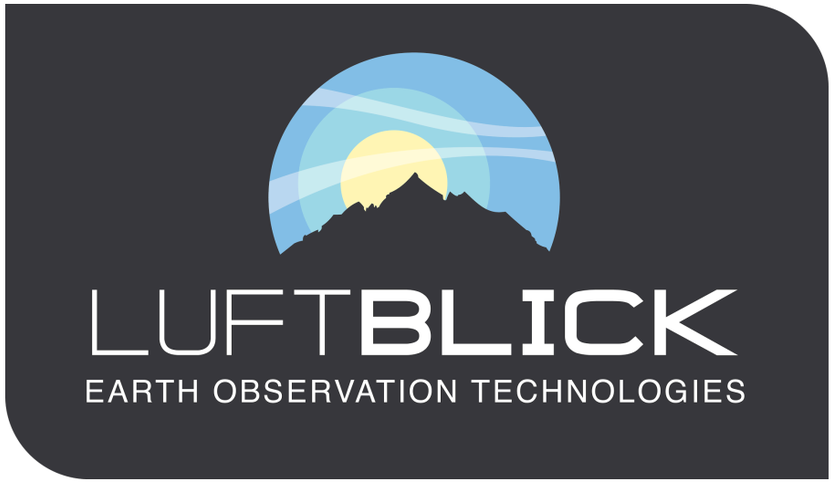


LuftBlick Report 2025004

Activity 2 Report

on L2 Retrieval Improvement and Validation Tasks

Version 3.0 - 28 Apr 2026

Prepared by

Name	Institution
Martin Tiefengraber	LuftBlick[1]
Manuel Gebetsberger	LuftBlick[1]
Markus Kilian	LuftBlick[1]
Axel Kreuter	LuftBlick[1]

[1] *LuftBlick Earth Observation Technologies, Innsbruck, Austria*

Contents

1 Document Change Record

Version	Date	Section	Observations
1	16 July 2025	All	First version
2	30 Nov 2025	All	Major additions to all sections.
3	28 Apr 2026	All	Major additions to all sections. Some parts have been reorganized to Appendix sections.

2 Introduction

This Report, Version 1, is part of Deliverable (D) Group 5 under Work Package (WP) 1 - Instrument Stability and Spectral AOD. WP1 is the initial activity within the EUMETSAT "Service to Support the Pandonia Global Network for Copernicus Mission Cal/Val - Algorithm, Monitoring and Analysis" (PanAMA) framework.

Specifically, this report discusses relevant WP1 tasks derived from the comprehensive pool of identified PanAMA "L2 retrieval improvements and validation tasks."

The document is split into two main chapters: L2 retrieval improvement studies and validation studies. Each chapter reports on completed and in-progress studies, as outlined in the corresponding living-planning document (in the form of a Trello board).

2.1 Acronyms and Abbreviations

AOD	Aerosol Optical Depth
AMF	Airmass factor
BSS	Blick Software Suite
DIFF	Diffuser
DIFF_CORR	Diffuse correction factor
DOY	Day of year
FOV	Field of View
GAM	Generalized Additive Model
GAMLSS	Generalized Additive Model for Location, Scale and Shape
RedLE	Reduced Langley Extrapolation technique
RTM	Radiative Transfer Model
SC	Slant Column
SLSPAT	Spatial Straylight
SLSPEC	Spectral Straylight
SSA	Single Scattering Albedo
SZA	Solar Zenith Angle
MinAMF	Parameter AMF minimization technique
LE	Langley Extrapolation
LUT	Lookup Table
rLE	Running Langley Extrapolation
VC	Vertical Column

3 L2 Retrieval Improvement Studies

3.1 Spectral AOD Algorithm (WP1.6)

Due to the nature of aerosol extinction, accurate retrieval of aerosol optical depth (AOD) depends strongly on knowing the absolute level of the L1 signal and its temporal stability. This contrasts with common trace-gas retrievals, where algorithms can rely on complex spectral absorption structures. In such cases, the requirements on absolute L1 signal accuracy can be relaxed, while high demands remain on spectral accuracy, for example dispersion and resolution.

Corrected-count L1 level accuracy and stability are not intrinsic strengths of the Pandora instrument. This is illustrated by the example multi-year comparison at the PanSel site Tel Aviv, where Pandora data are compared with the co-located Cimel AERONET instrument at three wavelengths (Figure 3.1-1). Multi-month stable periods are often interrupted by sudden jumps of up to 20 %, or by smooth half-year drifts of several percent.

As a consequence, no operational AOD product is currently available within the PGN. However, the ability to retrieve high-quality trace-gas products allows gas absorption to be constrained at L1. In principle, this makes it possible to retrieve AOD from any wavelength, commonly referred to as spectral AOD.

The primary objective of the Activity 2 tasks within this WP is therefore to apply methods that establish high spectral radiometric accuracy and maintain this accuracy over time, thereby enabling accurate AOD retrievals. We address this objective by combining two complementary methods into a tandem algorithm concept, hereafter referred to as TanAOD.

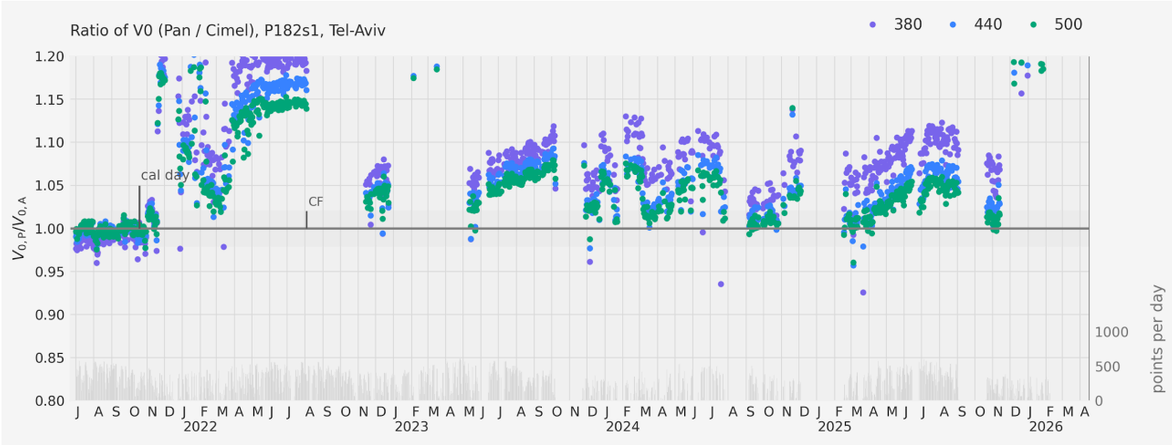


Figure 3.1-1 Calibration stability of Pandora 182s1 in Tel Aviv, retrieved from comparison with the co-located Cimel AERONET instrument.

