**WPs 2250-2251: "DOAS-BO: Towards a new FRM4DOAS-compliant site - Phase 2"**

# [D-1] Report on the MAX-DOAS analysis chain

| | |
|---|---|
| **Document reference:** | FRM4DOAS-BO_Phase2_D1 |
| **Document Issue:** | 2.0 |
| **Document Issue date:** | 06/10/2023 |
| **Document authors and affiliations:** | P. Pettinari[1], M. Valeri[2], E. Papandrea[1], E. Castelli[1], L. Di Liberto[1], A. Marinoni[1], S. Decesari[1]<br><br>1 CNR-ISAC<br>2 Serco Italia S.p.A |

**AMENDMENT RECORD SHEET**

The Amendment Record Sheet below records the history and issue status of this document.

| ISSUE | DATE | REASON |
|---|---|---|
| 1.0 | 05/09/2022 | D-1 document of project "WPs-2250-2251 Phase 2: DOAS-BO: Towards a new FRM4DOAS-compliant site'". |
| 2.0 | 06/10/2023 | Updated taking into account the introduction of the DEAP retrieval code in the processing chain and the provision of figures for the DOAS-BO website |

# Contents

# 1. List of Acronyms

| | |
|---|---|
| ACTRIS | Aerosol, Clouds and Trace Gases Research Infrastructure |
| AKM | Averaging Kernel Matrix |
| AMF | Air Mass Factor |
| ARPAE | Agenzia Regionale per la Prevenzione, l’Ambiente e l’Energia |
| AOD | Aerosol Optical Depth |
| CCD | Charge-Coupled Device |
| CNR | Consiglio Nazionale delle Ricerche |
| DEAP | DOAS optimal Estimation Atmospheric Profile retrieval |
| DOAS | Differential Optical Absorption Spectroscopy |
| FOV | Field Of View |
| FRM4DOAS | Fiducial Reference Measurements for DOAS |
| ISAC | Istituto delle Scienze dell’Atmosfera e del Clima |
| LIDAR | Light Detection and Ranging |
| MAX-DOAS | Multi-AXis DOAS |
| NDACC | Network for the Detection of Atmospheric Composition Change |
| OE | Optimal Estimation |
| SCD | Slant Column Densities |
| SNR | Signal to Noise Ratio |
| SPC | San Pietro Capofiume |
| SZA | Solar Zenith Angle |
| UV | Ultra-Violet |
| VCD | Vertical Column Density |
| VIS | Visible |

# 2. Introduction

This document is the report of the activities performed in the frame of the project “WPs-2250-2251: DOAS-BO: Towards a new FRM4DOAS-compliant site - Phase 2’’. Here, we briefly introduce the facility of San Pietro Capofiume and the instruments within the site that we exploit for the activities foreseen in this project. In particular, we describe the analysis chain implemented to achieve the expected output (i.e., calibrated spectra, slant column densities, etc.) from diffuse solar spectra measured by the SkySpec-2D in San Pietro Capofiume.

# 3. Location and Instruments

## 3.1. “Giorgio Fea” observatory at San Pietro Capofiume

The "Giorgio Fea" observatory (https://www.isac.cnr.it/it/node/7803, Latitude: 44.65° N, Longitude: 11.62° E, Altitude: 11 m a.s.l.) is located at the rural site of San Pietro Capofiume - Bologna, Italy (hereafter SPC). The site, founded in the early 1980s, is managed by the Agenzia Regionale per la Prevenzione, l’Ambiente e l’Energia (ARPAE, https://www.arpae.it/it/arpae/arpae) of Emilia Romagna, while CNR-ISAC operates in the field under the umbrella of a long-term agreement with ARPAE. The station is equipped for in-situ monitoring of trace gases and particulate matter sampling for atmospheric chemical speciation. ARPAE also runs radar measurements, radio soundings and operates a phenological station. The station is part of the European Research Infrastructure ACTRIS, as Mt. Cimone - Po Valley facility (CMN-PV). Currently, a MAX-DOAS instrument and an Automatic LIDAR/Ceilometer are operating on the site, a Raymetrics LIDAR is in the setting up phase.


***Figure 1**: Location of the “Giorgio Fea” observatory at San Pietro Capofiume (Bologna, Italy) and position of the instruments within the site.*

## 3.2. SkySpec-2D

SkySpec-2D is a MAX-DOAS instrument developed by Airyx GmbH (https://airyx.de/wp-content/uploads/2021/03/SkySpec-all_2021-03-09.pdf), fully compliant with the FRM4DOAS standard requirements, which allows performing reliable atmospheric measurements. The SkySpec-2D system is composed of a telescope, a case containing two spectrometers and a PC. The light acquired by the telescope is collected in the two spectrometers by an optical fiber. The two spectrometers allow for the measurement of two simultaneous spectra: one in the VIS (from 400 to 550 nm) and the other in the UV (from 300 to 400 nm) spectral range. The PC drives the instrument through the acquisition software MS-DOAS, provided by Airyx GmbH.

***Figure 2**: SkySpec-2D in "Giorgio Fea" observatory at SPC.*

## 3.3. Automatic LIDAR Ceilometer

The VAISALA LD40 Lidar-Ceilometer (Fig. 3) is an active remote sensing sensor that uses laser pulses backscattered from the atmosphere to determine cloud height and vertical aerosol distribution. The VAISALA LD40 is a simple system with a single wavelength at 855 nm but is able to evaluate the distribution of aerosols and clouds in the first 4-5 km. The instrument measures the vertical distribution of particulate matter with high vertical and temporal resolution from which the mixing layer height (MLH) and the cloud base, as well as the presence and vertical extension of desert dust, fire plume, volcanic ash, mists, and fine clouds.

The system is installed inclined at 45 degrees in order to have better information on the lower layers. The ceilometer measurements were performed 24/7 to monitor and characterize the evolution of the atmospheric layers.

***Figure 3:*** *Ceilometer VAISALA LD40 in "Giorgio Fea" observatory at SPC.*

The PI of this instrument is Luca Di Liberto (CNR-ISAC). In the frame of this project, we will use instrument's products as is. The VAISALA LD40 backscattered signals will be inter-compared against profiles retrieved using the retrieval code under development and the SkySpec-2D MAX-DOAS observations [WP2250-1]. The entire analysis will be performed in liaison with the PI of the instrument and other CNR-ISAC experts in order to fully exploit expertise and know-how related to the VAISALA LD40 instrument and the SPC site.

## 3.4. Raymetrics LIDAR

Despite the presence of the VAISALA LD40, considering state of the art for studying the distribution of aerosols and clouds in the atmosphere, the lidar is the best performing instrument to provide high spatial and temporal resolution information on the layers of the atmosphere. The high-resolution Raymetrics LiDAR LR221-D300 works with a Nd YAG laser, with emission at 355 and 532 nm and is fully compliant to the minimum requirements of ACTRIS guidelines for the aerosol remote sensing component. It has been installed at the beginning of 2022 at the SPC observatory (Fig. 4). The Raymetrics Raman LIDAR system is an active laser remote sensing instrument that measures the vertical profile of suspended aerosols and atmospheric molecules. The LIDAR system is a powerful tool for monitoring the evolution of various meteorological and atmospheric parameters that vary with high spatial and temporal resolution.
Measurements of aerosol extinction, backscatter and depolarization-ratio profiles are performed at both mentioned wavelengths (355 and 532 nm). Technical system parameters such as laser power, telescope aperture, receiver bandwidth and data acquisition system were chosen such that profiles can be acquired throughout the troposphere up to the lower stratosphere with the required accuracy and temporal

and spatial resolution. A separate near-range receiving system is installed for observations in the lower planetary boundary layer.
The Raymetrics LIDAR system is capable of detecting the Raman shifted backscattered radiation. The use of a Raman Lidar is a well-established method of determining the vertical profile of the aerosol extinction coefficient independently from the aerosol backscatter coefficient, resulting in a dramatic reduction in the associated uncertainties. When we are writing, the Raymetrics system is still in the configuration phase, but in the operational phase the ACTRIS high-power aerosol lidar will be operated continuously, in the absence of precipitation or fog.


