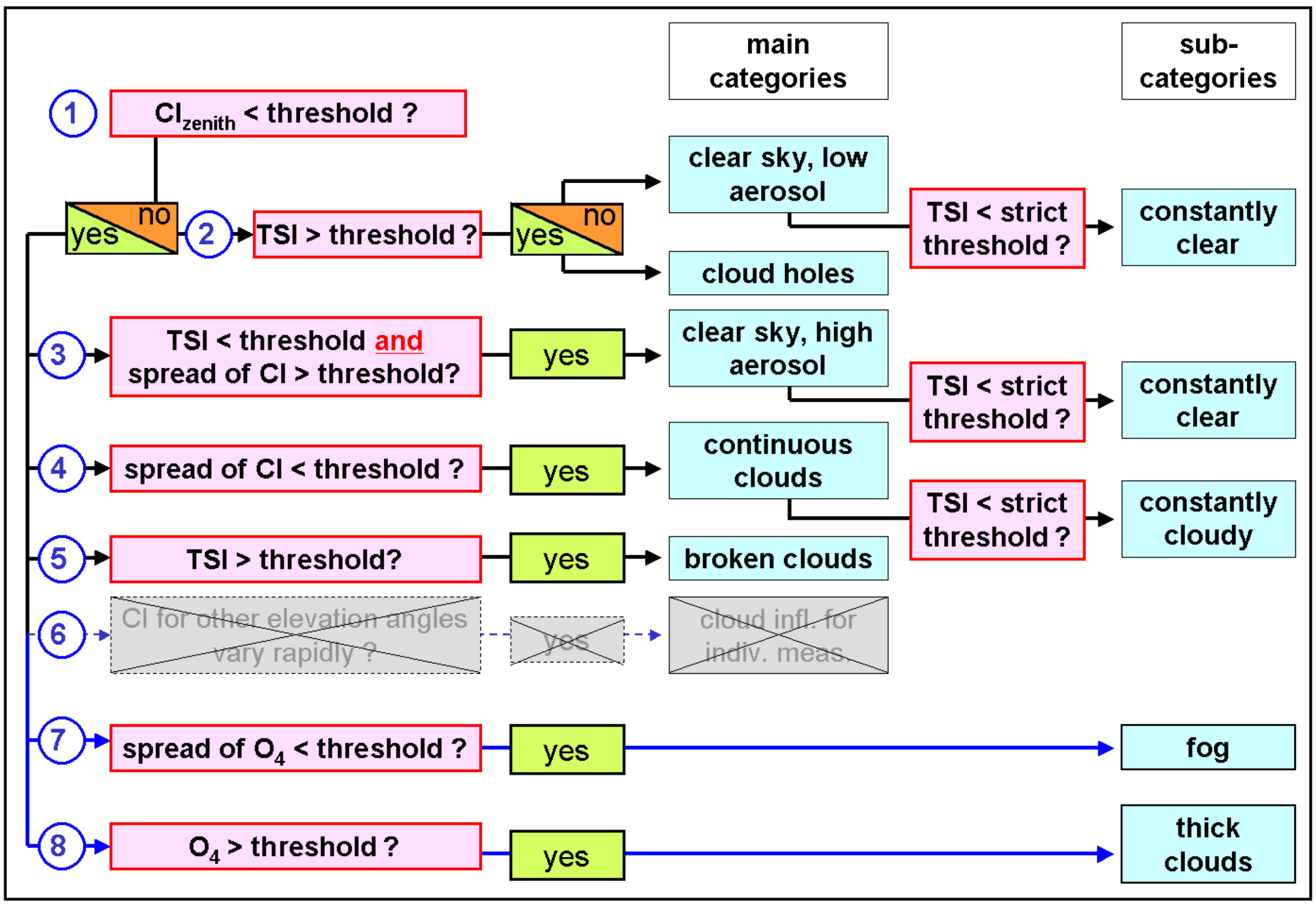

Fig. 7 Flow chart of the updated cloud classification scheme. The basic structure is the same as in Fig. 5, but two new sub-categories ('constantly clear' and 'constantly cloudy') are introduced.

