WPs-2250-2251: DOAS-BO: "Towards a new FRM4DOAS-compliant site"

D-2 Report on the description of the MAX-DOAS profiles retrieval code and its validation

Document reference: FRM4DOASBO_Ph2_D2

Document Issue: 1.0

Document Issue date: December 2, 2022

Document authors and affiliations: E. Castelli[1], P. Pettinari[1,3], E. Papandrea[1], M.Premuda[1], L. Di Liberto[1], M. Valeri[2]

[1] CNR-ISAC, [2] SERCO S.p.a., [3] UNIBO

Contents

1 List of Acronyms

AKM	Averaging Kernel Matrix
ALC	Automatic Lidar Ceilometer
AMF	Air Mass Factor
AOT	Aerosol Optical Thickness
Arpae	Agenzia Regionale per la Prevenzione, L'Ambiente, l'Energia
CINDI	Cabauw Intercomparison of Nitrogen Dioxide Measuring Instruments
CNR	Consiglio Nazionale delle Ricerche
DOAS	Differential Optical Absorption Spectroscopy
DEAP	Doas optimal Estimation Atmospheric Profile retrieval algorithm
DOF	Degree of freedom
FOV	Field Of View
FM	Forward Model
FRM4DOAS	Fiducial Reference Measurements for DOAS
ISAC	Istituto di Scienze dell'Atmosfera e del Clima
LOS	Line of Sight
MAPA	Mainz Profile Algorithm
MAX-DOAS	Multi AXis - DOAS
MMF	Mexican MAX-DOAS Fit
OE	Optimal Esimation
OMI	Ozone Monitoring Instrument
PGN	Pandora Global Network
QA	Quality Assurance
QA4EO	Quality Assurance For Earth Observation
RTM	Radiative Transfer Model
S-5P	Sentinel-5 Precursor
SCD	Slant Column Densities

STD Standard Deviation

SWIR Short Wave Infra-Red

SZA Solar Zenith Angle

SAA Solar Azimuth Angle

RAA Relative Azimuth Angle

TROPOMI TROPOspheric Monitoring Instrument

UV UltraViolet

VIS Visible

VCD Vertical Column Density

VCM Variance-Covariance matrix

XS Cross Section

2 Introduction

This document is the report of the activities performed in the frame of WPs 2250-1.2 and 1.3 of the IDEAS-QA4EO Phase II WPs-2250-2251 DOAS-BO: "Towards a new FRM4DOAS-compliant site'. The WP 2250-1.2 and 1.3 are centered on the development and validation of NO_2 and aerosol CNR-ISAC retrieval code for MAX-DOAS measurements.

The starting point of this work is the code developed and used during Phase I, [R-1]. This code assumed the aerosol profile as fixed and just retrieves NO_2 through the use of box-AMF. In this new phase, we have implemented the retrieval of the aerosol profile in the retrieval code.

The algorithm validation has been performed using the package of synthetic SCDs used in the frame of the FRM4DOAS algorithm Round-Robin exercise and freely available from https://frm4doas.aeronomie.be/index.php/documents. In addition, we applied our code to a couple of examples of real data and compared the retrieved NO_2 profiles to the ones obtained from the same spectra by the FRM4DOAS centralized processing.

3 Algorithm description

The algorithm we developed at CNR-ISAC for the retrieval of profiles from MAX-DOAS measurements is the DEAP (DOAS optimal Estimation Atmospheric Profile retrieval) algorithm.

The DEAP code is an OE algorithm that exploits the SCIATRAN code [R-2], as FM and a two-step approach. First, the aerosol profile is retrieved from O_4 SCDs with an iterative procedure. Then the retrieved aerosol profile is used to calculate the NO_2 or HCHO BOX-AMFs, and the gaseous profiles are obtained from this inversion (no iteration). The DEAP code uses O_4 and NO_2 (or HCHO) SCDs only, no spectral intensity is used as input.

A through description of the different steps of the algorithm is given here below.

The first step is, as said, the retrieval of aerosol profile from O_4 SCDs. Aerosol profiles are calculated through:

$$\begin{aligned} x_{i+1} = x_i + \\ (K^T S_y^{-1} K + S_0^{-1} + g K^T S_y^{-1} K)^{-1} \\ (K^T S_y^{-1}(y - y_i) - S_0^{-1}(x_i - x_0)) \end{aligned} \tag{3.1}$$

Where x_i is the retrieved profile at iteration i, x_{i+1} is the retrieved profile at iteration i+1, x_0 is the initial guess profile, K is the Jacobian matrix calculated from SCIATRAN, S_0 is the a-priori variance-covariance matrix (VCM), g is the lambda of Marquardt damping factor, S_y is the SCDs VCM matrix, y_i is the vector containing the SCIATRAN O_4 SCDs at different elevation angles, y is the vector of measured O_4 SCDs at different elevation angles.

The iterative procedure stops when the convergence is reached e.g. the difference between the chi-square calculated at two consecutive iterations is below a certain threshold. The chi-square is defined as :

$$\chi^2 = (y - y_i)^T S_y^{-1}(y - y_i) \tag{3.2}$$

and its minimization (the minimization of the difference between simulation and observations weighted by the noise) is the aim of the iterative method.

Once the aerosol profile is obtained, the gaseous profile can finally be retrieved. Although both these steps need to use an RTM (SCIATRAN), the aerosol part requires iteratively calculating the derivative of O_4 SCD with respect to aerosol extinction. By contrast, the gaseous part requires only one calculation of BOX-AMF (a linear problem with no iteration needed), which is the fastest part of the code. Gas profiles are calculated via:

$$x_i = x_0 + S_0 K^T (K S_0 + S_0 K^T + g S_y)^{-1} (y - K x_0) \tag{3.3}$$

Where x_i is the retrieved profile, x_0 is the initial guess profile used for BOX-AMF calculation, K is the BOX-AMF matrix, S_0 is the a-priori variance-covariance matrix (VCM) matrix, g is a damping factor, S_y is the SCDs VCM matrix, y is the vector containing the measured NO_2 or HCHO SCDs at different elevation angles minus the corresponding SCD at 90°.

The a-priori matrix is generally made using a percentage of the a-priori profile in the diagonal elements and a correlation length for extra-diagonal terms.

The derivatives of O_4 SCDs with respect to the aerosol profiles that compose the Jacobian matrix are calculated numerically with SCIATRAN.

In order to account for possible large differences between the initial guess and the real scenarios, the a-priori profile used for aerosol retrieval is scaled at each iteration according to the values of the AOD of the retrieved profile.

4 Algorithm validation using FRM4DOAS dataset from algorithm Round-Robin exercise

To validate the DEAP algorithm, we use the dataset and set up of the FRM4DOAS Round Robin algorithm exercise.

In this section, we describe the set up, the different tests and the outcome of this validation.

Following the work reported in [R-3], we compared the retrieved profiles with the reference ones to asses the performances of the developed retrieval code. In addition, we compare the total AOD and VCD and surface values of extinction and concentration.

4.1 Retrieval set-up and used atmospheric conditions

The package of synthetic SCDs, used in the frame of the FRM4DOAS algorithm Round-Robin exercise [R-3], consists on a dataset of O_4 (360 and 477 nm), NO_2 and HCHO dSCDs and their corresponding errors. The dSCDs are simulated at 9 elevation angles for a total of 990 different combinations of sun position (SZA and RAA), trace gas profiles and aerosol profiles. Two versions of dSCDs are provided, one without any noise and one with a 5% noise, in order to account for possible atmospheric variabilities.

The retrieval settings used for the code validation are reported in table 8 of [R-3]. To summarize them we recall the main characteristics here: aerosol profiles are retrieved from O_4 SCDs at 360 and 477 nm, NO_2 profile from 460 nm and HCHO from 343 nm. The a-priori error matrix is composed as follows: 50% of a priori profile for diagonal elements and Gaussian functions with correlation length of 200 m for extra-diagonal terms.

We named "v1" the version of the retrievals performed with the SCDs with no noise and FRM4DOAS exercise retrieval settings and "v1n" the version with SCDs with noise added, consistently with [R1].

The reference atmosphere is the one used to simulate the synthetic SCDs. SCDs of HCHO and NO_2 are modelled using the same set of vertical profiles since they are of the same magnitude. O_4 SCDs are modeled using 11 aerosol profiles and no trace gases. HCHO and NO_2 SCDs are simulated using all the combinations of 10 trace gases and 11 aerosols profiles (110 model atmospheres). For this study, we avoided SCDs relative to aerosol profile number 10 that is an extreme case with cloud located above 4 km. For each atmospheric scenario, the combination of 3 SZA (40°, 60°, 80°) and 3 RAA (0°, 90°, 180°) are used. Temperature and pressure profiles are taken from "FRM4DOAS_Atmospheres.dat", the surface albedo is set to 0.06, the single scattering albedo to 0.92 and the asymmetry parameter to 0.68 in accordance with table 6 of [R-3].

4.2 Starting point: NO_2 profiles retrieval only

We started from the code developed in Phase I, where no aerosol was retrieved and only NO_2 profiles were obtained. Fig. 4.1 shows the results obtained without any data filtering. Here, we report the results of NO_2 profile retrievals for the 10 target gas scenarios. The blue line represents the initial guess profile, the reference one is plotted in green, while the median of the retrieved profiles with STD is in red. In the last panel, the various colors represent the different aerosol profiles as indicated by the legend. SCDs simulated using different SZA and RAA are used for the retrievals. As can be seen, the better results are obtained for Aerosol profile 1 (in cyan). This is not surprising since we are not retrieving any aerosol profile and the aerosol profile 1 is the closest to the a-priori profile. In general, the NO_2 retrievals perform quite well even when the NO_2 reference profile is really different from the a-priori one. However, the effect of a fixed aerosol profile produces biases in the retrieved NO_2. This can be clearly seen from the comparison between reference and retrieved NO_2 at the surface and tropospheric VCD in Fig.4.2. NO_2 at the surface is underestimated in our retrievals and and the same happens to the tropospheric VCDs, even if with a better correlation and a less extent.

4.3 DEAP Aerosol, NO_2 and HCHO retrieval using "v1" dSCDs dataset

The DEAP code was firstly applied to the dSCD dataset named "v1" for validation. Following the work in [R-3] and [R-4], we retrieved aerosol profiles at 360 and 477 nm , HCHO profiles at 343nm and NO_2 at 460nm. The used retrieval grid is a point every 0.2 km from 0 up to 4 km. The same grid was adopted in the SCIATRAN code used for Jacobians and block-AMF calculations.

The AER10 scenario was already excluded from our tests. Still, since all the statistics reported in [R-3] and [R-4] also exclude AER8 and AER9 scenarios, to avoid possible biases, we will also exclude those two scenarios from the following analysis.

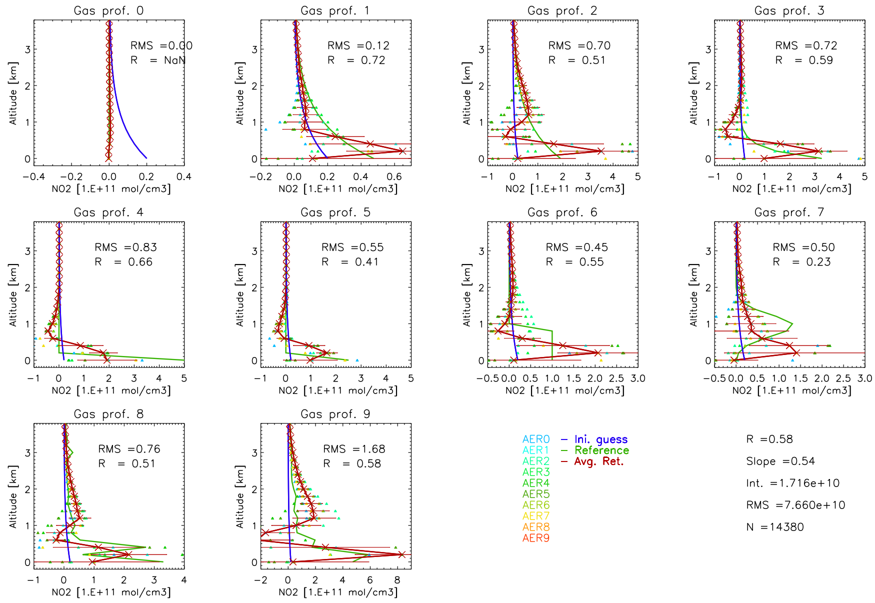

Figure 4.1: Results of NO_2 profiles retrieval for the 10 target gas scenarios. The blue line is the initial guess profile, the green is the reference one, the red is the median of retrieved profiles with STD. The different colors represent the different aerosol profiles, as indicated in the legend.