To achieve radiometric accuracy for the first TanAOD component, TanAOD-L, an adjusted Langley extrapolation method is applied independently at each wavelength. The more stable the atmosphere is, meaning the more constant the atmospheric transmittance remains during the calibration period, the more reliable the extrapolation becomes. This stability requirement is the intrinsic limitation of the Langley extrapolation technique. In the present approach, it is relaxed by reducing variable atmospheric constituents from the spectra before extrapolation.

To maintain the accuracy gained from the Reduced Langley Extrapolation (RedLE), RedLE is applied in a smoothed, running manner to each day of a time series. Details are provided in the section “Reduced Langley method”.

Meeting the Langley stability requirements remains intrinsically limiting in both time and space, because sufficiently stable atmospheric conditions are only sparsely available during typical calibration periods, usually one day or half a day. The second TanAOD component, TanAOD-P, therefore follows an independent and novel approach: it establishes and maintains radiometric

accuracy through post-processing of DOAS-like spectral fitting closure polynomials. Details are provided in the section “Post-processing of spectral fitting closure polynomials”.

The Appendix sections provide further technical details on the TanAOD methods, as well as associated tools and supporting tasks:

- A newly developed running application of trace-gas calibration, used to ensure the highest-quality input data for gas-signature reduction within RedLE. Beyond this application, the indispensable role of such a running technique for operational processing is demonstrated in Appendix C.1, “Develop a framework for an automated rLE to the PGN”.
- Integration of AOD reference data from AERONET. To facilitate seamless integration into the workflow, tools were developed that use the AERONET data-access APIs. More details are provided in Appendix C.4, “Collect the AERONET AODs”.
- Selection criteria for the PanSel sites are described in Appendix C.2. These sites provide the baseline datasets used for all studies.
- The development of the profile-sky tropospheric AOD product and O2, O2O2 retrieval setting optimization requires radiative transfer capabilities. Potential candidates are discussed in Appendix C.3.

3.1.1 Reduced Langley Method and Running Application (TanAOD-L)

For the proposed Langley method, spectra are reduced by absorption and extinction features that alter atmospheric transmittance. These are mainly:

A. trace gases: O3, H2O, NO2, SO2, HCHO, O2O2, and O2;
B. Rayleigh scattering;
C. spectral tropospheric AOD.

The first two components relate to products retrievable from direct-sun measurements, although O2O2 and O2 are still under development. The third component, tropospheric AOD, would be a novel product retrieved from profile sky-radiance measurements. Its development, together with the associated O2 and O2O2 fitting setups, will be postponed to the follow-up WP. A key prerequisite for these developments is the ability to perform radiative-transfer calculations. Within this WP, efforts have therefore been made to identify the most suitable method; see Appendix C.3, “Research about suitable RTM”.

In the ideal case, the variability-reduced spectrum, F_{RED}, would be free from atmospheric absorption features that can vary during the selected Langley period. Considering the rearranged monochromatic Beer-Lambert-Bouguer law,

$$\ln(F_{RED}) = \ln(F_{MEAS}) + \tau_{SCA} \cdot m_{SCA} + \sum_i \tau_{GASi} \cdot m_{GASi} + \tau^*_{AER} \cdot m_{AER} =$$
$$= \ln(F_0) + \Delta\tau_{AER} \cdot m_{AER}$$

with

- F_{RED} variability-reduced direct-sun irradiance;
- F_{MEAS} measured direct sun irradiance;
- F_0 extraterrestrial direct sun irradiance;
- τ_{SCA} Slant optical depth of Rayleigh scattering;
- τ_{GAS} Slant optical depth of trace gas i;
- τ^*_{AER} Slant optical depth of tropospheric AOD;
- $\Delta\tau_{AER}$ Slant optical depth of free tropospheric and stratospheric AOD;
- m_x Airmass factors,

If tropospheric AOD is known and removed, the only remaining atmospheric constituent is $\Delta\tau_{AER}$. Constrained by atmospheric advection processes, this component is expected to remain approximately constant over half days. Therefore, half-day periods, split at the lowest solar zenith angle into AM and PM branches, are preferred Langley periods.

The slant optical depths for trace gases are calculated by scaling pre-convolved gas cross-sections with slant columns obtained directly from PGN trace-gas retrievals. Specifically, we use direct-sun

slant-column retrievals from the retrieval setups listed below. Setup descriptions are provided in the PGN data products readme.

gas	NO2	O3	SO2	H2O	HCHO	O2O2*	O2***
retr. setup	rnvs3	rout2/rous1	rsus1	rwvt1/rwvs0	rfus5	rdvs0	rxvs0

* No official data product yet; retrieval setup refinement ongoing.
** Retrievable from the VIS channel, s2, only.

To achieve the best possible trace-gas reduction across the full spectral range, approximately 300 to 900 nm, the two spectrometers must operate as synchronously as possible. More details are provided in the document LuftBlick_PanAMA-WP1_Activity1-L1&InstrumentImprovement_Report_2025003. For example, O2 can only be retrieved from the Pandora VIS channel, s2, whereas O3 must be constrained for VIS spectra but can be retrieved at highest quality from the UV channel, s1.

An example of spectral reduction is shown in Figure 3.1-2 for noon spectra measured on 23 November 2022 at Izaña, Tenerife, Spain, using Pandora 209 s1, the UV channel, and s2, the VIS channel. Two reduction levels are visible: first, reduction of Rayleigh scattering only; second, additional reduction of the trace-gas absorptions listed above. Tropospheric AOD is not yet considered because the product is still under development. The apparently stronger structure in the s1 spectrum is caused by its better, and therefore narrower, spectral resolution compared with s2.

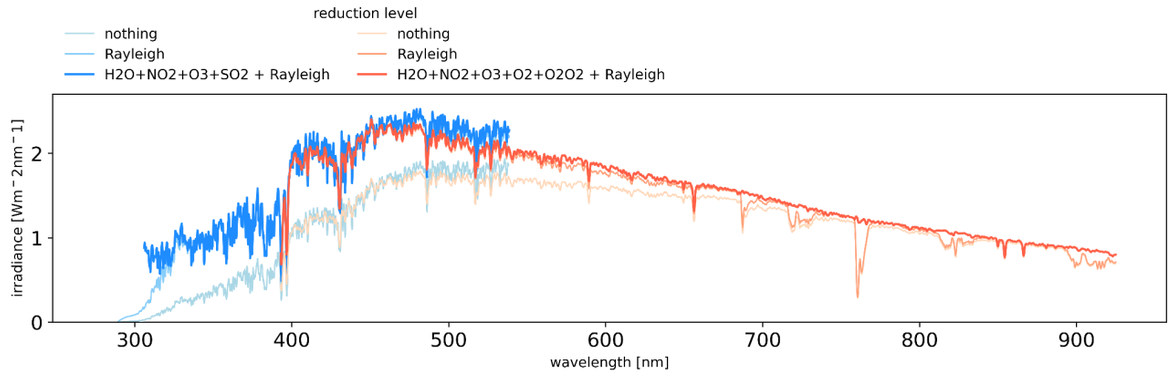


Figure 3.1-2: Variability reduction applied to noon spectra from Pandora 209 s1, shown in bluish colors, and s2, shown in reddish colors, measured on 23 November 2022.