***Figure 4:*** *Raymetrics LIDAR in "Giorgio Fea" observatory at SPC.*

The Raymetrics LIDAR is managed by CNR-ISAC in collaboration with Angelo Lupi (CNR-ISP). Similar to the VAISALA LD40, since instrument setup and running are out of this project's scope, we will use the instrument's products (i.e., aerosol profiles) as is, exploiting the support of the instrument's PI as necessary. The Raymetrics LIDAR system products (interpolated at the DOAS wavelength if needed) will be inter-compared against similar products retrieved using the retrieval code under development and the SkySpec-2D MAX-DOAS observations [WP2250-1]. As highlighted in the previous section, we will perform the entire analysis in strictly collaboration with the PI of the instrument and other CNR-ISAC experts to fully exploit expertise and know-how related to the new instrument and the SPC site.

# 4. SkySpec-2D Processing Chain

## 4.1. Observation strategy

SkySpec-2D has been continuously acquiring MAX-DOAS measurements at the "Giorgio Fea" observatory in SPC since the 1st of October 2021. During all these months, the measurement strategy has remained unchanged except for the azimuth directions selected. Indeed, we decided to modify them because the telescope, on the 23rd March 2022, was moved and installed in its permanent position (a few meters away from the previous one) and the previous chosen viewing directions were not free from obstacles anymore. The SkySpec-2D was permanently installed on the roof of the shelter containing the PC and spectrometers, as can be seen in Fig. 2.

According to the settings chosen in the acquisition software MS-DOAS, fully compliant with the FRM4DOAS guidelines [R1] [R2], SkySpec-2D performs three different kinds of measurements every day: calibration spectra, atmospheric spectra and horizon scans.

- The **calibration spectra** are automatically measured during the night, precisely when the SZA becomes higher than 100° (when the sun is 10° below the horizon). Once acquired, SkySpec-2D stops to measure until the SZA does not reach the chosen value to start the atmospheric acquisition (94°). The calibration spectra are important to correct the atmospheric spectra properly and consist of offset, dark current and emission lines of a Hg lamp mounted inside the instrument.
  The routine measurements of Hg emission lines are important to calibrate the measurements spectrally. All the Hg emission spectra are measured with a total integration time of 1 minute. It's important to mention that every acquired spectrum is the sum of several single spectra, with the purpose of increasing the SNR. Every Hg spectrum is acquired with an exposure time of 0.28 s (0.36 s) in the VIS (UV) to obtain well-defined emission peaks and prevent the CCD sensor's saturation. The different exposure times in VIS and UV are due to the different emission lines intensities in the two spectral ranges. As a consequence of these measurement times, a Hg spectrum in the VIS (UV) is the sum of about 210 (160) single spectra. However, since we need at least three different emission peaks per spectral range for a precise spectral calibration and since the actual exposure times lead to only two well defined peaks in the VIS, we need to measure the VIS Hg lines with a higher exposure time of 10 s (only 6 co-added spectra). This allows us to measure a third less intense emission line. From now on, we will always refer to the total measurement time of a spectrum (composed of the sum of single spectra) as "total integration time" and to the measurement time of a single spectrum as "exposure time". The offset and dark current are needed to correct the counts in the acquired atmospheric spectra. The offset is a signal automatically added to the spectra during the acquisition, which must be removed. It's constant and does not depend on the exposure time, therefore it can be estimated through a dark measurement with a very low exposure time. We estimate it using a total integration time of 1 minute and an exposure time of 0.01 s (6000 co-added spectra). The dark current signal must be removed from the atmospheric spectra as well; however, it's not constant and increases linearly with the exposure time. We measure it with a total integration time of 60 s and an exposure time of 10 s (6 co-added spectra). Information on the calibration measurements is summarized in Table 1.

| | OFFSET | DARK CURRENT | Hg | Hg* (only VIS) |
|---|---|---|---|---|
| Total integration time (s) | 60 | 60 | 60 | 60 |
| Exposure time (s) | 0.01 | 10 | 0.28 (VIS), 0.36 (UV) | 10 |
| Coadded spectra | 6000 | 6 | 210 (VIS), 160 (UV) | 6 |

***Table 1:*** *Technical parameters for calibration measurements during nighttime*

- The **atmospheric spectra** start to be acquired every morning when the SZA becomes lower than 94° (the sun is 4° below the horizon). In the beginning, the SkySpec-2D starts to acquire only zenith-sky spectra. When the SZA becomes lower than 85°, SkySpec-2D starts to perform MAX-DOAS measurements. During MAX-DOAS acquisitions, SkySpec-2D measures in three different azimuth directions: 120°, 225° and 300° from the 1st of October 2021 to the 23rd of March 2022, and 135°, 250° and 315° afterwards (Fig. 5). For each azimuth direction, spectra are acquired at the following elevation angles: 1°, 2°, 3°, 5°, 10°, 30° and 90°.

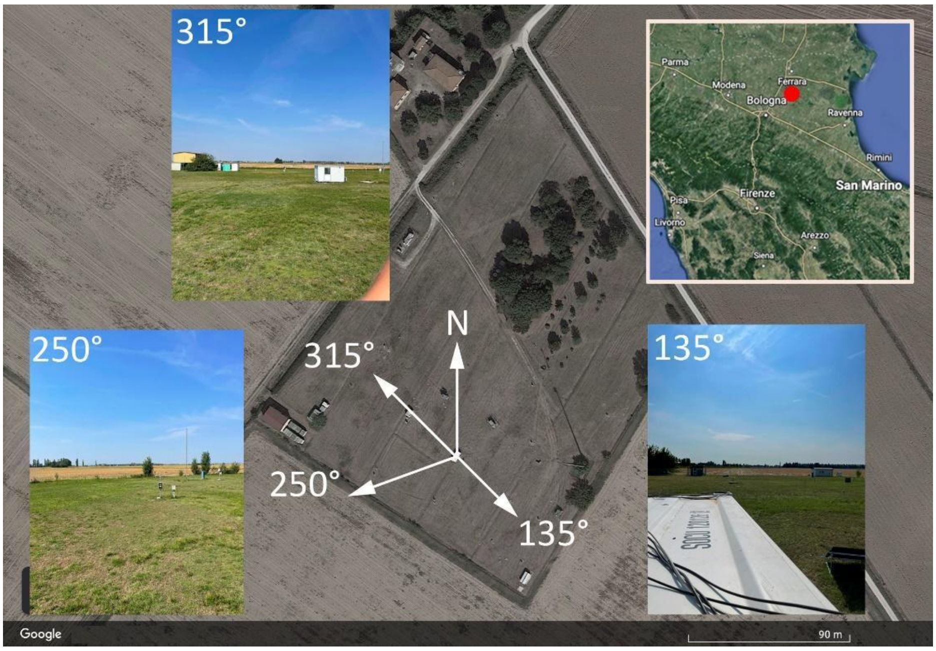


***Figure 5:*** *Different views from SkySpec-2D in the three azimuth directions (135°, 250° and 315° after the 23rd of March 2022) adopted during MAX-DOAS measurements.*

The acquisition system automatically avoids measuring when the instrument viewing direction is close to the sun position (less than 5°).
Each spectrum is acquired in a total integration time that depends on the SZA. We chose 50 s in the SZA range between 94° and 92°, 40 s between 92° and 90°, 30 s between 90° and 80°, and 20 s when SZA is lower than 80°. The purpose of increasing the total integration time at high SZAs is to increase the SNR of those measurements. Indeed, at high SZAs, the exposure time increases (less light) and the number of the co-added single spectra to sum up decreases, leading to a lower SNR.
Indeed, the exposure time of the single summed spectra is variable and is automatically set in order to prevent the CCD sensor saturation. As a consequence, also the number of the co-added single spectra is variable and depends on both the total integration time and exposure time. All the information is summarized in Table 2.