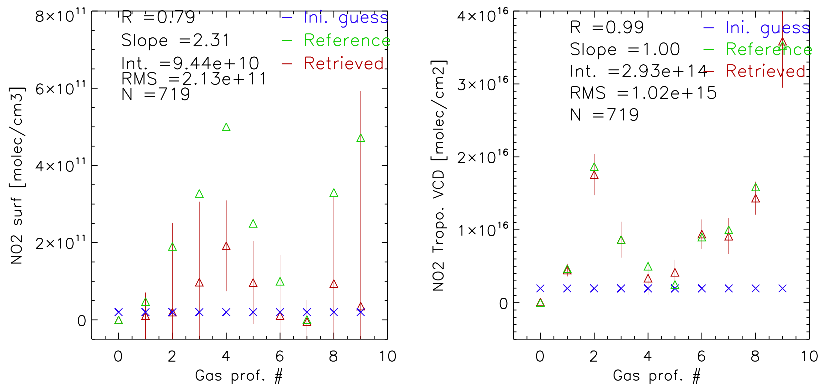

Figure 4.2: Results of NO_2 surface retrieved profiles for the10 trace gas profiles (in red) against initial guess values (in blue) and reference (in green) on the left. Same for NO_2 retrieved Tropospheric VCDs on the right. All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

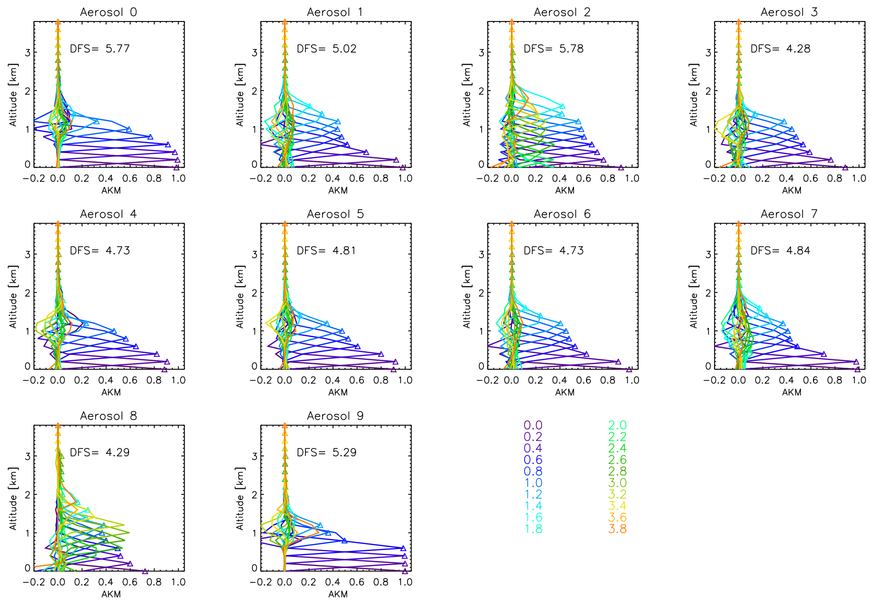

Figure 4.3: Aerosol extinction at 360 nm AKM for the 10 aerosol profiles with indication of DOFs.

4.3.1 Averaging Kernel Matrix

As an example of performances, AKM for aerosol profiles at 360 and 477 nm, HCHO and NO_2 profiles are reported in figures 4.3 to 4.6. Here we plot the median values of AKM. The dependency on SZA and RAA is really low. AKM suggest that no information above 2km is obtained when using a zenith-sky reference spectrum from the same scan as reported in [R-4]. DOFs for aerosols ranges from 4 to 7: UV wavelengths have less DOFs with respect to Visible in agreement with atmospheric opacity, lower DOFs values are obtained for fog (AER(8)) scenario. Trace gases DOFs ranges from 1 to 3. These values and behaviour are in line with what reported in [R-4], even if we found slightly higher values for aerosol DOFs possibly due to a very low Marquardt lambda value used in our retrievals.

Our AKM are similar to the ones reported in [R-4] especially for trace gases. In case of aerosols our AKM are less smoothed with respect to [R-4], probably because of the very low Marquardt lambda value used in our retrievals.

4.3.2 Slant Column Densities

The quality of the convergence of the retrievals can be established comparing the "measured" dSCDs with the one that results from the last simulation of the forward model produced with the retrieved target profiles.

We have the possibility of filtering out results that are produced from modeled dSCDs that exceed the measured ones of a certain threshold. For this analysis the threshold were: $0.5*10^{44}$ on O4 dSCDs (similarly to what reported in [R-4]), $0.4*10^{16}$ on NO_2 and HCHO.

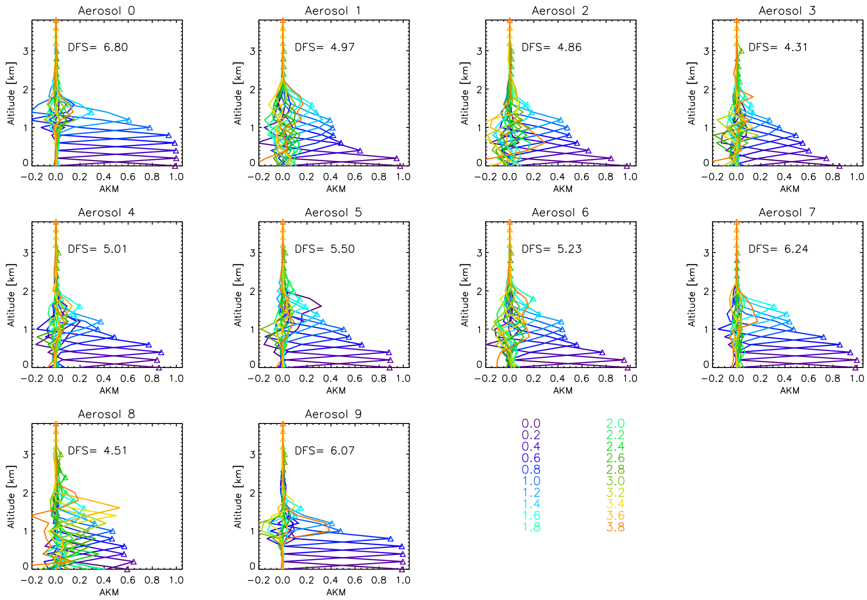

Figure 4.4: Aerosol extinction at 477 nm AKM for the 10 aerosol profiles with indication of DOFs.

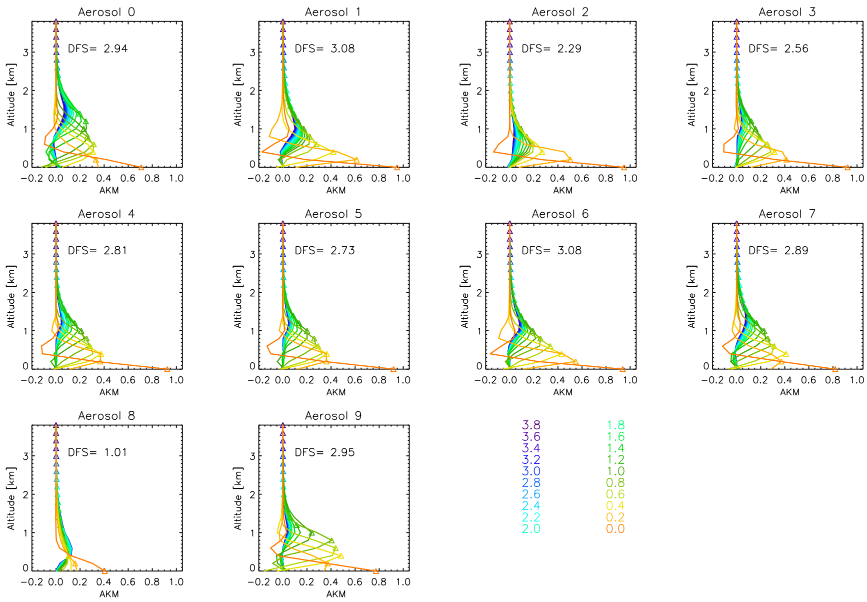

Figure 4.5: NO_2 AKM for the 10 aerosol profiles with indication of DOFs.

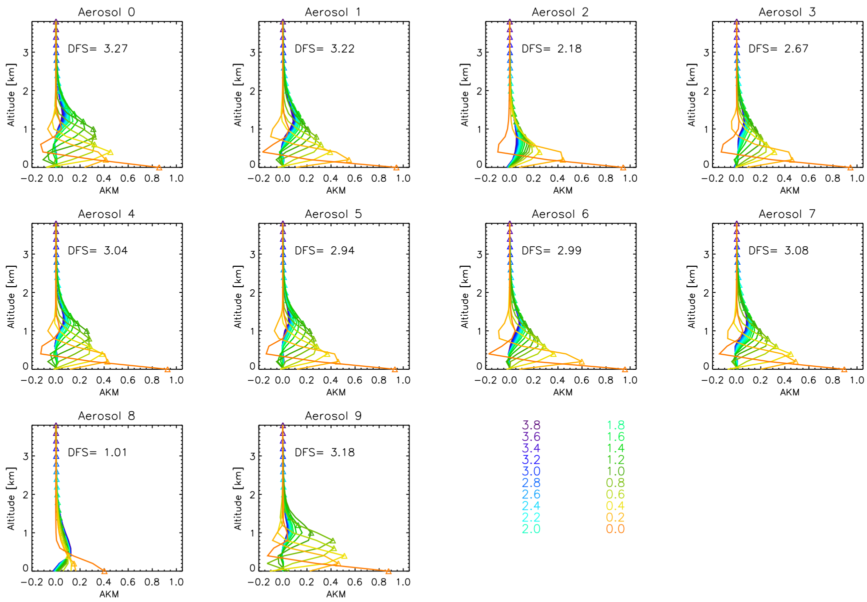

Figure 4.6: HCHO AKM for the 10 aerosol profiles with indication of DOFs.

The results of the comparison between a-posteriori and "measured" dSCDs is reported in Figure 4.7. As can be noticed, comparing the results with the ones in [R-4] figure 10, we can find a similar good agreement for aerosol retrievals (apart from AER0) at both 360 and 477 nm, while a slightly worst, but still not bad, agreement is obtained for gaseous dSCDs.

4.3.3 PROFILES

Results of profiles retrieval are reported in figures 4.9 (aerosol extinction at 360nm), 4.8 (aerosol extinction at 477nm), 4.10 (HCHO), 4.11 (NO_2).