Table 3 Thresholds used in the updated cloud classification scheme. Changes compared to the old version are marked in red.

Quantity	Normalisation	Threshold (S is the SZA in degrees **divided by 90**)
CI	no	$= -0.654 \cdot S^6 + 0.367 \cdot S^5 + 2.647 \cdot S^4 - 6.006 \cdot S^3 + 3.576 \cdot S^2 - 0.094 \cdot S + 0.779$
TSI♦	$= 13.144 \cdot S^6 - 39.445 \cdot S^5 + 47.422 \cdot S^4 - 31.335 \cdot S^3 + 10.383 \cdot S^2 - 0.5107 \cdot S + 1.0026$	$= 2.5 \cdot 0.06 \cdot (13.66 \cdot S^6 - 32.298 \cdot S^5 + 28.522 \cdot S^4 - 14.468 \cdot S^3 + 4.644 \cdot S^2 - 0.288 \cdot S + 0.304)$ ♦♦
spread of the CI♦♦♦	$= -0.38 \cdot S^6 + 3.4836 \cdot S^5 - 9.186 \cdot S^4 + 6.2471 \cdot S^3 + 0.0513 \cdot S^2 - 0.318 \cdot S + 0.3716$	0.5
Spread of the O_4 dAMF	No	0.37
O_4 AMF	no	$= -81.975 \cdot S^6 + 197.773 \cdot S^5 - 172.649 \cdot S^4 + 64.482 \cdot S^3 - 7.832 \cdot S^2 + 0.964 \cdot S + 2.11$

♦*TSI is now calculated according to equation 6*

♦♦for the new sub-categories *'constantly clear' and 'constantly cloudy' the ,old threshold' as in table 1 is used.*

♦♦♦*the spread is only useful for SZA > 40°*

6 Validation results

The original and updated cloud classification algorithms were validated using selected MAX-DOAS measurements at 5 locations in Europe (table 4). The validation was performed by visual classification of the sky conditions from simultaneously recorded sky images (for details see RD-2). For the original version of the cloud classification algorithm, only about 62% of all classification results agreed with the results from the visual classification (Fig. 8). After applying the changes described in section 5, the success rate increased to more than 90%. Here it should be noted that the largest improvement was due to the application of the change flag. The application of this flag strongly reduced the total number of measurements, especially measurements under broken cloud conditions. But such measurements are in general not well suited for MAX-DOAS profile inversions.

Table 4 Selected days used for the validation study at the different locations

Location	**Mostly clear day**	**Mostly cloudy day**	**Days with mostly broken clouds**
Bremen	19.07.2022	22.07.2022	23.07.2022, 28.08.2022
Cabauw	11.08.2022	09.06.2022	16.08.2022 17.08.2022
Uccle	19.07.2022	21.07.2022	03.07.2022 15.07.2022
Mainz	22.09.2022	29.09.2022	19.09.2022 23.09.2022
Thessaloniki	07.07.2022	18.06.2022	20.06.2022 26.06.2022