| | ZENITH ACQUISITION | MAX-DOAS ACQUISITION |
|---|---|---|
| SZA range (°) | 94-85 | < 85 |
| Total integration time (s) | 50 (92°< SZA < 94°), 40 (90°< SZA < 92°), 30 (85°< SZA < 90°) | 30 (80° < SZA < 85°), 20 (SZA < 80°) |
| Exposure time (s) | automatic | automatic |
| Coadded spectra | automatic | automatic |
| Azimuth directions (°) | / | 120, 225, 300 (1/10/2021 – 23/03/2022) 135, 250, 315 (23/03/2022 - nowadays) |
| Elevation angles (°) | 90 | 1, 2, 3, 5, 10, 30, 90 |

***Table 2:*** *Technical parameters for atmospheric measurements during daytime.*

- **Horizon scans**. In the observation strategy, we implemented the horizon scans to be performed daily around noon. Assuming that the true horizon position in the three different azimuth angles (see Table 2) does not change with time, horizon scans allow us to assess the pointing stability of the SkySpec-2D, which is very important for reliable MAX-DOAS measurements. At this scope, for

each of the three azimuth directions, VIS and UV spectra are acquired within the elevation angles range of ±3°, with a step of 0.2°, a fixed total integration time of about 1 s and 100 co-added spectra. This measurement strategy is applied twice for each azimuth direction, the first from -3° to +3° (upwards) and the second from +3° to -3° (downwards), in order to assess if systematic differences in the telescope movements occur between upwards and downwards scans. At the end of this process, 4 sets of data are available for each azimuth direction: two scans, upwards and downwards, each containing two other datasets, one in the UV and the other in the VIS.
The spectra obtained in these scans are used to estimate the horizon position, according to the SkySpec-2D, every day. Fig. 6 shows an example of this procedure for 02/04/2022. Each dot represents the mean value of a certain spectrum with respect to the elevation angle. Each set of dots is fitted by an "error function" (the Gaussian integral function) around the maximum slope region that is assumed to represent the horizon position. Indeed, we expect to observe low spectra values when the SkySpec-2D FOV completely crosses the ground (see low elevation angles), and higher values when the telescope is fully observing the sky. The transition region occurs at elevation angles where the instrument FOV partially crosses the sky and the ground. We estimate the center of the fitted "error function" (equivalent to the center of the derivative Gaussian function) as the horizon position. As we can see, our method states that the horizon position during 02/04/2022 at the 135° azimuth direction is 0.2° for all the 4 different datasets.

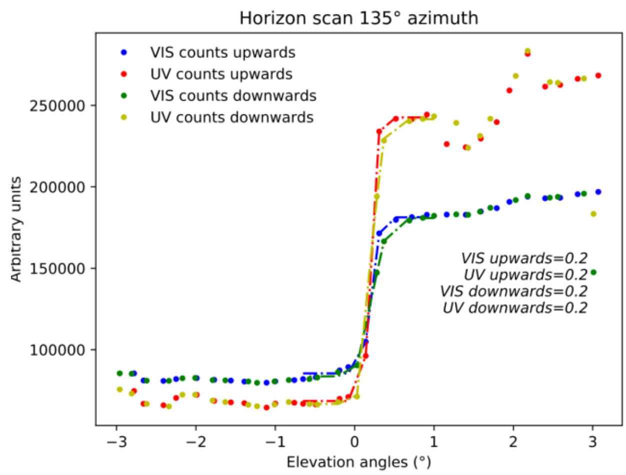


***Figure 6:*** *Mean values of the spectra (dots) for the 4 different datasets acquired during a horizon scan at 135° azimuth direction, during the day 02/04/2022. The dashed lines represent the theoretical "error functions" used to fit the data. The values in the legend represent the horizon positions, estimated as the centers of the theoretical functions, for the 4 datasets.*

## 4.2. Data processing

The processing chain is composed of different steps. At first, spectra are calibrated and converted into NetCDF files, compliant with the FRM4DOAS guidelines [R3]. Then, the NetCDF files are provided as input of the QDOAS software (https://uv-vis.aeronomie.be/software/QDOAS/) to estimate the $NO_2$ SCDs. At the end of the process, the $NO_2$ SCDs can be used to compute $NO_2$ VCDs, from zenith-sky measurements only, and the vertical profiles of $NO_2$ and aerosol extinction coefficient from MAX-DOAS measurements. Here, we will describe all these steps.

### 4.2.1. Conversion of the measured spectra into NetCDF files

The MS-DOAS software acquires the spectra according to the measurement strategy described in Sect. 4.1 and store each spectrum in a binary file that contains the spectrum values and all the information regarding the exposure time, number of co-added spectra, total acquisition time and observation geometry.
Our conversion tool reads these binary files, calibrates the spectra and writes them into NetCDF files, two for each day (one for VIS and one for UV). The final NetCDF files are created according to the FRM4DOAS standard guidelines, as described in [R3]. The conversion tool processes one day of measurements at a time and repeats all the following procedures for each day.

- **Wavelength calibration**

  The first operation performed by the conversion tool is to search the calibration measurements performed during nighttime, as explained in section 4.1. The Hg lamp measurements are used to calibrate the measured spectra spectrally. This step consists of creating a mathematical relation which links each point of the spectrum (each pixel of the CCD sensor) to a wavelength value. Regarding the VIS spectra, 3 Hg emission peaks occurring at known wavelengths (435.8, 491.6, 546.0 nm) are used. This occurrence allows us to create a 2-degree polynomial relation between wavelengths and CCD pixels.
  The same method is applied to calibrate the UV spectra. The only difference is that the Hg lamp presents 4 well-defined peaked emission lines in the UV. Hence, the UV wavelength grid is built through a 3-degree polynomial relation.

- **Counts (intensity) calibration**

  All the other calibration measurements are used to correct the values of the counts in the spectra. In particular, the offset and dark current must be properly subtracted. During this procedure, we also correct the spectra counts for the non-linearity of the two sensors (provided by the Airyx company). During this operation, we must consider that all the saved atmospheric spectra are the sum of a certain number of co-added single spectra, which is not equal to the number of the co-added spectra used in the measured calibration spectra. For this reason, we apply all the corrections to every spectrum after scaling all the calibration spectra. In the beginning, we subtract the constant offset; then, the spectra values are corrected for the sensor non-linearity (which depends on the spectra values). In the end, the dark current signal is subtracted from the spectrum. For this step, since the dark current linearly depends on the exposure time, the measured dark signal must be appropriately scaled to the exposure time of the measured atmospheric spectrum.

- **NetCDF file creation**

All the calibrated spectra and relative information acquired in a day are saved into two NetCDF files (one for the VIS channel and the other for the UV channel). The name of these NetCDF files follows the FRM4DOAS naming convention [R1]. An example of a NetCDF file containing VIS spectra and relative to the day 30/06/2022 is:

*ESA-FRM4DOAS-L1-CNR.ISAC-SAN.PIETRO.CAPOFIUME-1695-1-20220630T030848Z-20220630T192541Z-fv001.nc*

The name contains the affiliation (*CNR.ISAC*), the location of the station (*SAN.PIETRO.CAPOFIUME*), a unique identification number provided by the FRM4DOAS community (*1695*), the information of the VIS channel (*1, 2* for UV channel), the day and UTC time for the start (*20220630T030848Z*) and the end (*20220630T192541Z*) of atmospheric measurements, and the version (*fv001*).