The comparison is performed with true profiles contained into the reference atmosphere. No AKM convolution, that should improve the comparison, has ben performed here in line

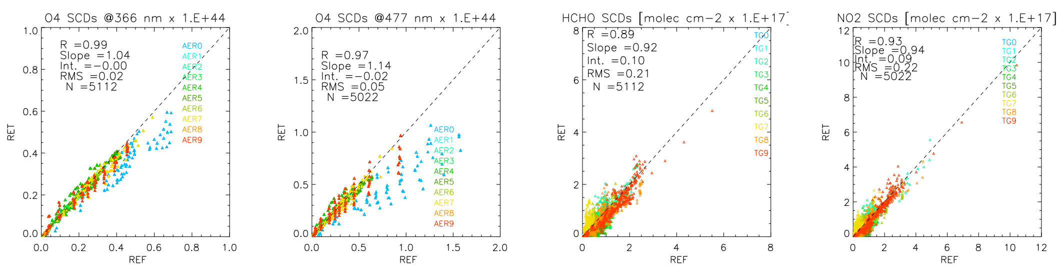

Figure 4.7: SCDs from FRM4DOAS versus the a-posteriori ones for O_4 at 360 nm, O_4 at 477 nm, HCHO at 343nm and NO_2 at 460nm for v1. All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

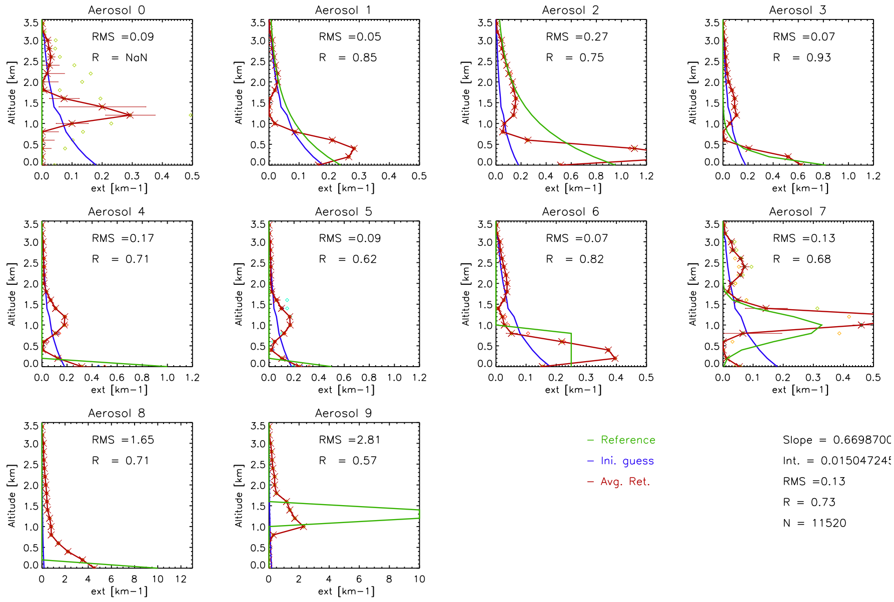

Figure 4.8: Results of aerosol extinction at 360 nm profiles retrieval for the 10 aerosol scenarios and v1. Blue line is the initial guess profile, green the reference one, red the median of retrieved profiles with STD.

with what done in [R-3] and [R-4].

The blue lines are the a-priori profiles used in the retrievals, the green lines are the reference profiles, and the red line is the average of the retrieved profiles for each scenario. Individual profiles are indicated with colored symbols.

The aerosol profile retrievals show good results even if some oscillations are present. Averaging over all possible gaseous, SZA and RAA case for each aerosol scenarios, we get correlations from 0.57 for AER9 to 0.91 for AER3 in cases of 360 nm and from 0.45 for AER9 to 0.98 for AER3 in the visible. The RMS values ranges from 0.04 to about 3 for AER9 scenario, in line with what is reported in figures 11-12 of [R-4].

Similar considerations apply to HCHO profiles: the R value ranges from 0.56 to 0.94 and the RMS from 0.01*10^{11} to 1.83*10^{11} molecules/cm^3. For NO_2, we find R values from 0.52 to 0.98 and RMS from 0.01*10^{11} to 1.58*10^{11} molecules/cm^3.

The shape of the profiles agrees quite well with the reference in all cases apart from high aerosol load. The retrieved profile is able to reproduce the amount of target even in cases when the real value is far away from the initial guess.

In case of trace gas retrievals, the shape of the profiles is much more smooth with respect to the aerosols one in agreement with AKM (e.g. see figure 4.5).

Globally the profiles correlations range from 0.7 to 0.83, the slope from 0.63 to 0.82.

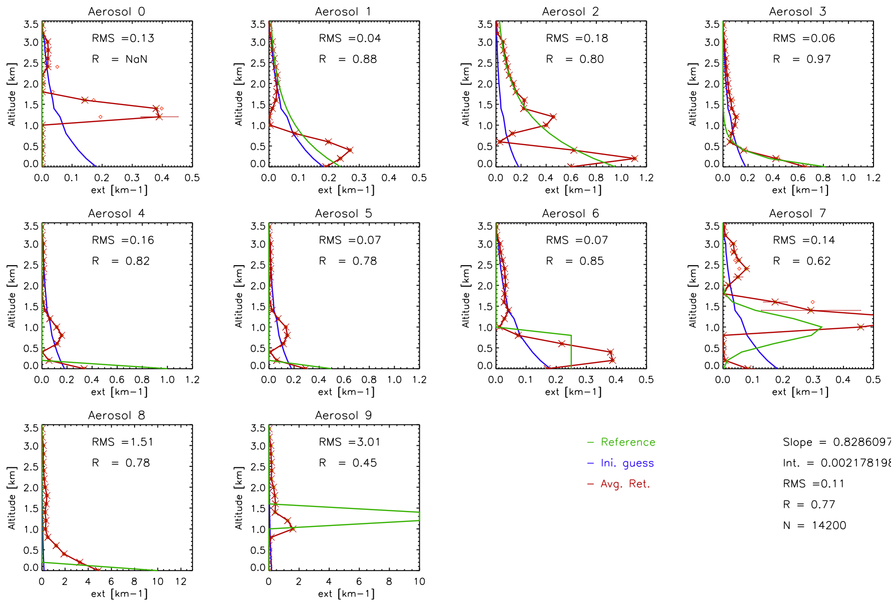

Figure 4.9: Results of aerosol extinction at 477 nm profiles retrieval for the 10 aerosol scenarios and v1. Blue line is the initial guess profile, green the reference one, red the median of retrieved profiles with STD.

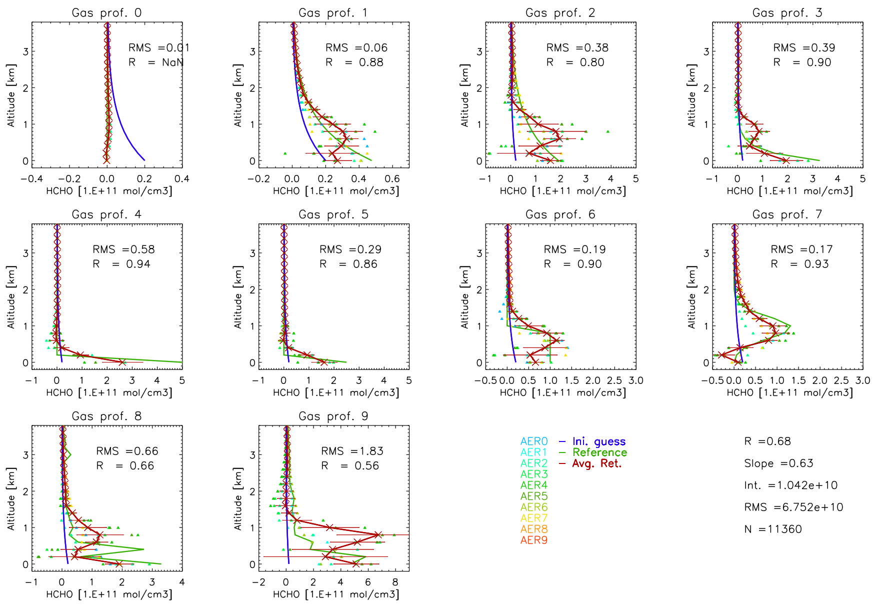

Figure 4.10: Results of HCHO profiles retrieval for the 10 target gas scenarios and v1. Blue line is the initial guess profile, green the reference one , red the median of retrieved profiles with STD. The different color represent the different aerosol profiles as indicated in the legend.

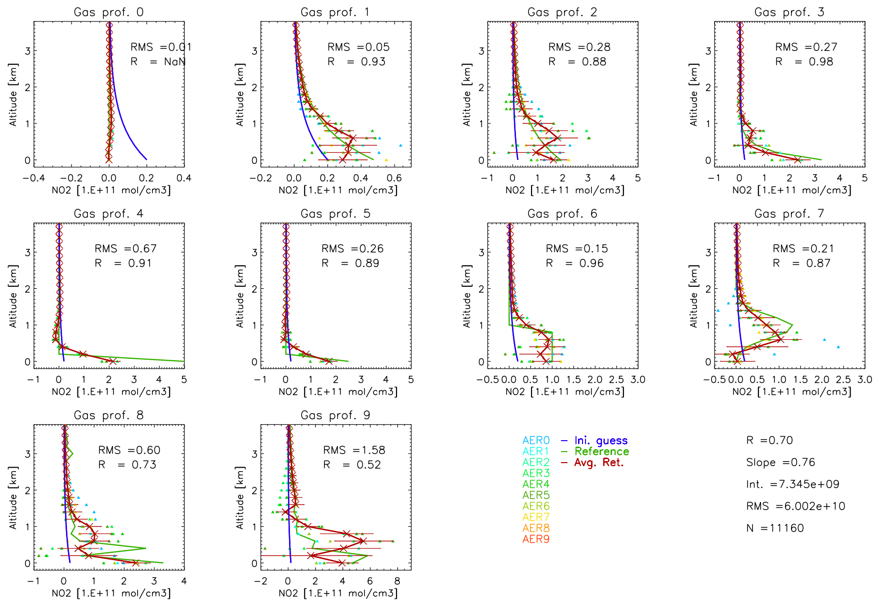

Figure 4.11: Results of NO_2 profiles retrieval for the 10 target gas scenarios and v1. Blue line is the initial guess profile, green the reference one , red the median of retrieved profiles with STD. The different color represent the different aerosol profiles as indicated in the legend.

4.3.4 SURFACE VALUES

Surface values are interesting for comparisons with in-situ ground-based data. We recall here that we performed a rough comparison with reference profiles without considering the AKM. A proper inter-comparison between high resolution and lower altitude resolution data should account for that.

In Figure 4.12, we report the retrieved surface values for aerosol and trace gases for the different scenarios. Blue crosses represent the initial guess value, the green triangles the reference one and in red the retrieved values.

As can be noticed, in some cases the retrieved values are lower than the reference ones. The slope is > 1, ranging from 1.08 to 1.65. The correlation, however, is always high with values between 0.87 and 0.92.

4.3.5 TROPOSPHERIC COLUMNS

The tropospheric columns are, so far, the most critical product of MAX-DOAS retrievals. They are used for satellite validation purposes and, therefore, their accuracy is crucial for this task.

In Figure 4.13, we report the retrieved tropospheric columns for aerosol and trace gases for the different scenarios. Blue crosses represent the initial guess value, the green triangles the reference one and in red the retrieved values.