The Langley extrapolation itself is performed by linearly fitting the logarithm of the quasi-reduced spectra, $\log F_{RE}$, against the aerosol effective airmass factor (following Kasten, 1965). Accordingly, when tropospheric AOD is not reduced, the dominant driver of atmospheric-transmittance variability remains the full AOD.

Two example applications at 500 nm are shown in Figure 3.1-3 for the background site Izaña, Tenerife, Spain, and the suburban site La Porte, Texas, USA. The relatively stable atmosphere at Izaña is reflected in the intercepts, meaning calibration constants, which agree within 1 % for a large group of Langley periods selected within one month. In contrast, the suburban site shows higher half-day AOD variability, leading to the expected larger spread of intercepts, around 15 %, for the selected Langley periods.

Although the standard Langley extrapolation is already meaningful for the remote site, we argue that considering all reduction levels, including tropospheric AOD, would also enable RedLE applications at sites with variable atmospheric transmittance during the Langley period, such as the suburban example.

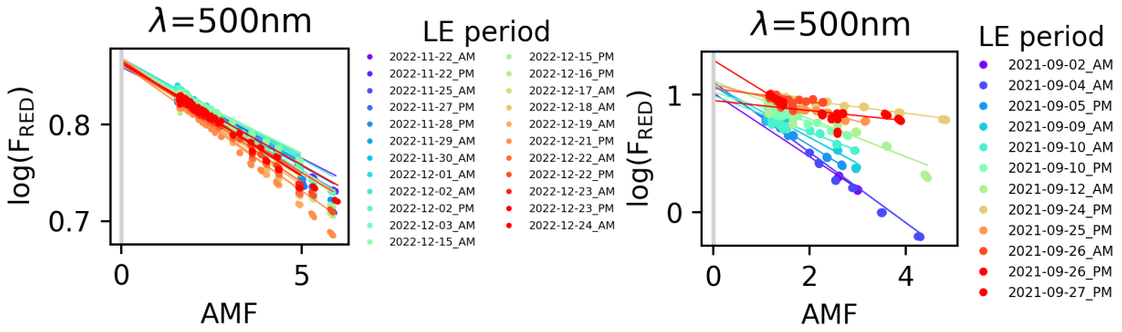


Figure 3.1-3: RedLE applied to selected Langley periods, half days, at Izaña for Pandora 209, left panel, and La Porte, Texas, USA, for Pandora 58, right panel. The dots represent the reduced irradiance at 500 nm, and the lines show the linear fits extrapolated to AMF = 0.

One consequence of a reliable Langley extrapolation is that the estimated extraterrestrial spectrum, F_0, can be used to correct radiometric sensitivity, thereby improving L1 data quality. The AOD itself is unaffected by this correction because the sensitivity cancels in the transmission ratio.

The final AOD is then calculated by solving the reduced Beer-Lambert-Bouguer law for τ and applying the Sun-Earth distance correction. Example AOD retrievals for the same remote and suburban datasets, Izaña and La Porte, are shown in Figure 3.1-4 for standard AERONET wavelengths up to 500 nm. Pandora data are shown as blue dots, and Cimel AERONET data are shown in red. The Pandora data at 340 and 380 nm were measured using a U340 bandpass filter in both systems, and the Pandora 58 data are fully corrected for spectral stray light.

The Pandora data were calibrated to AERONET in order to emphasize potential systematic differences and temporal evolution rather than calibration inconsistencies. Overall, the agreement is high. However, even over these short time scales, an AOD drift is observed for the Izaña dataset from Pandora 209, particularly at shorter wavelengths. For La Porte, Pandora 58, we likely encountered a filter-wheel issue on the second-last day, where the expected, and therefore corrected, filter was not in place during the measurement.