### 4.2.2. $NO_2$ SCDs estimation

The $NO_2$ SCDs are calculated by the software QDOAS [R4]. It analyzes all the input spectra contained in the NetCDF files with respect to a reference spectrum, exploiting the DOAS technique. We chose to analyze all our data with respect to a fixed reference spectrum acquired by the SkySpec-2D in the zenith direction, during the day 11/08/2021 and with a SZA of 29.4°. More information regarding the DOAS technique can be found in [R5]. The QDOAS settings used for the analysis are reported in Table 3 and an example of retrieved NO2 SCDs is given in Fig. 7.

| | $NO_2$ VIS | Ref. Cross section |
|---|---|---|
| Calibration spectral range | 420-500 nm | |
| Retrieval spectral range | 430-490 nm | |
| Considered XS | $NO_2$ 298K | From Van Daele [R6] |
| | $NO_2$ 220K (orto. to $NO_2$ 298K) | From Van Daele [R6] |
| | $O_3$ 223K | From Bogumil [R7] |
| | $O_3$ 293K (orto. to $O_3$ 223K) | From Bogumil [R7] |
| | $O_4$ | From Hermans [R8] |
| | Ring | Computed according to [R9] |
| | $H_2O$ | From Hermans [R8] |
| | CHOCHO | From Volkamer [R10] |
| Other fits | Polynomial deg. 5 | |
| | Linear offset order 1 | |

***Table 3:*** *QDOAS settings for the retrieval of the $NO_2$ SCDs.*

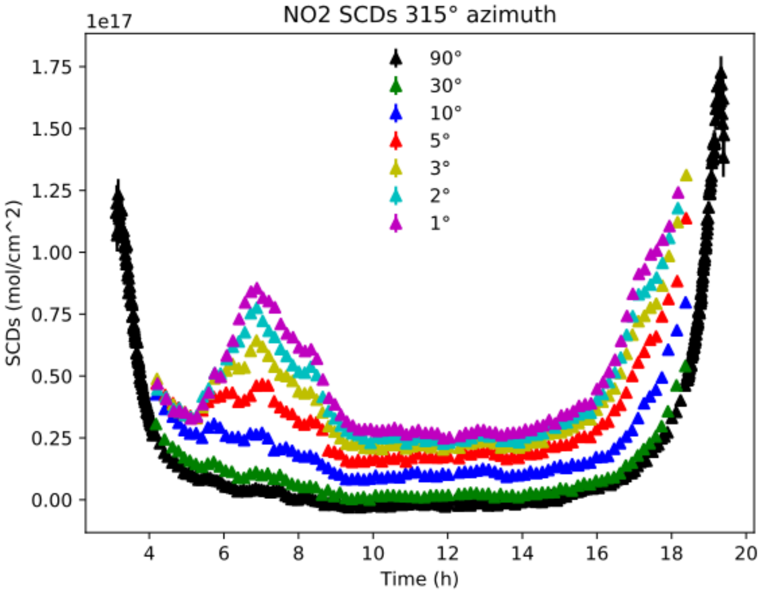


***Figure 7:*** *$NO_2$ SCDs estimated by the QDOAS software, according to the settings of Table 3, for spectra acquired during 20/06/2022, at 315° azimuth and all the elevation angles. The reported time is UTC.*

### 4.2.3. $NO_2$ VCDs from zenith-sky spectra

The $NO_2$ VCDs are computed from the $NO_2$ SCDs, retrieved only from zenith-sky measurements, according to the following equation:

$$VCD = \frac{SCD + SCDref}{AMF}$$

Where SCDref is the contribution due to the reference spectrum used for the QDOAS analysis. It must be estimated and added to the SCDs because the $NO_2$ SCD relative to the reference spectrum is not negligible. AMF is the Air Mass Factor; it must be simulated by a radiative transfer model and links the slant to the vertical columns.

#### *4.2.3.1. AMFs simulation for $NO_2$ SCDs from zenith-sky measurements*

The AMFs are dimensionless quantities. They represent the ratio between the SCDs and VCDs and are essential to interpret the SCDs fitted by the QDOAS properly. The AMFs, relative to zenith-sky $NO_2$ SCDs, depend on the input profiles, mainly the trace gas we are interested in ($NO_2$) and aerosol, the SZA and the surface albedo.
We simulated them with the SCIATRAN code [R11]. The simulations are performed for several values of SZAs with input standard profiles included in the SCIATRAN code that

account for monthly and latitudinal variations, no aerosol content and a surface albedo of 0.3. The input $NO_2$ vertical profile is the parameter that more significantly affects the accuracy of the simulated AMFs. Errors higher than 100% can occur in high polluted conditions since our input profiles are more representative of the stratosphere and contain low $NO_2$ concentration in the troposphere. However, we verified that our simulated AMFs are very close to the Network for the Detection of Atmospheric Composition Change (NDACC) standard ones.

#### *4.2.3.2. Estimate of the reference spectrum contribution*

From the equation defined in section 4.2.3, it's easy to see that, if the VCD remains constant, the SCD, as a function of the AMF, is a straight line with the VCD as line slope and SCDref the intercept. This means that the SCDref can be estimated as the inverse of the intercept of the line that fits the retrieved SCDs, relative to a constant VCD, as a function of the simulated AMFs; this is called the Langley plot. However, the assumption of a constant $NO_2$ VCD is difficult to satisfy. For this reason, we decided to create a Langley plot with one month of data, dividing all the data into bins of AMFs 0.1-wide and considering the lowest SCDs in each bin for the Langley plot. This method assumes that the lowest SCD values are representative of low tropospheric $NO_2$ concentrations and that the relative VCDs can be considered constant. Indeed, stratospheric concentrations are usually affected by lower temporal variability than tropospheric ones. As a result of this method, the reference content due to the reference spectrum used for QDOAS analysis, for SkySpec-2D measurements in SPC, is 9.7 x $10^{15}$ molec/cm$^2$, as shown in Fig. 8.

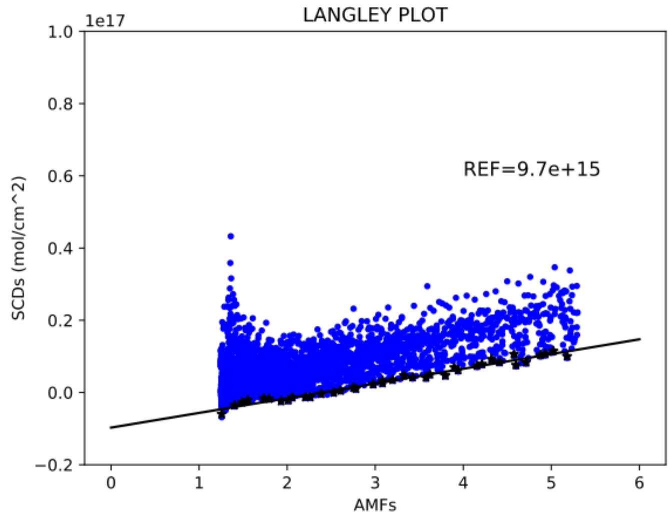


***Figure 8:*** *Langley plot used to estimate the reference contribution in the $NO_2$ SCDs for measurements in SPC. Blue dots are the $NO_2$ SCDs fitted by the QDOAS with respect to the simulated AMFs.*

#### 4.2.3.3. Cloud and aerosol filtering

As written in Sect. 4.2.3.1, we simulate the AMFs considering an atmosphere with no amount of aerosol. Furthermore, the physical processes considered in those simulations are relative to clear-sky conditions, hence with no presence of clouds.
However, aerosols and clouds can heavily affect the scattering processes in the atmosphere, leading to high variations in the SCDs and AMFs. For this reason, we need to filter out all these cloudy measurements to avoid the retrieval of $NO_2$ VCDs affected by high errors. To do that, we exploit the information from the $O_4$ SCDs that we fit in our analysis (see Table 3), as $O_4$ presents absorption lines in the chosen spectral window. The idea is that $O_4$, an oxygen dimer, is well distributed and shows low spatial and temporal variability. As a consequence, high deviations of $O_4$ SCDs are linked to different atmospheric paths crossed in the atmosphere due to scattering processes.
We adopted a filtering mechanism based on an iterative process that exploits the information from the $O_4$ SCDs and $NO_2$ VCDs retrieved in the current iteration. This type of filtering is applied to zenith measurements only. The first step consists of dividing all the $O_4$ SCDs into 3°-wide SZA bins and computing the $O_4$ SCD median value for each bin. Since most of the days at SPC are clear, we assume that all these median values correspond to clear-sky conditions. At this point, to classify clear and cloudy data, we need to define, for each SZA bin, a maximum distance from the $O_4$ SCD median value. In this way, data having $O_4$ SCDs which fall outside the chosen range are filtered out. This threshold is estimated through an iterative process. In the first step, the criterium is very stringent, leading to a low number of clear-sky data. During every iteration, the $O_4$ range in each SZA bin is increased, leading to more clear-sky data. In each SZA bin, the process stops when an important difference arises between the retrieved $NO_2$ VCDs labelled as clear and cloudy. In this way, we are confident in filtering out the $NO_2$ VCDs, labelled as cloudy, that are systematically biased compared to the clear-sky ones.