The retrieved and reference values agrees always very well with really high correlation values

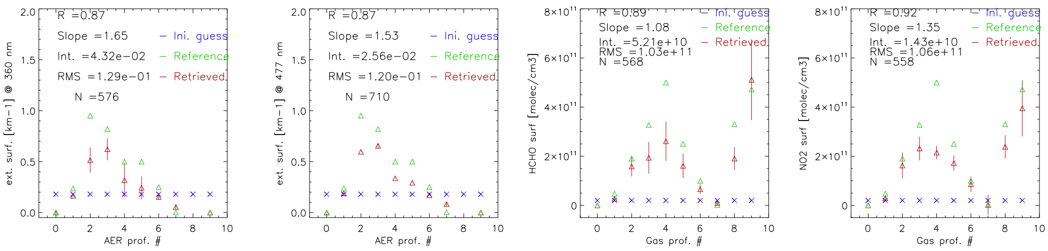

Figure 4.12: Results of aerosol extinction at 360 nm, aerosol extinction at 477nm, HCHO and NO_2 for v1 surface retrieved values (in red) against initial guess values (in blue) and reference ones (in green) All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

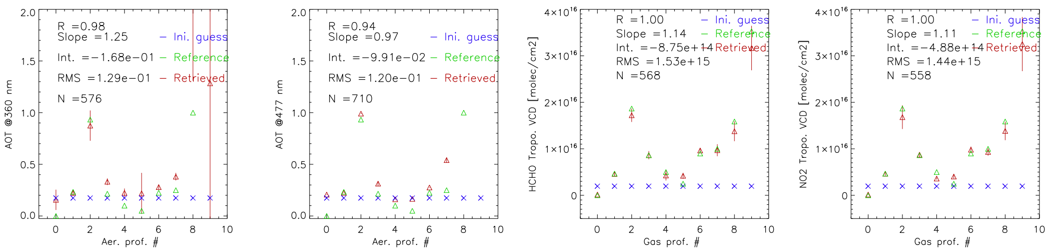

Figure 4.13: Results of aerosol extinction at 360 nm, aerosol extinction at 477nm, HCHO and NO_2 for v1 column retrieved values (in red) against initial guess values (in blue) and reference ones (in green) All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

from 0.94 to 1 and slope from 0.97 to 1.25, the RMS is lower than 10% for HCHO and NO_2.

4.4 DEAP Aerosol, NO_2 and HCHO retrieval using "v1n" dSCDs dataset

We repeat the validation exercise using the ”v1n” dSCDs dataset. This dataset has an additional noise values added to simulated dSDCs in order to mimic possible atmospheric effects.

Since they are really similar to the ones in the ”v1” case we do not show here the AKM.

4.4.1 Slat Column Densities

The results of the comparison between a-posteriori and ”measured” dSCDs for ”v1n” dataset are reported in Figure 4.14. As can be noticed, comparing the results with the ones in [R-4] figure 10, we can find a similar good agreement for aerosol retrievals (apart from AER0) at both 360 and 477 nm, and for gaseous dSCDs. The correlation is really high ranging from 0.94 to 0.98, the RMS is low (0.02*10^{44} molecules/*cm^2) for aerosols and slightly higher for gases (0.2*10^{17} molecules/*cm^2).

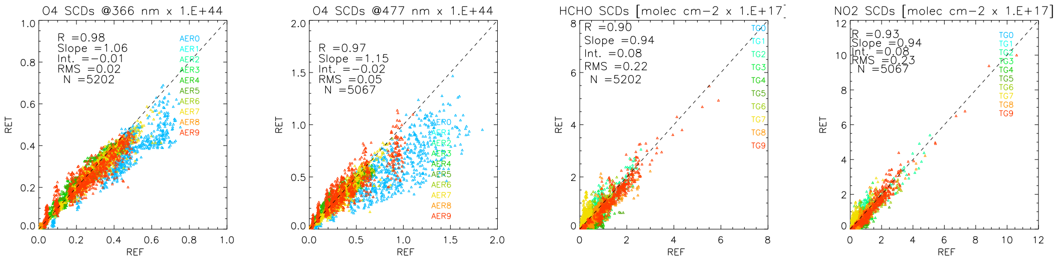

Figure 4.14: SCDs from FRM4DOAS versus the a-posteriori ones for O_4 at 360 nm, O_4 at 477 nm, HCHO at 343nm and NO_2 at 460nm for v1n. All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

Values of correlation, intercept, slope and RMS are in good agreement with what shown in figure 10 of [R-4].

4.4.2 PROFILES

Results of profiles retrieval are reported in figures 4.16 (aerosols at 343nm), 4.15 (aerosol at 477nm), 4.17 (HCHO), 4.18 (NO_2).

The aerosol profile retrievals show good results even if some oscillations are present. Averaging over all possible gaseous, SZA and RAA case for each aerosol scenarios we get correlations from 0.24 for AER9 to 0.95 for AER3 in cases of 360 nm and from 0.40 for AER9 to 0.96 for AER3 in the visible. The RMS values range from 0.02 to about 3 for AER9 scenario, in line with what reported in figures 11-12 of [R-4]. AER9 is one of the most critical scenarioes for almost all the codes involved in the algorithm round-robin exercise [R-4].

Similar considerations apply to HCHO profiles: the R value ranges from 0.4 to 0.97 and the RMS from $0.01*10^{11}$ to $1.87*10^{11}$ molecules/cm^3. For NO_2, we find R values from 0.61 to 0.94 and RMS from $0.01*10^{11}$ to $1.45*10^{11}$ molecules/cm^3.

The shape of the profiles agrees quite well with the reference in all cases apart from high aerosol load. The retrieved profiles are able to reproduce the amount of target even in cases when the real value is far away from the initial guess one.

Globally the profiles correlations range from 0.61 to 0.83, the slope from 0.62 to 0.9.

These results are really similar to the ones obtained with dSCDs dataset ”v1”.

4.4.3 SURFACE VALUES

Retrieved surface values are reported in figure 4.19. Blue crosses represent the initial guess value, the green triangles the reference one and in red the retrieved values.

As can be noticed, in some cases the retrieved values are lower than the reference ones. Actually, the slope is > 1, ranging from 1.27 to 1.5. The correlation, however, is always high with values between 0.90 and 0.97.

Comparing the results with the ones in figure 21 of [R-4], we can notice that the values we find for R, slope, intercept and RMS are in line with their results.

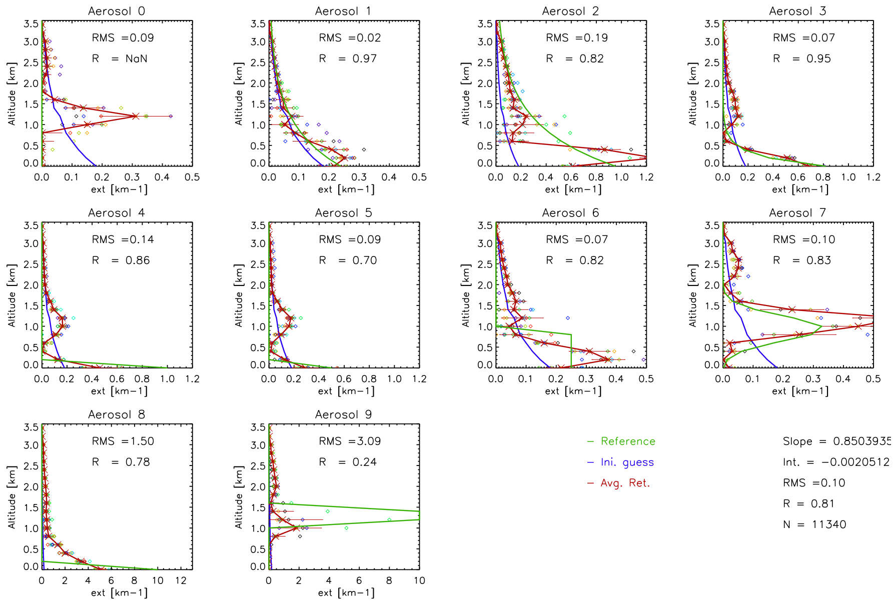

Figure 4.15: Results of aerosol extinction at 360 nm profiles retrieval for the 10 aerosol scenarios and v1n. Blue line is the initial guess profile, green the reference one, red the median of retrieved profiles with STD.

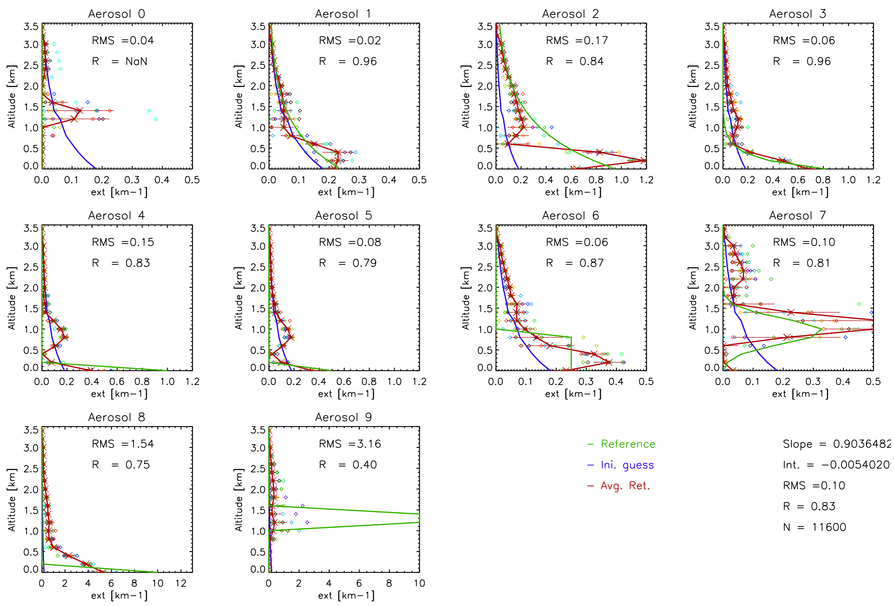

Figure 4.16: Results of aerosol extinction at 477 nm profiles retrieval for the 10 aerosol scenarios and v1n. Blue line is the initial guess profile, green the reference one, red the median of retrieved profiles with STD.

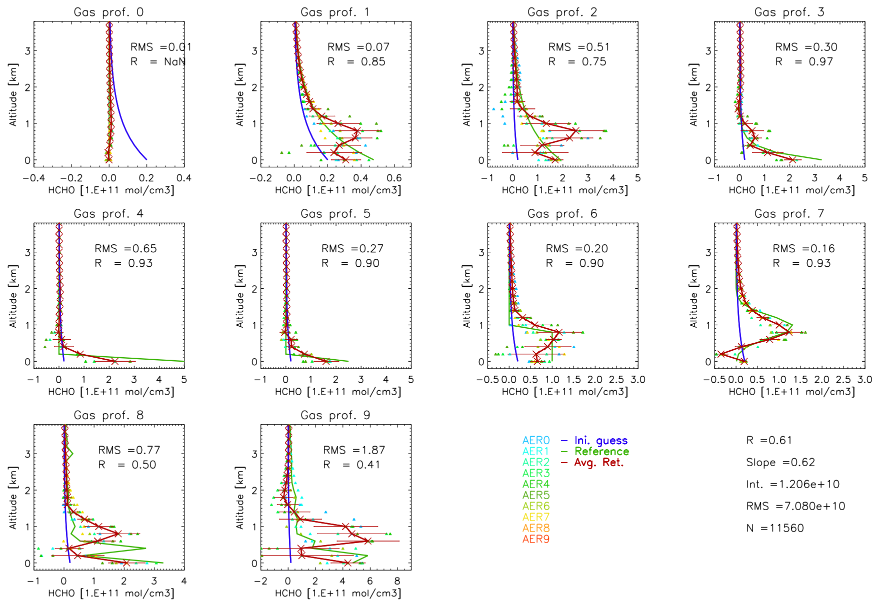

Figure 4.17: Results of HCHO profiles retrieval for the 10 target gas scenarios and v1n. Blue line is the initial guess profile, green the reference one , red the median of retrieved profiles with STD. The different color represent the different aerosol profiles as indicated in the legend.

4.4.4 Tropospheric columns

In Figure 4.20, we report the retrieved tropospheric columns for aerosol and trace gases for the different scenarios.

The retrieved and reference values always agree very well with really high correlation values from 0.97 to 1 and slopes from 1.1 to 1.27.

Comparing the results with the ones in figure 18 of [R-4], we can notice that the values we find for R, intercept and RMS are in line with their results, while the slope we get is slightly higher and above 1.