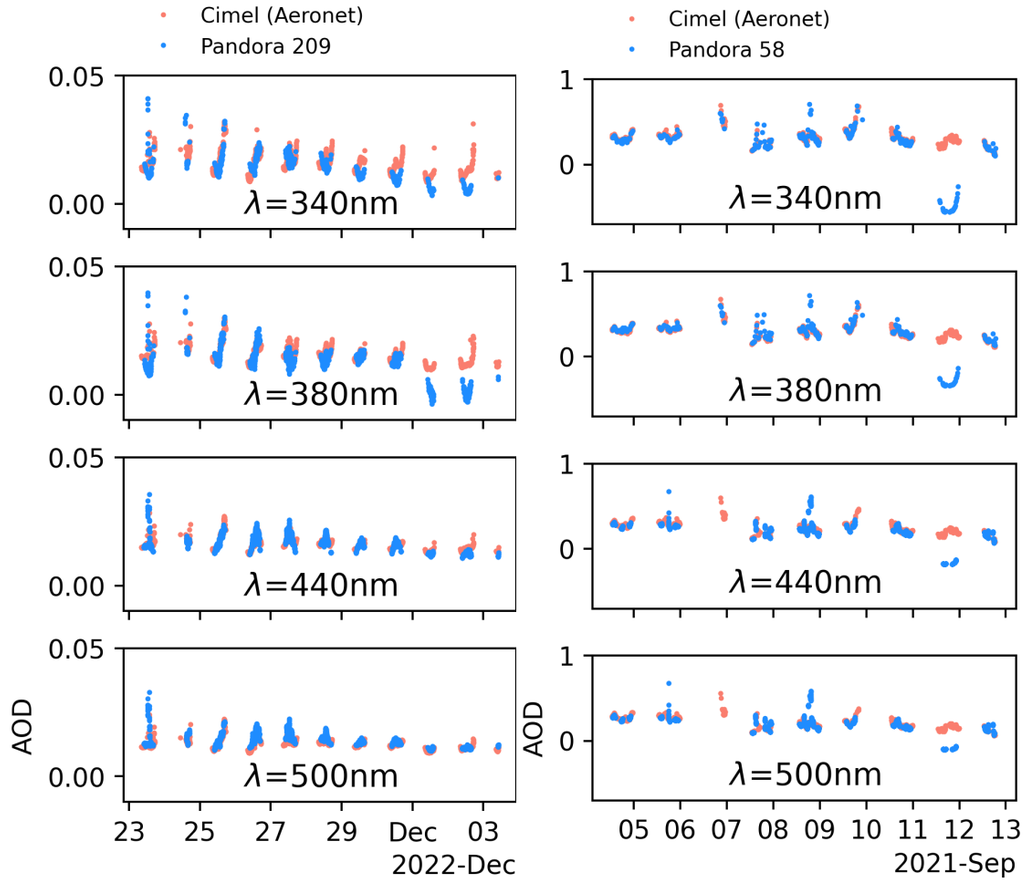


Figure 3.1-4: Example AOD at Izaña for Pandora 209, left panel, and La Porte, Texas, USA, for Pandora 58, right panel, at different AERONET wavelengths up to 500 nm. Pandora data are shown in blue and AERONET reference data in red.

We therefore conclude that even the most effective calibration method cannot ensure long-term Pandora AOD stability if applied only once. For this reason, RedLE is used in a running mode, TanAOD-L. In this context, F_0 and I_0 are treated synonymously.

The running procedure is:

1. RedLE is applied to each half day of a time series, split around solar noon, yielding two calibration constants, I_0, per day, one AM and one PM.
2. Each half-day extrapolated I_0 is either used directly, downweighted, or rejected.
3. The selected and weighted I_0 values are then smoothed in a running manner.

More technical details on the selection process are provided in Appendix A.

3.1.1.1 Test Application of Field Data

An example application of this selection process, steps 1 and 2, is illustrated for the Thessaloniki dataset at 500 nm in Figure 3.1-5. The resulting I0 time series is shown in the first panel of Figure 3.1-6. The color-coded dots represent AM/PM-combined daily Langley points, with color indicating fitting quality. After weighting and smoothing, the final running I0 is shown as a green line. For comparison, the same running I_0 time series is also shown for Tel Aviv, middle panel, and Izaña, bottom panel.

The need for this elaborate approach becomes evident from the strong variability in individual I_0 values, driven by unavoidably unstable fitting conditions at polluted sites such as Thessaloniki. A further challenge is the frequent occurrence of sudden changes in instrument sensitivity, reflected by I_0 jumps of several tens of percent.

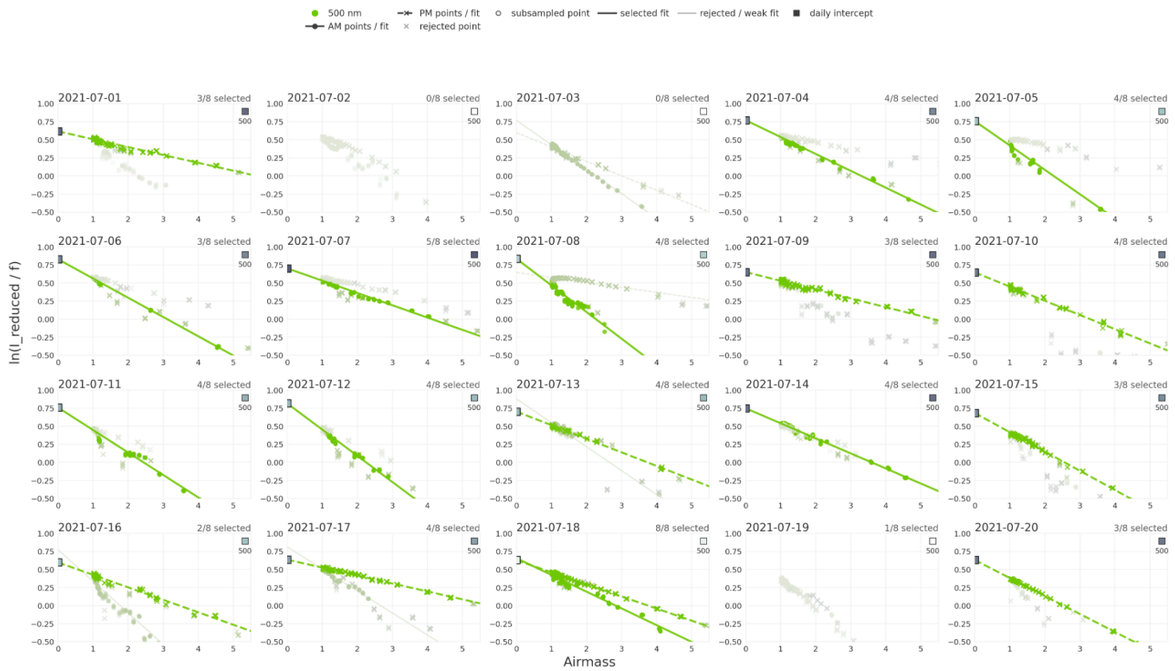


Figure 3.1-5: Example running RedLE application at 500 nm for Pandora 240s1 in Thessaloniki. Half-day intercepts are shown in blue, including uncertainty bars, and the red curve shows the result after smoothing and outlier filtering.