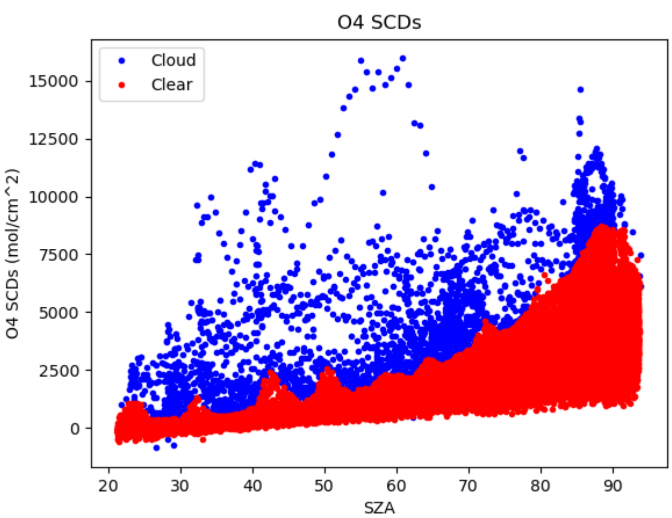


***Figure 9:*** *$O_4$ SCDs retrieved from zenith-sky measurements acquired from 1st October 2021 until nowadays. Red and blue dots refer to clear and cloudy (filtered data) conditions, respectively.*

### 4.2.4. $NO_2$ vertical profiles from MAX-DOAS measurements

The processing chain for the retrieval of aerosol extinction at 477 nm and $NO_2$ profiles starts from the $O_4$ and $NO_2$ SCDs determination described in Sect. 3.2.2. The unique difference with respect to the zenith sky processing is that the reference spectrum used for SCDs determination is not even longer the fixed one but is the one measured at the zenith sky for each elevation scan sequence. This results in SCDs being differential SCDs with respect to their corresponding zenith sky measurement. The $O_4$ and $NO_2$ SCDs together with their retrieval error at 1°, 2°, 3°, 5°, 10° and 30° are the input of the DEAP inversion algorithm.

The DEAP algorithm is a two-step code based on OE and the SCIATRAN forward model. A full description of the code and its validation using synthetic SCDs from the FRM4DOAS network, as well as comparisons with state-of-the-art codes on real data, are reported in [R15]. In the first step, the aerosol extinction profile for each measured scan is retrieved iteratively. Then, in the second step, this profile is used to compute (using SCIATRAN) the $NO_2$ box-AMF that is then used to obtain the $NO_2$ profile.

The outputs of the code for each scan consist of the following ascii files:

1. res_NO2_YYMMDD_scan_X_real.dat
2. res_O4_YYMMDD_scan_X_real.dat
3. dscdNO2_YYMMDD_scan_X_real.dat
4. scd_hyst_YYMMDD_scan_X_real.dat
5. akm_NO2_YYMMDD_scan_X_real.dat
6. akm_AER_YYMMDD_scan_X_real.dat

where YY is the year (“23” for “2023”), MM the month, DD the day and X the elevation of the scan.

The “res_…” files contain the retrieval results for aerosol and NO2 profiles together with their retrieval error and a-priori estimates and errors. The “akm_...” files contain the AKM. The “dscdNO2_...” and “scd_hyst_...” the information on modelled/measured SCDs, the iteration history and Chi-Square in case of aerosol extinction retrieval.

Once a full day has been processed, two additional ASCII files containing the total column values and the values at the surface for the whole day are produced. In detail:

1. YYMMDD_NO2_surf_tropo_DEAP.dat contains the time of the observation (decimal hours), the NO2 value at the surface (molecules/cm3), the NO2 Tropospheric VCD (molecules/cm2) and its error, the azimuth (deg) of the elevation scan and the cloud flag (0 for clear sky and 1 for cloudy sky).
2. YYMMDD_AER_surf_tropo_DEAP.dat contains the time of the observation (decimal hours), the value at the surface of the aerosol extinction at 477 nm ($km^{-1}$), the AOD and its error, the azimuth (deg) of the elevation scan and the cloud flag (0 for clear sky and 1 for cloudy sky).

These files were used for the comparisons reported in [R16] and [R17]. The SCDs are computed using the QDOAS software, while DEAP is written in Fortran. All the linking parts that compose the automatic retrieval chains are in python. The cloud filtering for retrieval profile processing follows a different approach with respect to the zenith sky ones: it is a colour index approach, as reported in [R16].

## 4.3. Outputs

In this section we describe the final outputs obtained from the SkySpec-2D spectra at SPC.

### 4.3.1. VIS and UV calibrated spectra

The main output of the processing chain is the VIS and UV calibrated spectra. This output is written in NetCDF files compliant with the FRM4DOAS requirements. The structure of the NetCDF files also follows the FRM4DOAS guidelines reported in [R3]. More details about these files are in Sect. 4.2.1. Basically, each file contains all the atmospheric spectra (one for VIS channel and one for UV channel) and the related information relative to one day of measurements. An example of the structure of the NetCDF file is reported in Figure 10.

| Name | Long Name | Type |
|---|---|---|
| ESA-FRM4DOAS-L1-CNR.ISAC-SAN... | Level 1 data for FRM4DOAS MAXDOAS proces... | Local File |
| INSTRUMENT_LOCATION | INSTRUMENT_LOCATION | — |
| altitude | altitude of the instrument above sea level | — |
| altitude_of_station | altitude of the station above sea level | — |
| latitude | latitude of the instrument (positive north) | — |
| longitude | longitude of the instrument (positive east) | — |
| RADIANCE | RADIANCE | — |
| GEODATA | RADIANCE/GEODATA | — |
| solar_azimuth_angle | solar azimuth angle, 0..360, measured towar... | 1D |
| solar_zenith_angle | solar zenith angle | 1D |
| viewing_azimuth_angle | viewing azimuth angle 0..360, measured tow... | 1D |
| viewing_elevation_angle | viewing elevation angle | 1D |
| OBSERVATIONS | RADIANCE/OBSERVATIONS | — |
| datetime | measurement date and time (UT YYYY,MM,DD... | 2D |
| datetime_end | end date and time (UT YYYY,MM,DD,hh,mm,s... | 2D |
| datetime_start | start date and time (UT YYYY,MM,DD,hh,mm,s... | 2D |
| exposure_time | exposure time | 1D |
| measurement_type | measurement type : 0-invalid, 1-offaxis, 2-di... | 1D |
| number_of_coadded_spectra | number of co-added spectra | 1D |
| radiance | sum of co-added spectra (count number) | 2D |
| radiance_quality_flag | pixel quality flags (currently 0 or 1 but proba... | 2D |
| total_acquisition_time | total acquisition time (the total time the detect... | 1D |
| total_measurement_time | total measurement time (should be the time d... | 1D |
| wavelength | wavelength grid (in air) | 2D |

***Figure 10:*** *Structure of the NetCDF file containing the calibrated spectra and other auxiliary information.*

## 4.3.2. $NO_2$ VCDs from zenith-sky spectra

The analysis method described in section 4.2.3, together with the QDOAS fit (section 4.2.2) and conversion of spectra into NetCDF files (section 4.2.1), is completely automated and is used to retrieve $NO_2$ VCDs exploiting zenith-sky spectra in the VIS channel. Fig. 11 shows an example of the diurnal variations of $NO_2$ VCDs over SPC for one day (03/07/2022).