4.5 Comparison with MMF and MAPA results

Since we want to have a quantitative assessment of the performances of the DEAP code, we try to consider the values of R, slope, intercept, RMS for each of the products listed above from [R-3] and [R-4]. Differently from [R-3], where the statistical values are reported in tables, it is not easy to extract the same statistical values from [R-4] since they are reported in plots. We do not know if the values in [R-4] and [R-3] are fully consistent, however, we can roughly estimate their consistency and use the values in tables in [R-3] for our comparisons.

MAPA (named MPIC-Param in [R-3], [R-5]) and MMF [R-5] are the two algorithms used in the frame of the FRM4DOAS centralized processing. We compare DEAP results against the ones from these two codes. We selected the " v1n" dataset for this task, but similar considerations

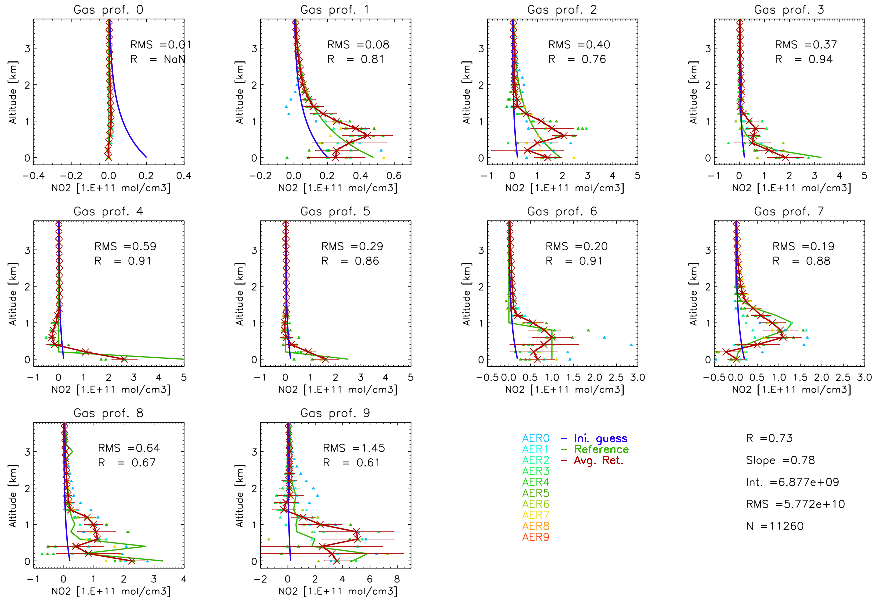

Figure 4.18: Results of NO_2 profiles retrieval for the 10 target gas scenarios and v1n. Blue line is the initial guess profile, green the reference one , red the median of retrieved profiles with STD. The different color represent the different aerosol profiles as indicated in the legend.

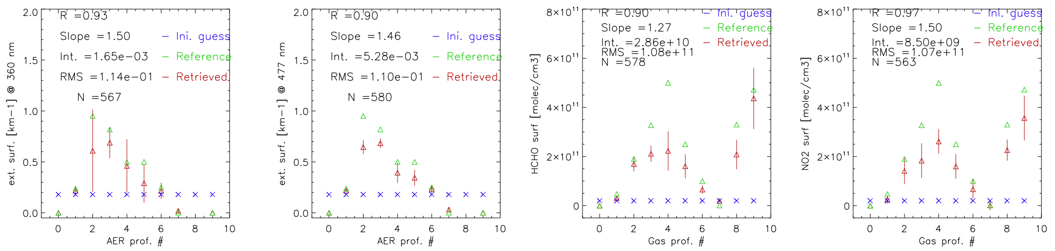

Figure 4.19: Results of aerosol extinction at 360 nm, aerosol extinction at 477nm, HCHO and NO_2 for v1n surface retrieved values (in red) against initial guess values (in blue) and reference ones (in green) All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

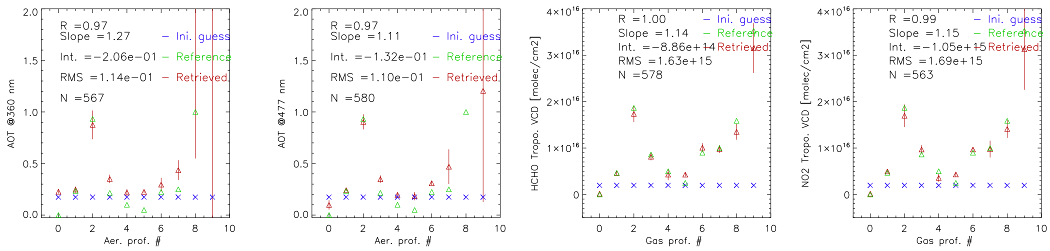

Figure 4.20: Results of aerosol extinction at 360 nm, aerosol extinction at 477nm, HCHO and NO_2 for v1n column retrieved values (in red) against initial guess values (in blue) and reference ones (in green) All the statistical parameters are calculated excluding AER8 and AER9 scenarios.

Table 4.1: Comparison of slope, intercept, correlation, RMS and number of points for MAPA, MMF and DEAP codes. Parameters are calculated for a-posteriori dSCDs of O_4 at 360 nm and 477 nm, HCHO and in NO_2.

	dSCDs - "v1n"											
	O4 @360 nm			O4 @477 nm			HCHO			NO2		
	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP
Slope	-	0.98	1.06	-	0.99	1.15	-	0.98	0.94	-	0.99	0.94
Intercept	-	$0.04*10^{43}$	$-0.1*10^{43}$	-	$0.04*10^{43}$	$-0.2*10^{43}$	-	$0.02*10^{17}$	$0.08*10^{17}$	-	$0.01*10^{17}$	$0.08*10^{17}$
R	-	1.0	0.98	-	1.0	0.97	-	1.0	0.90	-	1.0	0.93
RMS	-	$0.12*10^{43}$	$0.2*10^{43}$	-	$0.13*10^{43}$	$0.5*10^{43}$	-	$0.05*10^{17}$	$0.22*10^{17}$	-	$0.03*10^{17}$	$0.23*10^{17}$
N	-	6209	5202	-	5516	5067	-	6479	5202	-	5598	5067

are still valid when using the "v1" dataset.

The results of comparisons for a-posteriori dSCDs are reported in table 4.1. For MAPA, no results are reported in [R-3], but they are in figure 10 of [R-4]. At first sight, they seem quite similar to the MMF ones, even if the scale is quite large.

In comparison to MMF we have slightly larger slopes for O_4 dSCDs while a better agreement is obtained for NO_2 and HCHO dSCDs. On the other hand, correlations agree better for O_4 dSCDs than for NO_2 and HCHO dSCDs.

The results of comparisons for profiles are reported in table 4.2. In this case both results from MAPA and MMF are available. DEAP results show lower slope values with respect to both MAPA and MMF in all cases. The correlation of DEAP results is better than MAPA for aerosol extinction and similar to MMF for the same cases. It is slightly worse in the case of aerosol retrievals. This is also reflected in the RMSE values: better or comparable RMSE values with respect to MMF and MAPA are found for aerosol, while slightly higher values are found for gaseous profiles.

Surface values comparisons in table 4.3 show definitely higher values of the slope for DEAP (underestimation of values) in all cases while very good correlation, in some cases even better than MAPA and MMF is obtained. Regarding the RMS values, DEAP performs similarly or better than MAPA and MMF.

An important benchmark, since satellite columns validation is one of the primary objective of MAX-DOAS measurements especially in the frame of this project, is the comparison of DEAP

Table 4.2: Comparison of slope, intercept, correlation, RMS and number of points for MAPA, MMF and DEAP codes. Parameters are calculated for retrieved profiles of aerosols at 360 nm and 477 nm, HCHO and in NO_2.

	profiles - "v1n"											
	aer.@360 nm			aer.@477 nm			HCHO			NO_2		
	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP
Slope	1.04	0.90	0.85	1.084	0.91	0.90	1.0	0.90	0.62	1.026	0.871	0.78
Intercept	0.021	0.006	-0.002	0.02	0.005	-0.005	$-0.004*10^{11}$	$0.05*10^{11}$	$0.12*10^{11}$	$-0.006*10^{11}$	$0.04*10^{11}$	$0.07*10^{11}$
R	0.64	0.76	0.81	0.75	0.89	0.83	0.86	0.90	0.61	0.84	0.84	0.73
RMS	0.19	0.12	0.1	0.14	0.07	0.1	$0.38*10^{11}$	$0.28*10^{11}$	$0.7*10^{11}$	$0.42*10^{11}$	$0.32*10^{11}$	$0.6*10^{11}$
N	14400	13800	11340	14400	12260	11600	14400	14400	11560	14400	12440	11260

Table 4.3: Comparison of slope, intercept, correlation, RMS and number of points for MAPA, MMF and DEAP codes. Parameters are calculated for values at surface aerosols at 360 nm and 477 nm, HCHO and in NO_2.

	surface - "v1n"											
	aer.@360 nm			aer.@477 nm			HCHO			NO_2		
	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP
Slope	1.16	0.98	1.50	1.03	0.97	1.46	0.98	0.89	1.27	0.97	0.90	1.50
Intercept	-0.029	0.001	-0.002	0.009	0.013	-0.005	$0.01*10^{11}$	$0.11*10^{11}$	$0.28*10^{11}$	$0.12*10^{11}$	$0.11*10^{11}$	$0.08*10^{11}$
R	0.86	0.93	0.93	0.93	0.91	0.90	0.85	0.93	0.90	0.77	0.90	0.97
RMS	0.21	0.12	0.11	0.13	0.13	0.11	$0.86*10^{11}$	$0.55*10^{11}$	$1.0*10^{11}$	$1.18*10^{11}$	$0.59*10^{11}$	$1.07*10^{11}$
N	720	690	567	720	613	580	720	720	578	720	622	563

tropospheric columns values with official FRM4DOAS codes (reported in table 4.4). DEAP slope values are in general good, even if often higher than MAPA and MMF. The correlation is always very high and better than MAPA and MMF and the same is true for RMS. However, slightly less valid data are used in DEAP case, due to more stringent quality criteria used.

Generally, the DEAP code performs slightly worse against MAPA and MMF. Still, the results are entirely consistent with these reference codes and considering that it is the first version of the DEAP code, we can assert the validation of the code is successful.

Table 4.4: Comparison of slope, intercept, correlation, RMS and number of points for MAPA, MMF and DEAP codes. Parameters are calculated for tropospheric columns of aerosols at 360 nm and 477 nm, HCHO and in NO_2.

	Tropospheric Columns - "v1n"											
	aer.@360 nm			aer.@477 nm			HCHO			NO_2		
	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP	MAPA	MMF	DEAP
Slope	0.90	0.65	1.27	1.40	0.58	1.11	0.94	0.87	1.14	0.98	0.93	1.15
Intercept	0.12	0.1	-0.2	0.001	0.12	-0.13	$0.05*10^{11}$	$0.25*10^{11}$	$-0.009*10^{11}$	$0.02*10^{11}$	$0.12*10^{11}$	$-0.1*10^{11}$
R	0.59	0.85	0.97	0.80	0.94	0.97	0.86	0.98	1.0	0.97	0.98	0.99
RMS	0.36	0.15	0.11	0.34	0.13	0.11	$0.55*10^{11}$	$0.24*10^{11}$	$0.16*10^{11}$	$0.24*10^{11}$	$0.18*10^{11}$	$0.17*10^{11}$
N	720	690	567	720	571	580	720	720	578	720	622	563

5 DEAP retrievals of SkySpec-2D MAX-DOAS measurements at SPC: comparison with MAPA and MMF results

the SkySpec-2D spectra measured at San Pietro Capofiume (hereafter SPC, see [R-7] for more information) site have been provided to the FRM4DOAS community for centralized processing. Although we are still in the testing phase, some profiles retrievals performed with MAPA and MMF codes, are already available. The FRM4DOAS team kindly provided us the NO_2 and aerosol extinction (preliminary product) profiles and columns retrieved from SkySpec-2D SPC spectra in the VIS range using the two official retrieval codes.