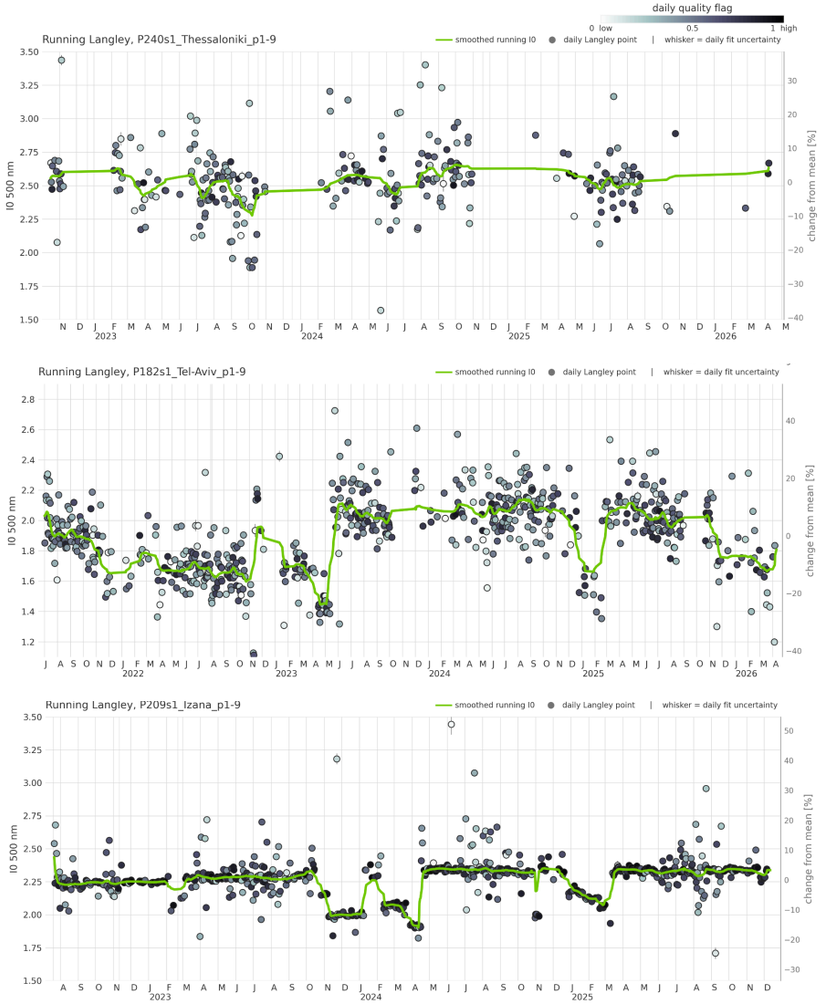


Figure 3.1-6: Example applications of the TanAOD-L I_0 selection and smoothing procedure at 500 nm for Thessaloniki, top, Tel Aviv, middle, and Izaña, bottom. Dot color indicates fitting quality; the green line shows the final selected I_0.

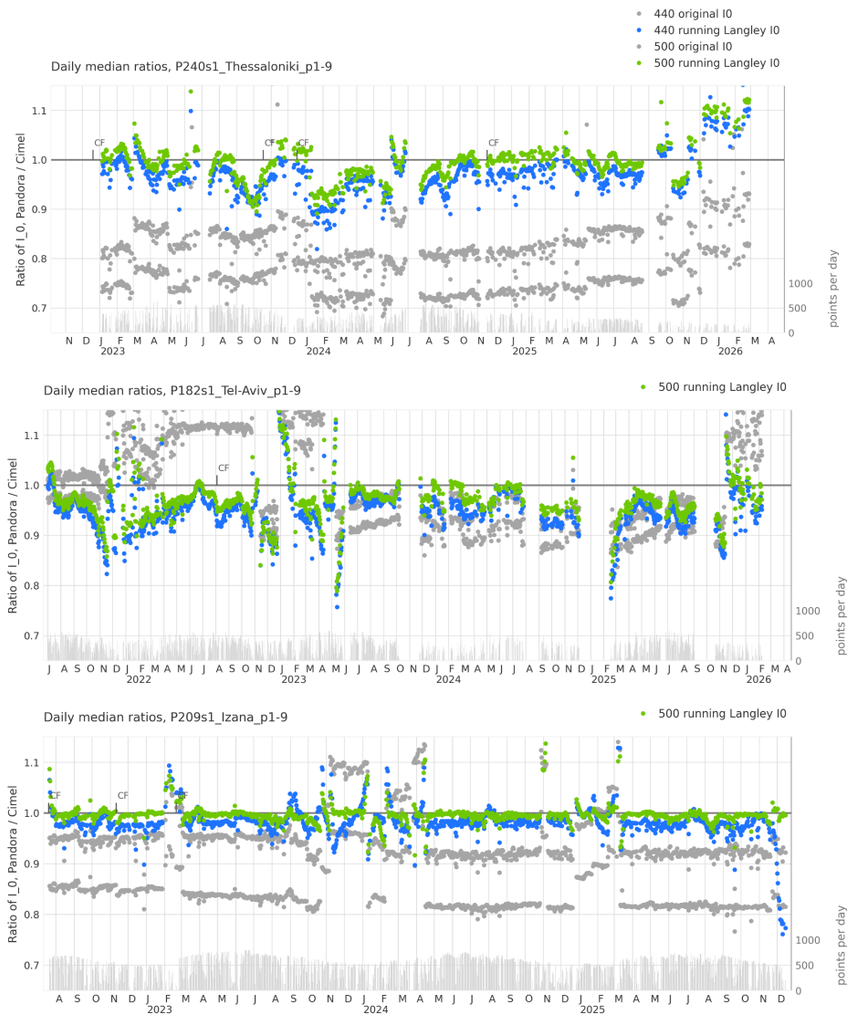


Figure 3.1-7: Example application of TanAOD-L results at 500 nm for Thessaloniki (top), Tel-Aviv (middle) and Izana. The color coding refers to different wavelengths. Untreated (original) data are shown in gray.

Following Beer's law, the relative difference between extrapolated constants I_0 from two sources can be related to the airmass-factor-corrected absolute AOD difference: $\frac{I_{0-P}}{I_{0-A}} = e^{(\tau_P - \tau_A)\cdot m}$,

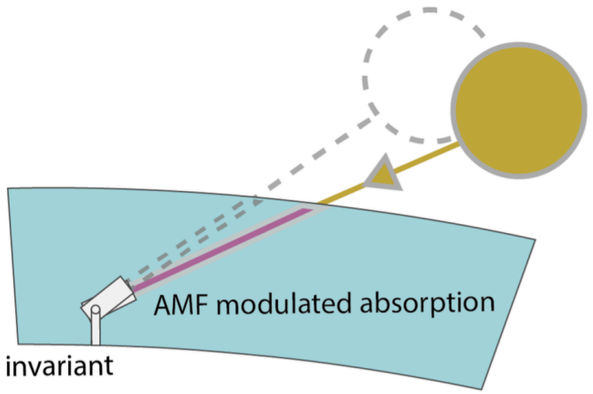
where P denotes Pandora, A denotes AERONET (Cimel instrument), τ is AOD, and m is the airmass factor.

Comparing Pandora-derived I_0 values with AERONET reference data for the three example sites reveals calibration biases and sensitivity jumps of up to 30 %, for example at 440 nm. This is visible in Figure 3.1-7 as gray dots.

Applying the TanAOD-L concept, i.e. running RedLE, largely removes systematic biases, including at polluted sites. The corrected values are shown by wavelength in the same figure. However, because of the running nature of the method, additional fluctuations can be introduced when multiple instrumental states occur on shorter time scales.

The stable atmosphere over Izaña enables a satisfactory overall accuracy of about 2 %. In a broader operational context, a realistic accuracy estimate is closer to 5 %.

3.1.2 Post-Processing of Spectral Fitting Closure Polynomials (TanAOD-P)



The fundamental paradigm of the second algorithm component is that atmospheric absorption recorded by a ground-based direct-sun instrument is modulated by the changing airmass over the day. Instrumental features that bias AOD, such as changes in radiometric throughput, are not modulated in the same way.