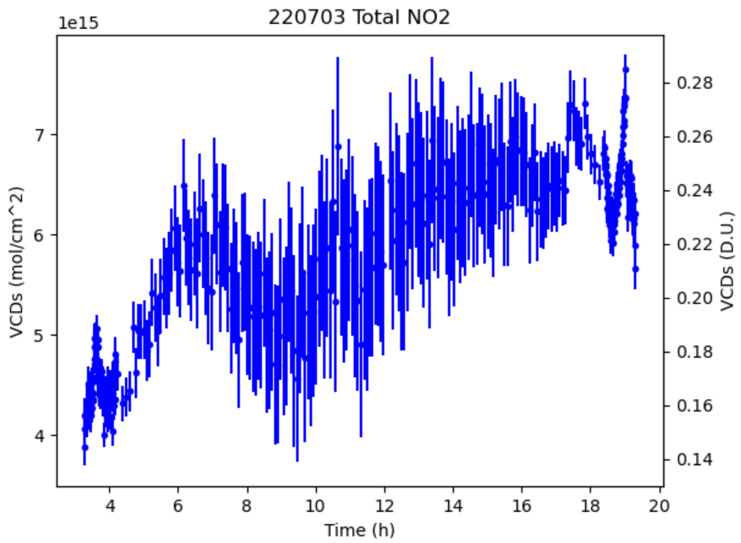


***Figure 11:*** *$NO_2$ VCDs retrieved from zenith-sky spectra during 03/07/2022*.

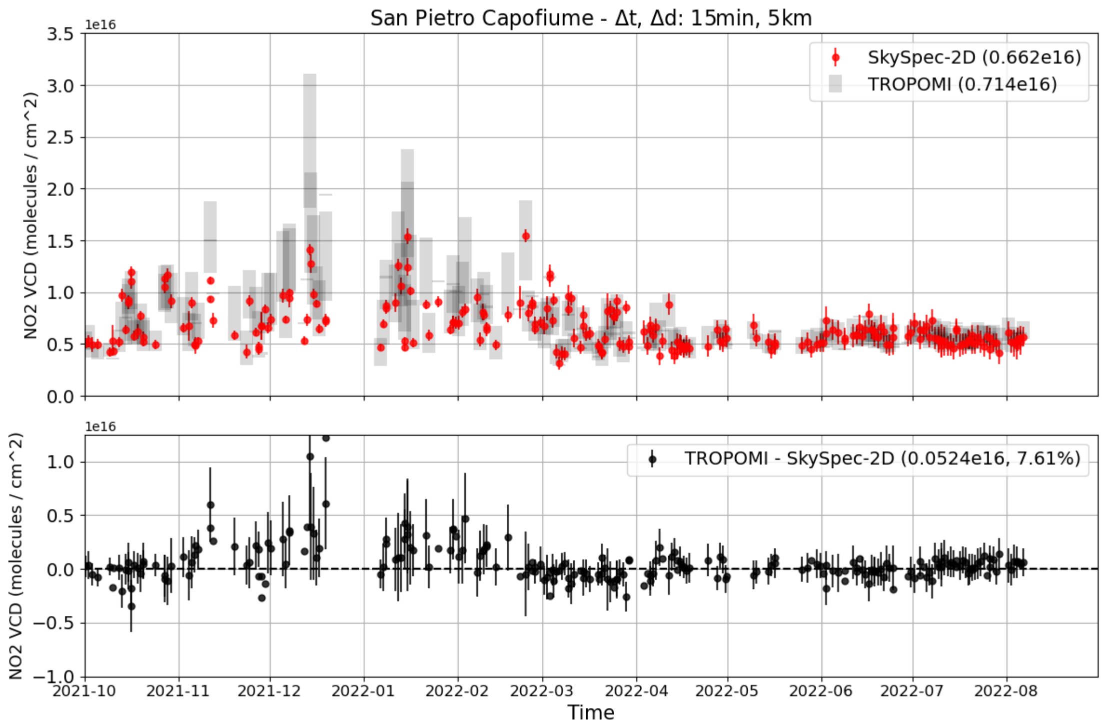


***Figure 12:*** *$NO_2$ VCDs retrieved from zenith-sky VIS spectra acquired by SkySpec-2D in SPC (red dots). They are $NO_2$ VCDs averaged in a 15 minutes time range around the*

*TROPOMI overpass. TROPOMI data (grey shadows) represent the $NO_2$ VCDs averaged within a 5 km radius around the SkySpec-2D position.*

In Fig. 12, we show the retrieved $NO_2$ VCDs timeseries. We can also observe a preliminary inter-comparison against Sentinel-5P TROPOMI $NO_2$ total (tropospheric plus stratospheric) column. SkySpec-2D $NO_2$ VCDs are averaged in a 15 minutes time range around the S-5P satellite overpass and TROPOMI $NO_2$ products are averaged within a 5 km radius around the SkySpec-2D position.

As reported in [R12], [R13], and [R14], the performance of SkySpec-2D and the quality of the $NO_2$ VCDs retrieved from zenith-sky spectra were assessed during two measurement campaigns in which the instrument took part: the first one in August 2021 in Bologna, where TROPOGAS, a research-grade custom-built MAX-DOAS instrument is located, and, the second one in September 2021 at the BAQUNIN facility at La Sapienza University (Rome) near the Pandora#117 instrument. During both campaigns, $NO_2$ VCDs retrieved from SkySpec-2D measurements revealed good agreement against TROPOGAS, Pandora#117 and other satellite (TROPOMI and OMI) products.

From October 1, 2021, the SkySpec-2D was permanently moved to SPC, where it has been continuously acquiring zenith and off-axis diffuse solar spectra. As shown in [R12], the quality of the retrieved products is regularly assessed against S-5P TROPOMI and OMI satellite products.

## 4.4. DOAS-BO web site

In order to provide information on the activities performed in the frame of the IDEAS-QA4EO project WP 2250-2251 “DOAS-BO: Towards a new FRM4DOAS site in the Po valley", we set up a prototype website (https://doas.isac.cnr.it/). The website provides a description of the MAX-DOAS system installed at the San Pietro Capofiume site (https://doas.isac.cnr.it/instruments/), information on the site itself (https://doas.isac.cnr.it/), and the main outcomes (papers/presentations/posters) of the WPs (https://doas.isac.cnr.it/documents/).

Together with this general information, a live monitoring section (https://doas.isac.cnr.it/live-monitoring/) has been set up to provide a quick outlook on the last available NO2 DEAP retrievals performed exploiting the MAX-DOAS system observations. Here, we reported some examples of those figures in this website section.

Fig.s 13 - 14, representing $NO_2$ tropospheric VCDs and $NO_2$ vertical profiles, respectively, will be updated (if available) daily. Indeed, the results of the DEAP retrieval exploiting the SkySpec-2D measurements of the day before will be uploaded every day. Instead, Fig. 15 will be uploaded less frequently (about weekly) since it includes an external dependence on S-5P/TROPOMI overpasses data, which need to be downloaded and extracted before being analysed.