The availability of these retrievals opens the possibility to compare DEAP retrieval results to MAPA and MMF on real data. We also compared obtained tropospheric NO_2 VCDs and NO_2 at surface values. For this exercise, we kept as valid all MAPA and MMF profiles flagged with 0 and 1 quality flags.

In addition, TROPOMI coincident Tropospheric NO_2 VCDs are used, when available, to evaluate the performances of ground-based versus satellite products. In-situ surface NO_2 hourly data routinely collected by Arpae of Emilia Romagna region and freely available from http://www.arpae.it/, are used for comparison of NO_2 surface MAX-DOAS retrievals.

One of the main objectives of project phase II is the exploitation of aerosol profiles informations from coincident ground-based remote sensing data at SPC. Backscatter signal for each days at high temporal resolution is available from ALC [R-7] from 0 to 2.5 Km. Even if the aerosol extinction and the signal from ALC cannot directly be compared, the ALC data can be used to infer informations on cloud and aerosol presence and their vertical distribution during the day that can be used to quantitatively assess DEAP performances in extinction profiles retrievals.

The test was performed on some case studies. We apply the DEAP algorithm to SkySpec-2D SPC MAX-DOAS observations obtained on 1, 7 and 10 October 2021 and 14 December 2021. We prefer to use data in Autumn (even if they are not the last measurements we have) due to the higher probability of finding higher NO_2 values and more peculiar dynamical conditions with respect to spring and summer data, as can be seen from NO_2 total columns in [R-8].

For this scope, we used observations performed in the west direction (300 degrees). Only the SCDs from the visible channel are used. The NO_2 and O_4 dSCDs at off-zenith observation angles were obtained from MAX-DOAS sequences using as a reference the zenith spectra corresponding to each sequence. The details on the DOAS analysis performed with QDOAS to obtain the SCDs are reported in [R-7].

The off-axis dSCDs are given in input to the DEAP code. We used for these analysis a slightly different configuration with respect to the one used for code validation with synthetic dSCDs: The a-priori error was set to 150 percent for aerosol retrievals and to 450 percent for NO_2 and the correlation lengths were set to 0.3 km for aerosol and 0.6 km for NO_2. The Marquardt lambda was set to 0.25. We relaxed the a-priori information in order to avoid constrain that can produce bias when the a-priori profile is really far from the real one. On the other hand, this implies that stronger oscillations and less smooth profiles are obtained. We do not apply any scaling to O_4 dSCDs even if we know that sometimes this approach is used for aerosol extinction retrievals. For this reason, in the comparison with MAPA we selected the results

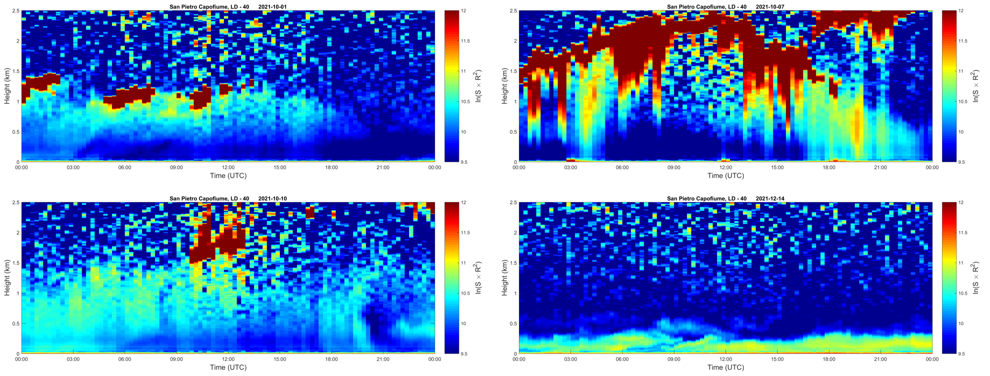

Figure 5.1: ln(SxR^2) as a function of time and altitude from ALC for 1, 7 , 10 October 2021 and 14 December 2021.

obtained with no O_4 scaling.

A-posteriori data filtering is performed using as an indicator the difference between retrieved and measured dSCDs: for O_4 dSCDs a RMSE of 10.5*10^{42} molec/cm^2 and for NO_2 dSCDS a RMSE of 1.5*10^{17} molec/cm^2 are used as thresholds. Thresholds on Chi-square values can be used as well.

5.1 EXTINCTION PROFILES

Aerosol extinction is the first step in many codes developed for MAX-DOAS retrievals. Their quality also impact the quality of the subsequent NO_2 profile retrievals. Figure 5.1 shows the logarithm of (SxR^2) from ALC as a function of time and altitude up to 2.5 km. High values (red color) show the evidence of clouds/thick aerosol layers, medium values (in light green) thin aerosol presence. The ALC measures 24 hours continuously. (**CREDO**)

Obviously we are interested only in day-time data. As can be seen, the only completely clear sky day is the 14 December 2021 when only aerosols at low levels are present. The 7 October is a cloudy day: we expect some criticalities from MAX-DOAS retrievals in this day. The 1 October has cloud/aerosols at the beginning of the day and then they disappear, while on the 10 October clouds/aerosols are present in the middle part of the day.

Aerosol extinction profiles retrieved by DEAP, as a function of altitude and time, are reported in Fig. 5.2 for the four days.

On 1 October 2021, the extinction profile at around 10 AM increased at about 1 km, in agreement with Fig. 5.1. THigh extinction values characterize this day during the morning and a clear sky condition after about midday. On 7 October, the extinction profiles show high values up to 0.8 km^{-1} almost all the day at altitudes of 1.2-1.8 km, quite in line with ALC data. On 10 October, we see aerosol presence at the beginning of the day then clouds are present around mid-day. On 14 December, we found an aerosol layer near the surface for the whole day.

Aerosol extinction profiles from MAPA are reported in Fig. 5.3 while MMF data in 5.4.

All the three codes agree on the aerosol layer at ground on the 14 December 2021. In the

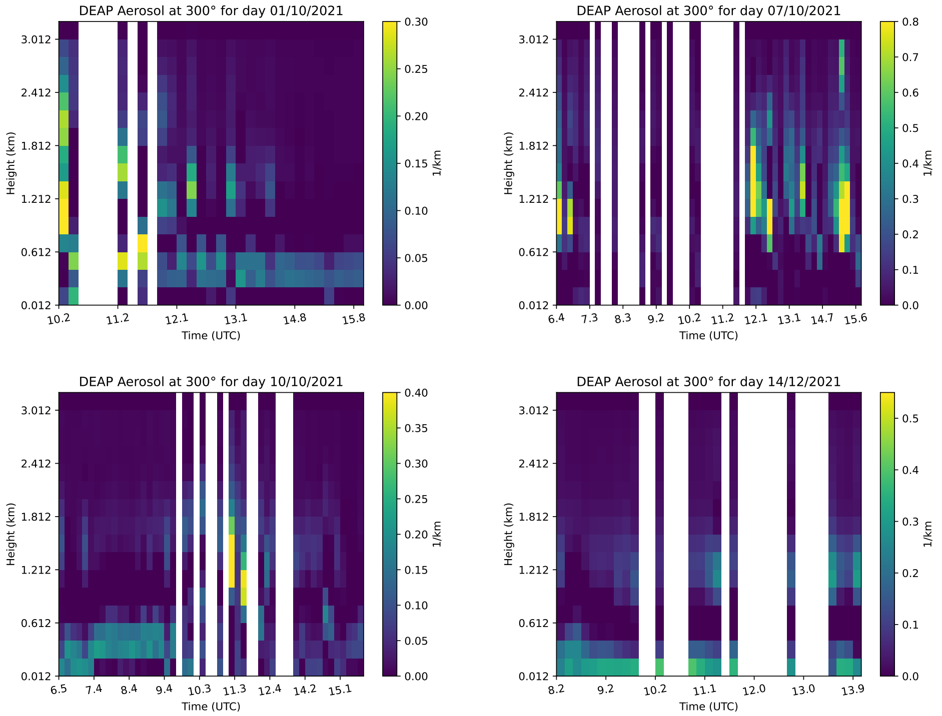

Figure 5.2: Extinction profiles a function of time and altitude from DEAP for 1, 7 , 10 October 2021 and 14 December 2021.

other days we can find some differences with DEAP being in slightly better agreement with MMF.

We recall here that MAPA and MMF aerosol retrievals are preliminary.

5.2 NO_2 PROFILES

NO_2 profiles retrieved from DEAP, MAPA and MMF are reported in Figs. 5.5-5.7. We observed a better agreement between the three codes considering the aerosol extinction results.

The DEAP profiles are the most oscillating due to the high a-priori error. Instead, the MMF profiles are the smoothest and characterized by the lowest values.

In general, the NO_2 behaviour is really similar for all the codes on the four days: on the 1 October we observe higher values in the morning at low altitudes until 9 A.M. On the 14 December, instead, the NO_2 values are much higher at the surface (up to a factor 10) with respect to the other days.

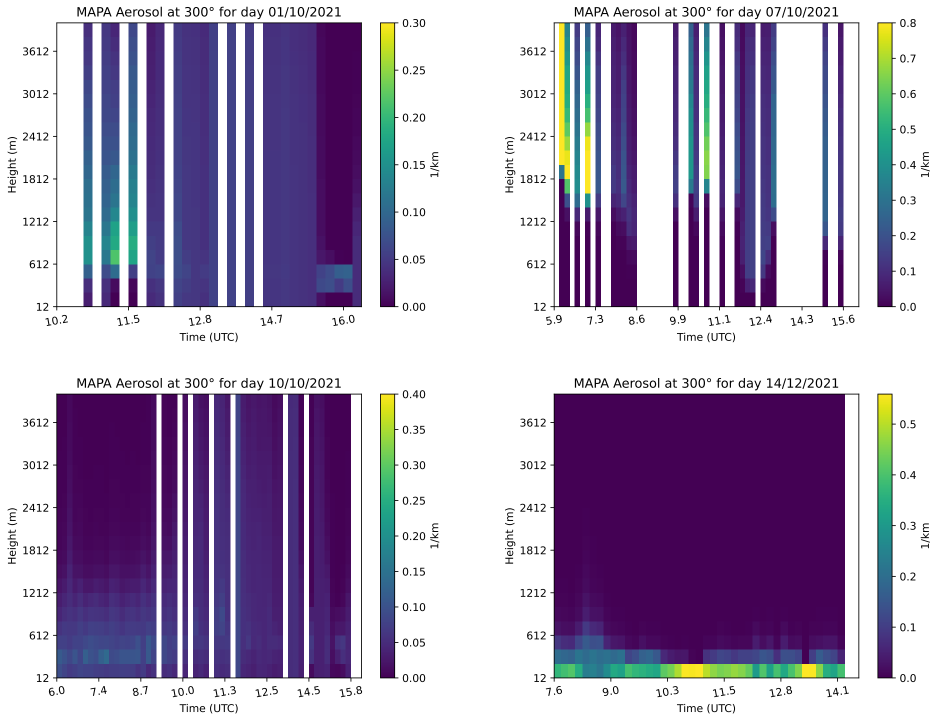

Figure 5.3: Extinction profiles a function of time and altitude from MAPA for 1, 7 , 10 October 2021 and 14 December 2021.