In principle, this concept allows any AOD retrieval to be post-processed by separating the instrumental component according to:

$$\tau_{AER}(\lambda, t) = \tau a_{AER}(\lambda, t) \cdot m_{AER}(t) + G_{INST}(\lambda, t)$$

with

- G_{INST} Instrumental component;
- τ_{AER} Slant optical depth of total AOD;
- τa_{AER} Vertical optical depth of total AOD;
- m_{AER} Airmass factors.

The essential physical assumption is that, within one day, the structured instrumental baseline is not multiplied by the aerosol airmass factor. This provides a time-domain criterion for separability.

For TanAOD-P, this paradigm is connected to DOAS spectral fitting. The fitting accounts for narrow-band absorbers, gases, and Rayleigh scattering, but also includes closure polynomials that compensate instrumental mismatches between measurement and reference. In particular, the smoothing polynomial, P_{SMO}, represents a broadband quantity that intrinsically includes τ_{AER}.

The schematics is shown below, starting from a simplified form of the spectral fitting equation:

$$\ln\left[(I_0/cal_0)(\lambda_j,t_i)\right]-\ln\left[(I/cal)(\lambda_j,t_i)\right]-\tau_{\mathrm{SFIX}}(\lambda_j,t_i)-\sum_{k=1}^{n_{\mathrm{GAS}}}\tau_{Sk}(\lambda_j,t_i)=P_{\mathrm{SMO}}(\lambda_j,t_i)$$

measurements — removed ODs (like Rayleigh) — fitted ODs — fitted broadband

stable instrument: $cal_0 = cal$

$$P_{\mathrm{SMO}}(\lambda_j,t_i)=m(t_i)\cdot\tau_{\mathrm{AER}}(\lambda_j,t_i)$$

standard instrument: $cal_0 \neq cal$

$$P_{\mathrm{SMO}}(\lambda_j,t_i)\approx m(t_i)\cdot\tau_{\mathrm{AER}}(\lambda_j,t_i)+\Delta\tau_{\mathrm{INST}}(\lambda_j,t_i)$$

$$\Delta\tau_{\mathrm{INST}}(\lambda_j,t_i)\approx G_{\mathrm{INST}}(\lambda_j)$$ instrument change constant over day

$$P_{\mathrm{SMO}}(\lambda_j,t_i)\approx m(t_i)\cdot\tau_{\mathrm{AER}}(\lambda_j,t_i)+G_{\mathrm{INST}}(\lambda_j)$$

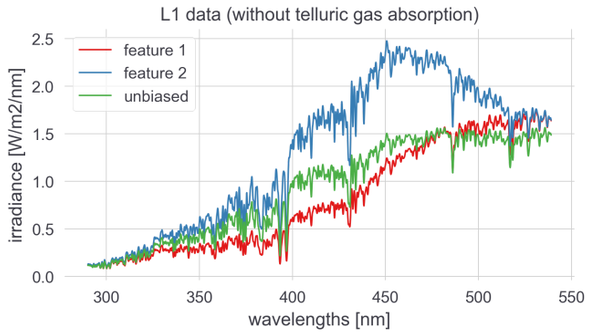
Concludingly, for wrongly calibrated or unstable instruments, G_{INST} absorbs the non-AMF-dependent instrumental component from P_{SMO}, allowing τ_{AER} to be inferred by simple AMF scaling. More technical details are provided in Appendix B.

Key aspects of TanAOD-P can be summarized as follows:

- To mitigate dependence on the exact fitting setup, smoothed spectral residuals are added to the evaluated P_{SMO}. Thus, if a higher-order instrumental effect is not captured by a lower-order smoothing polynomial, it remains in the residuals and is still considered.
- The retrieval of τ_{AER} is stabilized by constraining its spectral dependence with the three-parameter Ångström law.
- Pre-retrieval feasibility checks are used to select solver options or skip fitting the instrumental component when there is no evidence of instrumental bias.
- Multi-day smoothing of the instrumental contribution reduces potential day-to-day fitting noise.

3.1.2.1 Test Application to Synthetic Data

To prove the concept, TanAOD-P was applied to simplified synthetic data with no gas absorption and constant AOD over the day, spanning solar zenith angles from 1° to 82°. A correctly calibrated spectrum, shown as the green curve in the left panel, was independently distorted by two multiplicative, spectrally smooth biases, feature 1 and feature 2, mimicking changes in instrumental transmission.



A spectral fitting algorithm was applied to the full wavelength range from 300 to 500 nm. The broadband contribution, actually AOD plus instrument, was absorbed by a fourth-order P_{SMO}. The unbiased spectrum at SZA = 1° was used as reference.

Spectral AOD derived from post-processing P_{SMO} is shown in the left panel of Figure 3.1-8 for the two instrumental features and for the original unbiased case, feat0. The right panel shows the difference from the forward-model value.

The results clearly demonstrate that spectral AOD is well recovered in cases with instrumental biases, while no false bias is introduced in the unbiased case.

This is also reflected in the fitted versus forward-modeled instrumental features shown in Figure 3.1-9. The left panel demonstrates the zero test, while features 1 and 2 are reconstructed in the center and right panels, respectively.

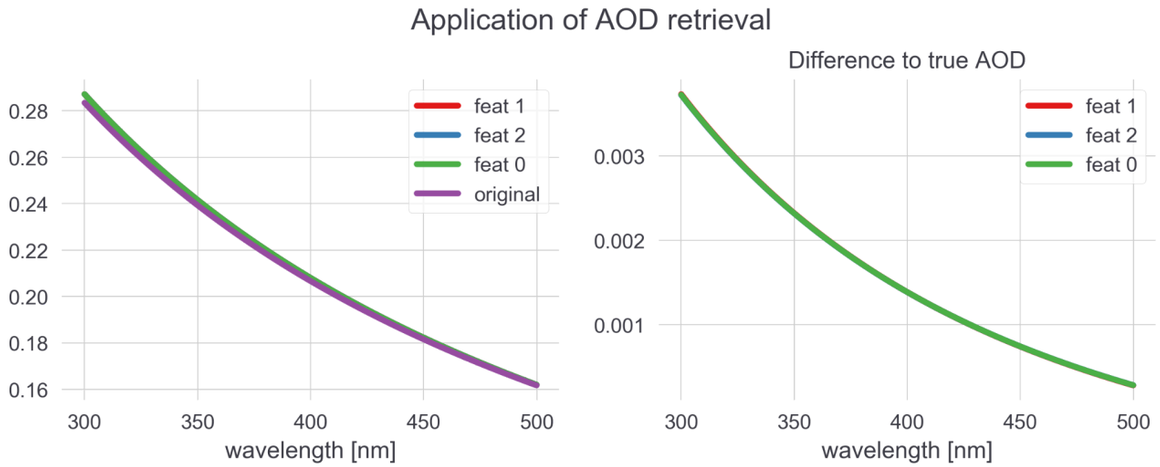


Figure 3.1-8: Retrieved AOD from synthetic data with added instrumental features.