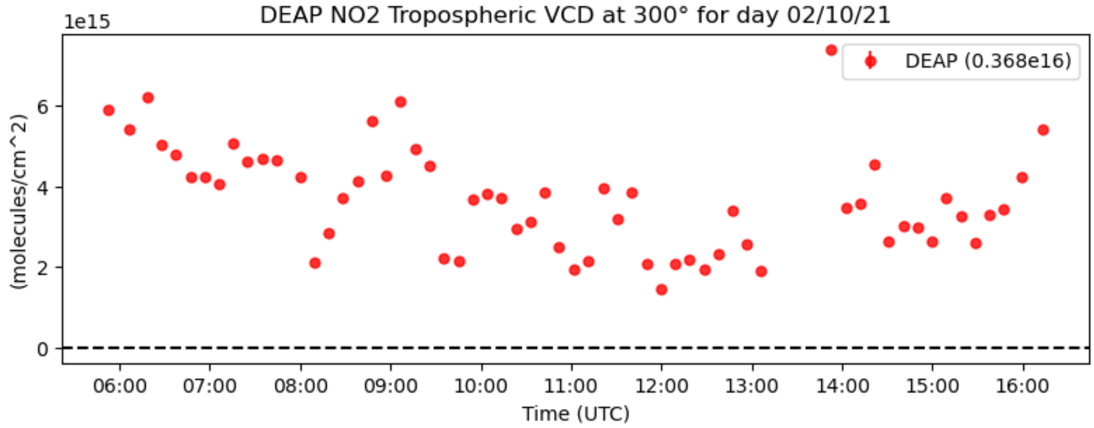


***Figure 13:*** *Example of* $NO_2$ *tropospheric VCDs (molecules/cm2) retrieved with DEAP from MAX-DOAS measurements for the azimuth (deg) angle of the elevation scan and for the day which are both reported in the title of the plot.*

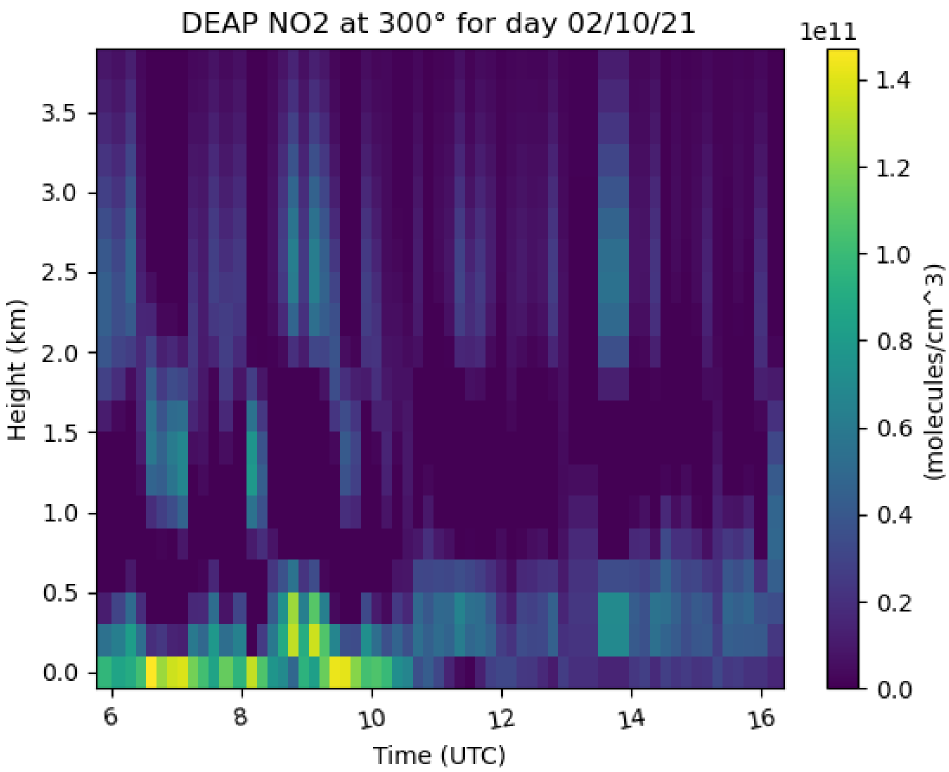


***Figure 14:*** *Example of* $NO_2$ *vertical distributions (molecules/cm3) retrieved with DEAP from MAX-DOAS measurements for the azimuth (deg) angle of the elevation scan and for the day which are both reported in the title of the plot.*

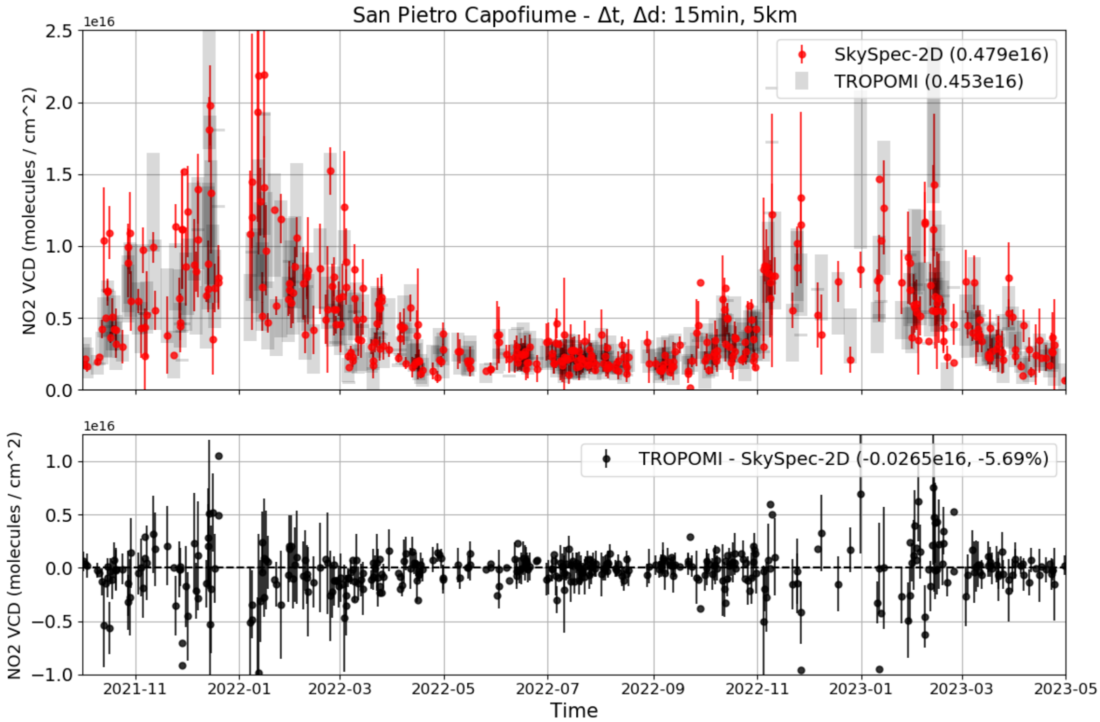


***Figure 15:*** *Example of $NO_2$ VCDs retrieved from zenith-sky VIS spectra acquired by SkySpec-2D in SPC (red dots). They are $NO_2$ VCDs averaged in a 15 minutes time range around the TROPOMI overpass. TROPOMI data (grey shadows) represent the $NO_2$ VCDs averaged within a 5 km radius around the SkySpec-2D position.*

## 4.5. Data dissemination

The outputs derived from SkySpec-2D analysis chain, such as the VIS and UV calibrated spectra (Sect. 4.3.1), and the $NO_2$ VCDs from zenith-sky spectra (Sect. 4.3.2) are routinely produced and stored within the CNR-ISAC servers. All the data are available upon request to the PI of the SkySpec-2D MAX-DOAS instrument, PhD Elisa Castelli (e.castelli@isac.cnr.it).

The list of the products available will be updated as soon as the retrieval code for the exploitation of the MAX-DOAS observation is available, and the $NO_2$ and aerosol profiles will be included. This document will be consequently updated, including details on the updated processing chain and the derived products.