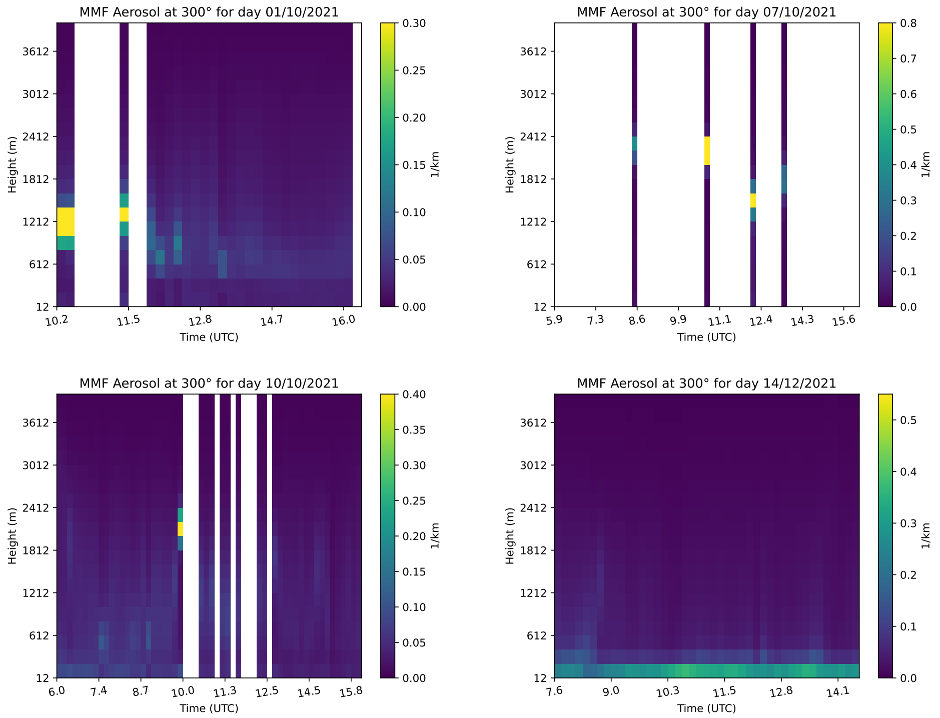

Figure 5.4: Extinction profiles a function of time and altitude from MMF for 1, 7 , 10 October 2021 and 14 December 2021.

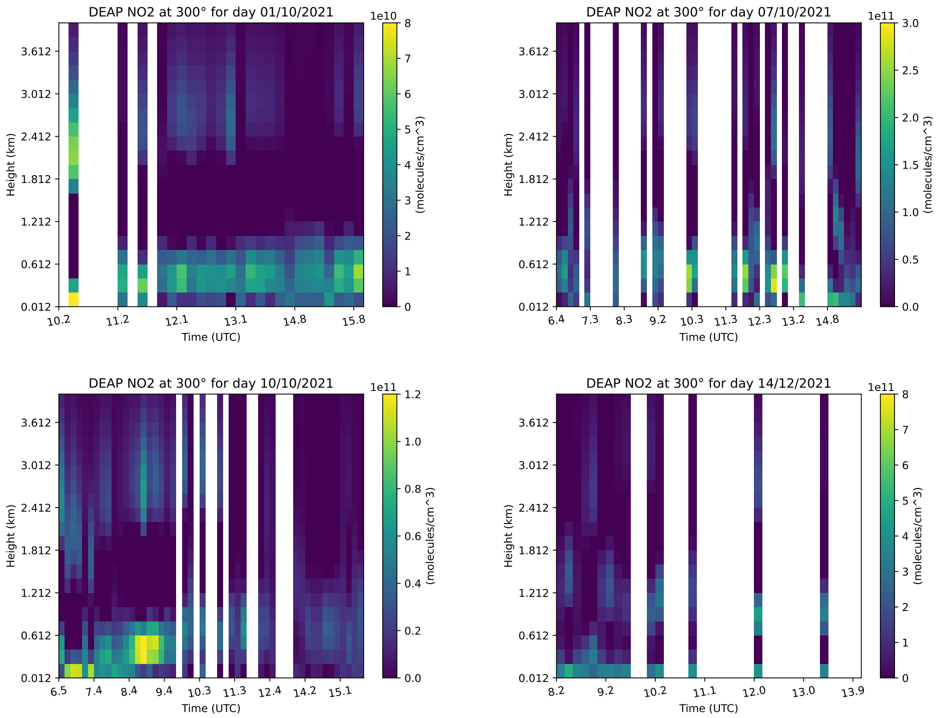

Figure 5.5: NO_2 profiles a function of time and altitude from DEAP for 1, 7 , 10 October 2021 and 14 December 2021.

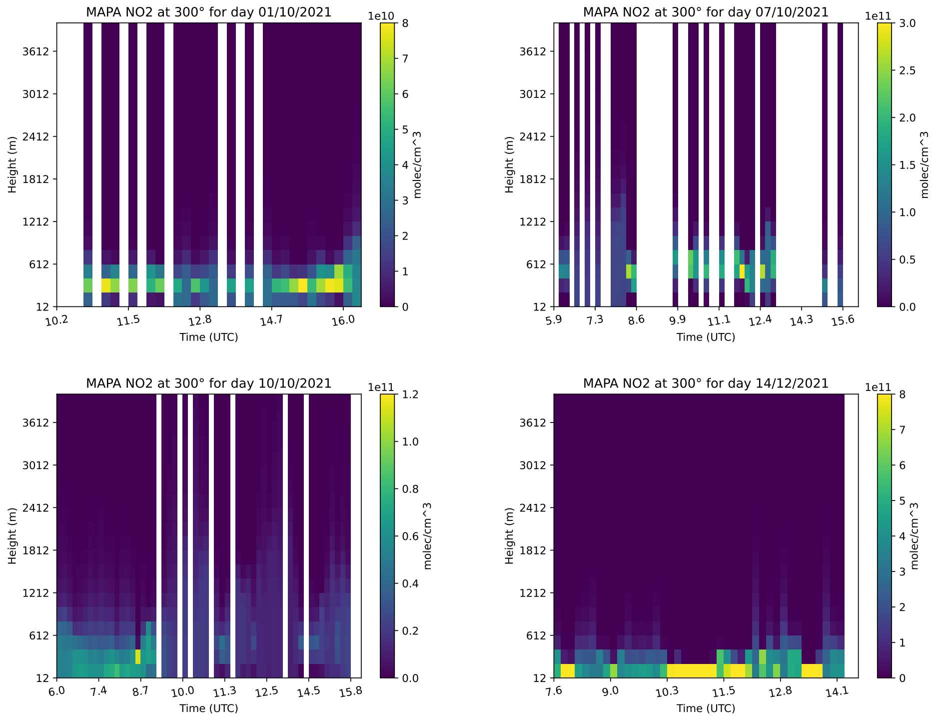

Figure 5.6: NO_2 profiles a function of time and altitude from MAPA for 1, 7 , 10 October 2021 and 14 December 2021.

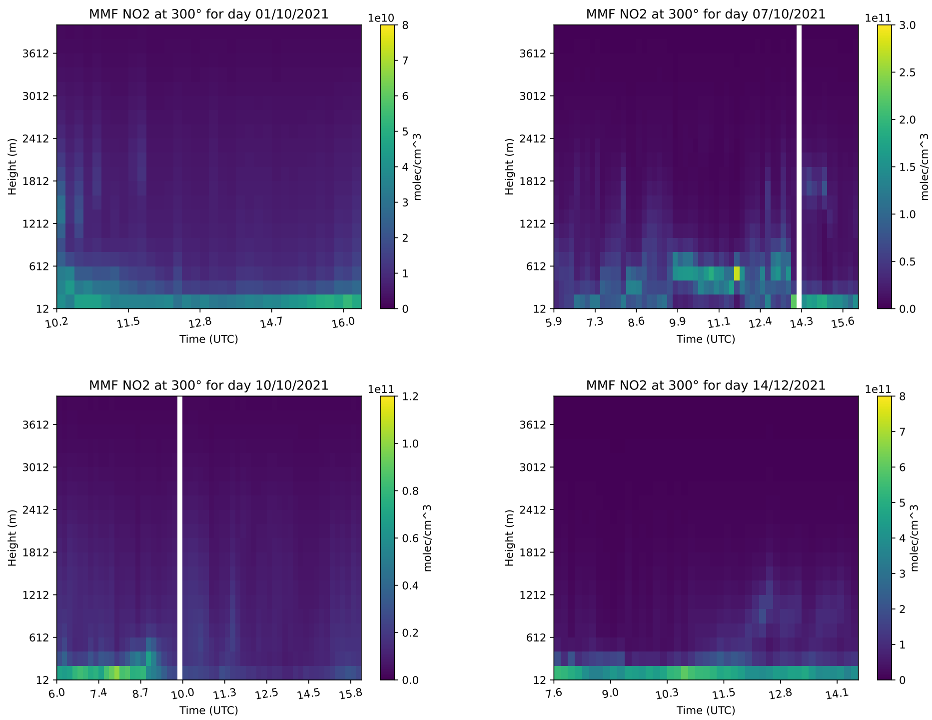

Figure 5.7: NO_2 profiles a function of time and altitude from MMF for 1, 7 , 10 October 2021 and 14 December 2021.

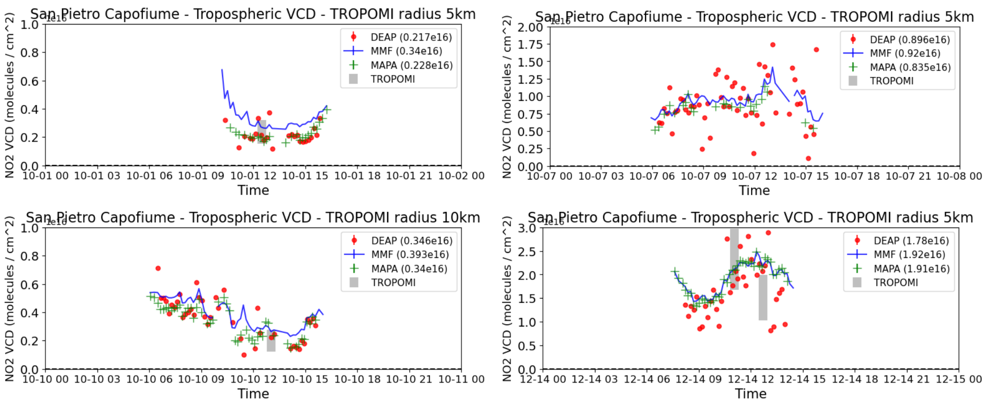

Figure 5.8: NO_2 tropospheric column retrievals for 1, 7, 10 October 2021 and 14 December 2021 from DEAP (red), MMF (blue), and MAPA (green) together with TROPOMI coincident values (when available) in grey.

5.3 DEAP NO_2 TROPOSPHERIC COLUMN RETRIEVALS: COMPARISON WITH MAPA, MMF AND TROPOMI RESULTS

NO_2 tropospheric column retrievals are the most important products for MAX-DOAS measurements when these measurements are used in the frame of satellite validation.

NO_2 tropospheric columns from DEAP (red dots), MAPA (green crosses), MMF (blue line) and TROPOMI (grey shadow representing the standard deviations of coincident observations) are reported in Fig. 5.8.

For TROPOMI we used data with quality flag above 0.75, in a radius of 5 km for 1 and 7 October and 14 December. Due to overcast conditions no TROPOMI coincidences can be found on the 7 October (even if we use a larger radius). On 10 October, we use a radius of 10 km in order to find a coincident point. This is due to cloud presence at around TROPOMI overpass time as shown in Fig. 5.1.

Results show a really good agreement of DEAP Tropospheric VCDs with MAPA VCDs on 1 October. MMF results are higher in this case. On the 7 October and 14 December DEAP shows more oscillations, while on 10 October the results are similar for all the three codes.

Scatterplots of VCDs obtained with different retrieval codes are reported in Fig. 5.9. The correlation of MMF and MAPA VCD is total with differences of the order of 5%. The DEAP results correlate similarly with the two official codes (correlation of 0.69) with differences of about 10%. An increase in scattering of results is found for high VCDs for DEAP retrievals.