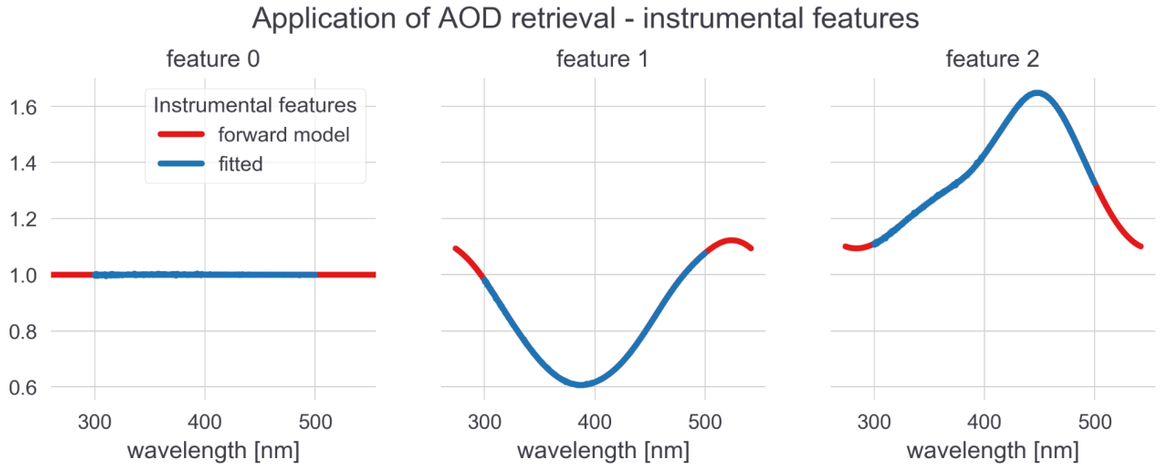


Figure 3.1-9: Recovered instrumental features as part of the AOD retrieval based on synthetic data.

3.1.2.2 Test Application on Field Data

For a first field-data demonstration, two spectral fitting configurations are compared with AERONET reference data, as listed below. UV wavelengths are retrieved from UV bandpass-filtered measurements. VIS wavelengths are extracted from regular measurements, but with a zero-order offset polynomial included.

Name	retrieve AOD at [nm]	measurements	Fitting window [nm]	P_{SMO} order	P_{OFFS} order
UV	340, 380	UV bandpass	340 to 380	2	none
VIS	440, 500	no filter	400 to 500	2	0*

* The VIS setup uses a constant offset polynomial to compensate for intensity-additive biases, such as out-of-range spectral stray light. This contribution is rejected by the bandpass filter in the UV.

We first examine daily results from Thessaloniki. All daily figures follow the same structure: AOD is shown in the top row, with different wavelengths in different columns, and the retrieved instrumental optical depth, IOD, is shown in the bottom row. Red denotes the AERONET reference, black denotes direct conversion of P_{SMO} without post-treatment, and blue denotes our best estimate.

On the first demonstration day in Thessaloniki, 29 July 2023, Figure 3.1-10, excellent agreement with AERONET is observed, even without post-correction of the spectral fitting output. The IOD is correctly found to be close to zero, but still shows slight continuous changes in the UV at both wavelengths and U-shaped variations (probably driven by stray light) at different levels for the VIS wavelengths.

On the second example day, 14 April 2024, Figure 3.1-11, a noon bias of about 0.05 is observed at both UV wavelengths. The algorithm successfully separates this bias into IOD. Both VIS AODs are already free of bias on that day.

A different situation is seen for Tel Aviv, where notable instrumental biases occur regularly. For the selected example day, 15 September 2022, Figure 3.1-12, the IOD contribution is of similar magnitude to the actual AOD.

At Izaña, interesting cases are also detected. The AOD load is typically low at this background site and can usually be retrieved well, as shown for 3 November 2022 in Figure 3.1-13. Two days earlier, on 1 November 2022, Figure 3.1-14, larger particles were still present after a Saharan dust event, leading to nearly vanishing wavelength dependence, with Ångström α < 0.1. This may explain why the UV AOD retrieval failed on that day, incorrectly redistributing most of the AOD into IOD.

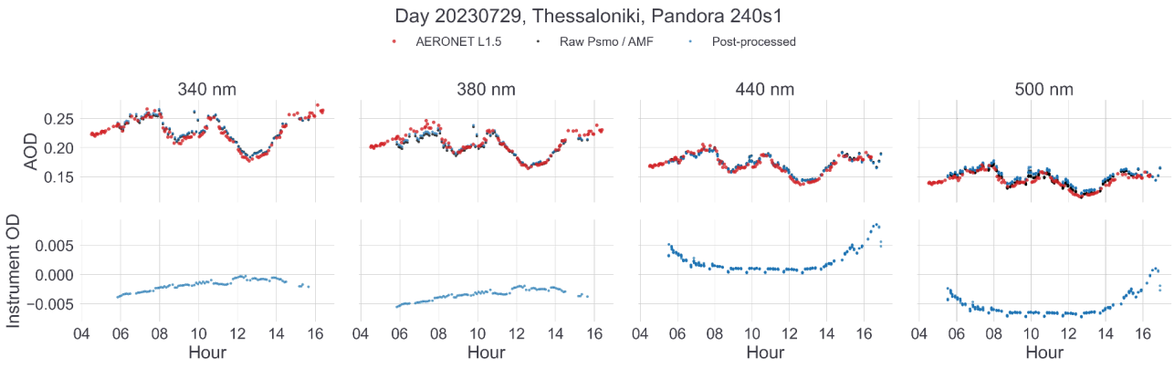


Figure 3.1-10: TanAOD-P application at Thessaloniki on 29 July 2023.