Since March 2022, we also routinely providing the VIS and UV SkySpec-2D calibrated spectra in NetCDF files to the FRM4DOAS community (https://frm4doas.aeronomie.be). At the moment, our spectra are under a testing phase. If the testing phase will be passed, the spectra will also be analyzed through the FRM4DOAS network central processing. In that case, once processed, the products (such as tropospheric $NO_2$ profiles, etc.) will be publicly available on the NDACC/Rapid Delivery repository and on the European Space Agency (ESA) - Atmospheric Validation Data Centre (EVDC) database.

## 4.6. Quality checks

### 4.6.1. Overall checks on processing chain

Since the SPC station is located in a remote area, probably the most challenging task is to guarantee the continuity of the remote connection and the electric power. The electric power's stability is attained through battery packs, which are recharged with the solar panels during the daytime and the electric line during the nighttime. Currently, the remote connection is guaranteed through a mobile 4G router. The request for connecting the entire area to the GARR ultra-broadband network has been submitted.

The MAX-DOAS measurements are stored on a laptop connected to the instrument at the station and then synchronized routinely (every 5 minutes) on a local machine at the CNR-ISAC in Bologna.

To ensure that the instrument is measuring and the connection is alive, an automatic script checks, from sunrise to dusk, with a defined time frame (e.g. 2 hours), if the measurements are acquired and present in the local machine. In case of expecting data missing, the script sends an alert email, also indicating the time passed since the last acquisition. The laptop is equipped with remote control software (i.e., AnyDesk) in order to be reached from external devices so that the user may take action in case of problems or modify the instrument settings from a remote.

### 4.6.2. Specific checks on spectra processing

During the first step of the processing chain, described in Sect. 4.2.1, all the acquired atmospheric spectra are converted into NetCDF files. During this procedure, the conversion tool checks, for every day of measurements, that all the required calibration measurements (see Sect. 4.1), acquired during nighttime, are present. If one or more calibration spectra are missing due to acquisition problems, the atmospheric spectra, relative to that day, are not automatically converted. In that case, we will evaluate whether performing the conversion manually, using the calibration spectra relative to another day, or not.

Other filters are then applied to the estimated SCDs in order to avoid problems due to the QDOAS analysis. In particular, all the SCDs derived from analyzes labelled as "failed", for those QDOAS provides an appropriate flag, and analyzes with final $\chi^2$ higher than $10^{-4}$ are filtered out.

The last filter is then applied during the conversion of $NO_2$ SCDs retrieved from zenith-sky spectra into $NO_2$ VCDs with the purpose of filter out the cloudy data. The filtering method is explained in Sect. 4.2.3.3.

In case of NO2 SCDs used for profile retrievals, the cloud filtering adopts a cloud index approach on spectra as reported in [RD16].

## 5. References

[R1] FRM4DOAS Instrument Operation and Calibration Guidelines, Date: 16/03/2018 Version: 1.1, ESA Contract No. 4000118181/16/I-EF

[R2] MAXDOAS Calibration and Operations Best Practices, Date: 10/01/2018 Version: 1.0, ESA Contract No. 4000118181/16/I-EF

[R3] FRM4DOAS Level-1 data, Date: 16/03/2018 Version: 1.1, ESA Contract No. 4000118181/16/I-EF

[R4] Danckaert T., Fayt C., Van Roozendael M., De Smedt I., Letocart V., Merlaud A., Pinardi G., QDOAS Software user manual, Royal Belgian Institute for Space Aeronomy.

[R5] Ulrich Platt, Jochen Stutz, 2008, Differential Optical Absorption Spectroscopy, Springer.

[R6] Vandaele, A.C.; Hermans, C.; Simon, P.C.; Carleer, M.; Colin, R.; Fally, S.; Merienne, M.F.; Jenouvrier, A.; Coquart, B. Measurements of the $NO_2$ absorption cross-section from 42 000 cm- 1 to 10 000 cm- 1 (238–1000 nm) at 220 K and 294 K. Journal of Quantitative Spectroscopy and Radiative Transfer 1998, 59, 171–184.

[R7] Bogumil, K.; Orphal, J.; Homann, T.; Voigt, S.; Spietz, P.; Fleischmann, O.; Vogel, A.; Hartmann, M.; Kromminga, H.; Bovensmann, H.; others. 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. Journal of Photochemistry and Photobiology A: Chemistry 2003, 157, 167–184.

[R8] Hermans, C.; Vandaele, A.C.; Carleer, M.; Fally, S.; Colin, R.; Jenouvrier, A.; Coquart, B.; Mérienne, M.F. Absorption cross-sections of atmospheric constituents: $NO_2$, O2, and H2O. Environmental Science and Pollution Research 1999, 6, 151–158.

[R9] Wagner, T. Satellite observations of atmospheric halogen oxides; Mensch & Buch, 1999.

[R10] Volkamer, R.; Spietz, P.; Burrows, J.; Platt, U. High-resolution absorption cross-section of glyoxal in the UV–vis and IR spectral ranges. Journal of Photochemistry and Photobiology A: Chemistry 2005, 172, 35–46.

[R11] Rozanov, V.; Buchwitz, M.; Eichmann, K.; de Beek, R.; Burrows, J. SCIATRAN - A new radiative transfer model for geophysical applications in the 240-2400 nm spectral region: The pseudo-spherical version. Remote Sensing of Trace Constituents in the Lower Stratosphere, Troposphere and the Earth's Surface: Global Observations, Air Pollution and the Atmospheric Correction; Burrows, J.; Takeucki, N., Eds. Comm Space Res, 2002, Vol. 29, Advances in Space Research-Series, pp. 1831–1835. A1 2 Symposium of COSPAR Scientific Commission A held at the 33rd COSPAR Scientific Assembly, WARSAW, POLAND, JUL, 2000, doi:10.1016/S0273-1177(02)00095-9.

[R12] Pettinari, P.; Castelli, E.; Papandrea, E.; Busetto, M.; Valeri, M.; Dinelli, B.M. Towards a New MAX-DOAS Measurement Site in the Po Valley: $NO_2$ Total VCDs. Remote Sens. 2022, 14, 3881. https://doi.org/10.3390/rs14163881.

[R13] Castelli, Elisa, Pettinari, Paolo, Papandrea, Enzo, Cristofanelli, Paolo, Busetto, Maurizio, & Valeri, Massimo. (2022). Report on measurements campaign within the ISAC-CNR (Bologna) site exploiting in-situ and the satellite-borne synergies. Zenodo. https://doi.org/10.5281/zenodo.5886896.

[R14] Castelli, Elisa, Pettinari, Paolo, Papandrea, Enzo, Cristofanelli, Paolo, Busetto, Maurizio, Di Liberto, Luca, & Valeri, Massimo. (2022). Report on the inter-comparison campaign within the BAQUNIN supersite in Rome. Zenodo. https://doi.org/10.5281/zenodo.5886950.

[R15] Castelli, Elisa, Pettinari, Paolo, Papandrea, Enzo, Premuda, Margherita, Di Liberto, Luca, & Valeri, Massimo. (2023). Report on the description of the MAX-DOAS profiles retrieval code and its validation. Zenodo. https://doi.org/10.5281/zenodo.8186938.

[R16] Valeri, Massimo, Castelli, Elisa, Pettinari, Paolo, Papandrea, Enzo, Di Liberto, Luca, Marinoni, Angela, & Decesari, Stefano. (2023). Report on the inter-comparison results between ground-based and satellite measurements. Zenodo. https://doi.org/10.5281/zenodo.8186954.

[R17] Castelli, Elisa, Pettinari, Paolo, Papandrea, Enzo, & Valeri, Massimo. (2023). Report on the inter-comparison results between MAX-DOAS Tropospheric VCDs and S-5P TROPOMI "corrected" NO2 Tropospheric VCDs. Zenodo. https://doi.org/10.5281/zenodo.8214537.