original cloud classification algorithm updated cloud classification algorithm

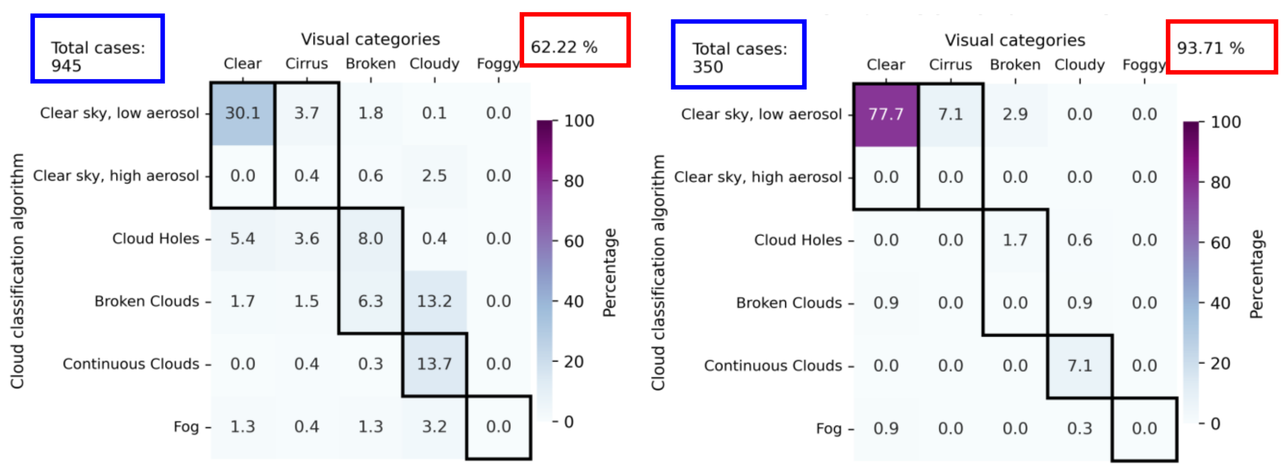

Fig. 8 Validation results for the original (left) and updated (right) cloud classification algorithm. The columns represent the visual classifications, the lines the results of the classification algorithm. The boxes surrounded by the solid lines represent successful assignments; the boxes outside the solid lines represent wrong assignments.

7 Summary

Motivated by the validation results described in RD-02, several changes were applied to the original cloud classification algorithm from Wagner et al. (2016). These changes largely improved the agreement (from about 62% to 93%) with the visual classification results of the sky conditions at 5 stations in Europe.

The basic structure (Fig. 5) of the algorithm remained unchanged, but two new sub-categories *'constantly clear' and 'constantly were introduced (Figs. 6 and 8).*

In addition, the definitions and threshold values for the TSI and spread of the CI were updated as described in section 5.

8 References

[RD-1] FRM4DOAS-2.0 deliverable D-2.1 'Technical Note on existing cloud classification algorithms'

[RD-2] FRM4DOAS-2.0 deliverable D-2.3 'Validation Report (VR) on the cloud classification algorithm'

Beirle, S., Borger, C., Dörner, S., Kumar, V., and Wagner, T.: Calculating the vertical column density of O4 during daytime from surface values of pressure, temperature, and relative humidity, Atmos. Meas. Tech., 15, 987–1006, https://doi.org/10.5194/amt-15-987-2022, 2022.

Bogumil, K., J. Orphal, T. Homann, S. Voigt, P.Spietz, O.C. Fleischmann, A. Vogel, M. Hartmann, H. Bovensmann, J. Frerik and J.P. Burrows, Measurements of Molecular Absorption Spectra with the SCIAMACHY Pre-Flight Model: Instrument Characterization and Reference Data for Atmospheric Remote-Sensing in the 230-2380 nm Region, J. Photochem. Photobiol. A., 157, 167-184, 2003.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. Photoch. Photobio. A, 168, 117–132, 2004.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O2–O2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15, 15371, doi:10.1039/c3cp50968k, 2013.

Vandaele, A. C., C. Hermans, P. C. Simon, M. Carleer, R. Colin, S. Fally, M.-F. Mérienne, A. Jenouvrier, and B. Coquart, Measurements of the NO2 Absorption Cross-section from 42000 cm-1 to 10000 cm-1 (238-1000 nm) at 220 K and 294 K, J. Quant. Spectrosc. Radiat. Transfer, 59, 171-184, 1997.

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., Puķīte, 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? An assessment of uncertainties of measurements and radiative transfer simulations for 2 selected days during the MAD-CAT campaign, Atmos. Meas. Tech., 12, 2745–2817, https://doi.org/10.5194/amt-12-2745-2019, 2019.