Respect to TROPOMI, the agreement is quite good for all the codes.

5.4 DEAP NO2 SURFACE VALUES RETRIEVALS: COMPARISON WITH MAPA, MMF RESULTS AND IN-SITU DATA

The last point of the retrieved profile can be used for comparisons with in-situ data at the surface. NO_2 at the surface from DEAP (red dots), MAPA (green crosses), MMF (blue line) and Arpae in situ data (blue dots) are reported in Fig. 5.10.

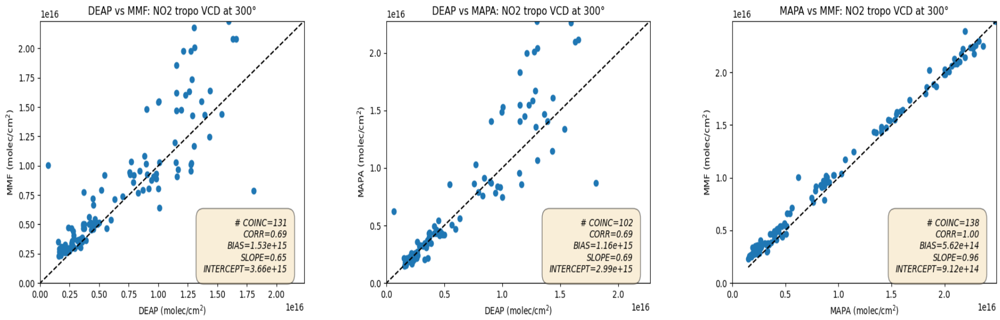

Figure 5.9: Scatterplots of NO_2 tropospheric column retrievals from DEAP, MMF and MAPA.

In general, DEAP performs similarly to MAPA and MMF. MMF gives the smoothest results while MAPA is higher and more oscillating than DEAP and MMF on 14 December. With respect to the outcomes of the validation with synthetic dSCDs (Sect. 4.4.3), DEAP does not show any underestimation with respect to the other codes.

The comparison with in-situ data highlights interesting features: during night the NO_2 values are higher than during the day when photolysis acts as a sink. This feature is valid for all the analyzed days apart for the 7 October, when the daily behaviour is reversed, with day-time data higher than the night-time ones. This behaviour is probably due to overcast conditions that prevent photolysis during the day. On the 14 December 2021, the NO_2 in-situ data show almost constant values over day and night. This is possibly related to stable conditions that cause high NO_2 values at the surface also in MAX-DOAS at surface values.

The surface values retrieved with the three codes agree quite well with in-situ data for 1 October and 14 December. On 10 October the agreement is really good with MAX-DOAS retrievals catching the NO_2 variations at the beginning of the morning. On 7 October the MAX-DOAS results agrees well apart from the central part of the day where probably the cloud effects affect the MAX-DOAS results.

Scatterplots of NO_2 at surface obtained with the three retrieval codes are reported in Fig. 5.11. With respect to tropospheric VCDs, the correlation of MMF and MAPA is lower in this case with differences for the two codes generated by high MAPA values retrieved on the 14 December. The DEAP results correlate differently with MAPA (correlation of 0.87) and MMF (correlation 0.90). The differences are of the order 10% for all the combinations.

6 Conclusions

We developed the DEAP code for the retrieval of NO_2, HCHO and aerosol extinction profiles from MAX-DOAS measurements obtained with the SkySpec-2D instrument located in the SPC and already exploited in zenith sky mode during WP2250 DOAS-BO Phase I project. The code is a two-step algorithm that exploits the Optimal estimation. In the first step, the aerosol extinction profile is retrieved by exploiting the SCIATRAN forward model. In the second step, the NO_2 (or HCHO) retrieval using Box-AMF is obtained with SCIATRAN and exploiting the

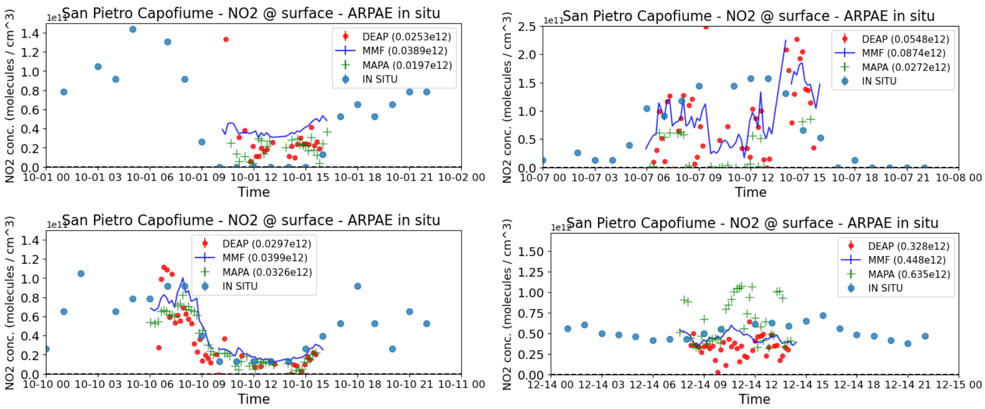

Figure 5.10: NO_2 at surface retrievals for 1, 7, 10 October 2021 and 14 December 2021 from DEAP (red), MMF (blue), and MAPA (green) together with Arpae hourly mean values (blue dots).

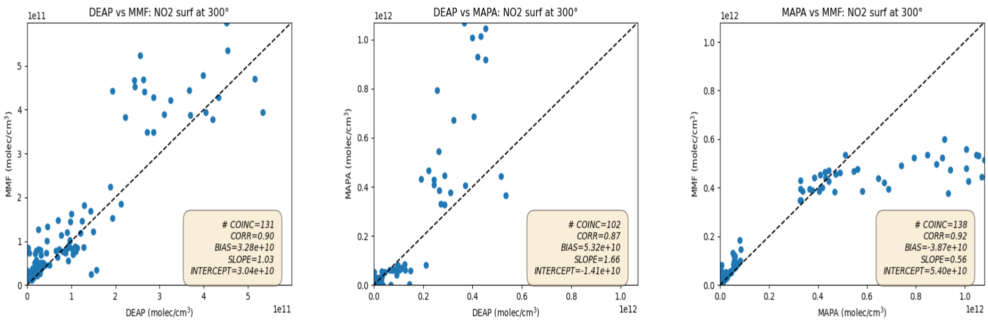

Figure 5.11: Scatterplots of NO_2 surface values retrievals from DEAP, MMF and MAPA.

aerosol extinction profile already retrieved in the first step.

The code has been validated using the synthetic dSCDs provided in the frame of the FRM4DOAS project.

The main products retrieved through the new code are the aerosol extinction profiles at 360 and 477 nm, the NO_2 and HCHO profiles, AOT, tropospheric columns and values at the surface. These products, obtained with the DEAP code, have been compared considering other state-of-the-art algorithms, in particular, MAPA and MMF, the two codes used for the FRM4DOAS centralized processing.The validation exercise shows that the DEAP results are fully consistent with both MAPA and MMF, even if the code performs slightly worse than the two operational ones. As it is the first version of DEAP, we can consider the validation of the code as successful.

Since the beginning of 2022, the SkySpec-2D SPC spectra have been provided to the FRM4DOAS community for centralized processing. Although still in the testing phase, the FRM4DOAS team kindly provided us with the NO_2 and aerosol extinction profiles and columns retrieved with MAPA and MMF using SPC spectra in the VIS range. The possibility of performing comparisons of results obtained with DEAP and official codes on real data is really important to understand code performances and features and give hints of possible future improvements. Tests were performed over four days in autumn/winter 2021 when particular atmospheric conditions, typical of the Po Valley were present. We used only the measurements acquired in the 300 degrees azimuthal direction in the visible spectral range.

Satellite data (e.g. TROPOMI NO_2 Tropospheric VCDs) and in-situ data (e.g. NO_2 at surface from Arpae) as well as aerosols profiles over the whole day from ground-based instruments (e.g. ALC) were used for comparisons.

The results obtained from the three codes agree well for NO_2 profile retrievals as well as for values at the surface and NO_2 Tropospheric columns. Slightly higher differences are found for aerosols extinction retrievals. However similar behaviors are found. In particular, the aerosol extinction profiles shape during the day follows quite well the profiles measured by ALC. NO_2 Tropospheric columns agree well with TROPOMI where coincident data are present. The agreement with in-situ data at the surface is also satisfactory.

The possibility of exploiting the synergies from simultaneous measurements from remote sensing (MAX-DOAS and ALC) and in-situ data at the SPC station is an important features of this project that will be further investigated.

7 Aknowledgements

For the MAPA and MMF retrievals we kindly acknowledge the FRM4DOAS and FRM4DOAS-2.0 projects (ESA contracts nř4000118181/16/I-EF and 4000135355/21/I-DT-Ir) and, in particular Caroline Fayt, Martina M. Friedrich, François Hendrick (IASB-BIRA) and Steffen Beirle (MPIC). Arpae NO_2 in situ hourly data have been dowloaded from http://www.arpae.it/.

REFERENCES

[R-1] E. Castelli, P. Pettinari, E. Papandrea, P. Cristofanelli, M. Busetto, M. Valeri, D-3 Report on measurements campaign within the ISAC-CNR (Bologna) site exploiting in-situ and the satellite-borne synergies, 2021, https://doi.org/10.5281/zenodo.5886896

[R-2] Rozanov, V.V., Rozanov, A.V., Kokhanovsky, A.A., Burrows, J.P.: Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN, Journal of Quantitative Spectroscopy Radiative Transfer, Vol. 133, 2014, pp. 13 - 71, DOI: 10.1016/j.jqsrt.2013.07.004

[R-3] Deliverable D5: MAXDOAS Algorithm Round-Robin Definition and Results Document, Date: 30/01/2018 Version: 5.0, ESA Contract No. 4000118181/16/I-EF.

[R-4] Friess, U., Beirle, S., Alvarado Bonilla, L., Bösch, T., Friedrich, M. M., Hendrick, F., Piters, A., Richter, A., van Roozendael, M., Rozanov, V. V., Spinei, E., Tirpitz, J.-L., Vlemmix, T., Wagner, T., and Wang, Y.: Intercomparison of MAX-DOAS vertical profile retrieval algorithms: studies using synthetic data, Atmos. Meas. Tech., 12, 2155–2181, https://doi.org/10.5194/amt-12-2155-2019, 2019.

[R-5] Beirle, S., Dörner, S., Donner, S., Remmers, J., Wang, Y., and Wagner, T.: The Mainz profile algorithm (MAPA), Atmos. Meas. Tech., 12, 1785–1806, https://doi.org/10.5194/amt-12-1785-2019, 2019.

[R-6] Friedrich, M. M., Rivera, C., Stremme, W., Ojeda, Z., Arellano, J., Bezanilla, A., García-Reynoso, J. A., and Grutter, M.: NO_2 vertical profiles and column densities from MAX-DOAS measurements in Mexico City, Atmos. Meas. Tech., 12, 2545–2565, https://doi.org/10.5194/amt-12-2545-2019, 2019.

[R-7] P. Pettinari, M. Valeri, E. Papandrea, E. Castelli, L. Di Liberto, A. Marinoni, S. Decesari: [D-1] Report on the MAX-DOAS analysis chain,WPs 2250-2251: DOAS-BO: Towards a new FRM4DOAS-compliant site - Phase 2, Issue 1, 05/09/2022

[R-8] Pettinari, P.; Castelli, E.; Papandrea, E.; Busetto, M.; Valeri, M.; Dinelli, B.M. Towards a New MAX-DOAS Measurement Site in the Po Valley: NO2 Total VCDs. Remote Sens. 2022, 14, 3881. https://doi.org/10.3390/rs14163881