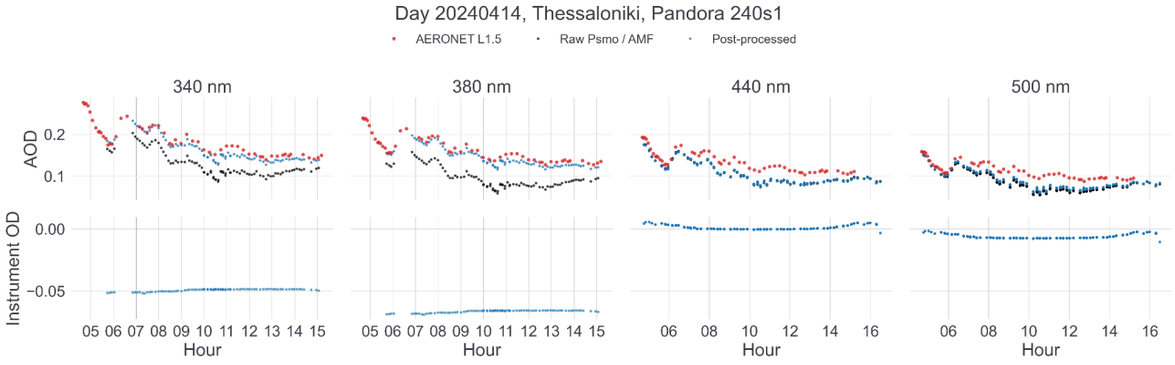


Figure 3.1-11: TanAOD-P application at Thessaloniki on 14 April 2024.

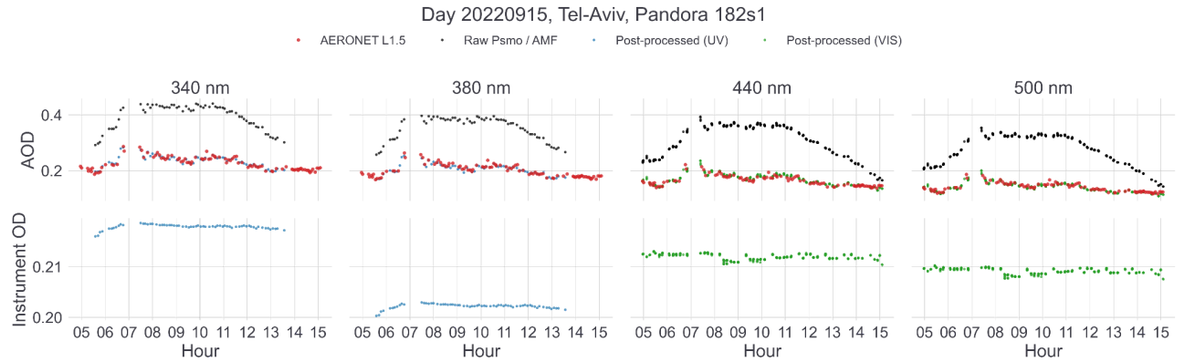


Figure 3.1-12: TanAOD-P application at Tel-Aviv on 15 September 2022.

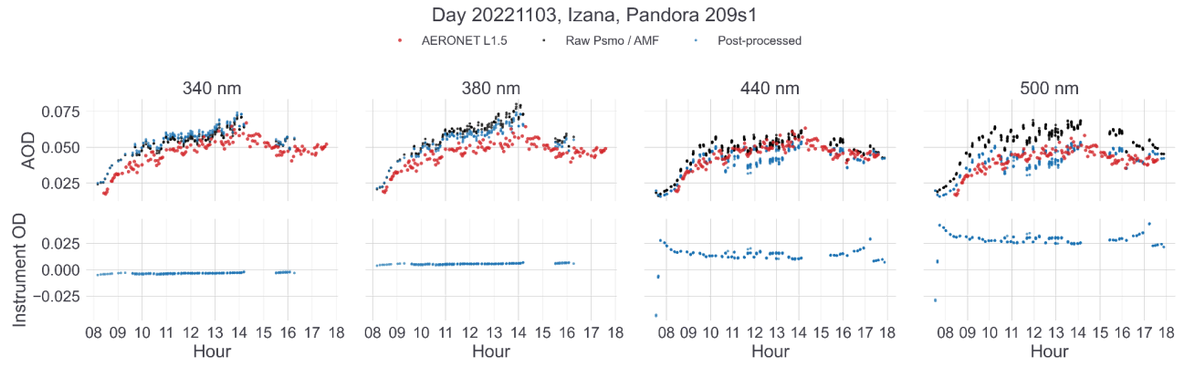


Figure 3.1-13: TanAOD-P application at Izana on 3 November 2022.

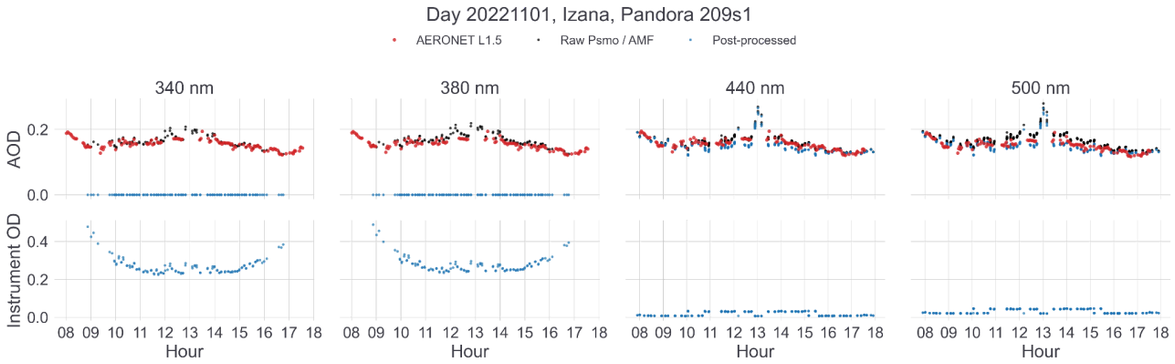


Figure 3.1-14: TanAOD-P application at Izana on 1 November 2022.

3.1.2.3 Mulit-Year Application of TanAOD-P

In the following we show multi-year timeseries of the three locations (Figures 3.1-15, 3.1-16 and 3.1-17, respectively) at the four wavelengths, expressed as relative differences to the Aeronet reference data. For comparison, we also report the "raw" form that is TanAOD-P without P_{SMO} post-processing.

We see a quite diverse range of results. For Thessaloniki, the direct extraction of AOD from the P_{smo} gives already quite good results. The post-processing is not adding much improvement - even sometimes produces outliers (see Oct 2024 for 440 and 500 nm). For Tel-Aviv, the strongly biased AOD from direct utilization of P_{smo} could be well recovered by the post-processing for the majority of the period. However there are also phases, where the post-processing made it worse or didn't have and effect (e.g. summer 2021 or Dec. 2022 for 440 and 500 nm). AOD data retrieved at Izaña is already very good in the "raw" format and periods with substantial instrumental biases could be well recovered in general. But also here some overcorrection can be observed (autumn 2024 for 440 and 500 nm).

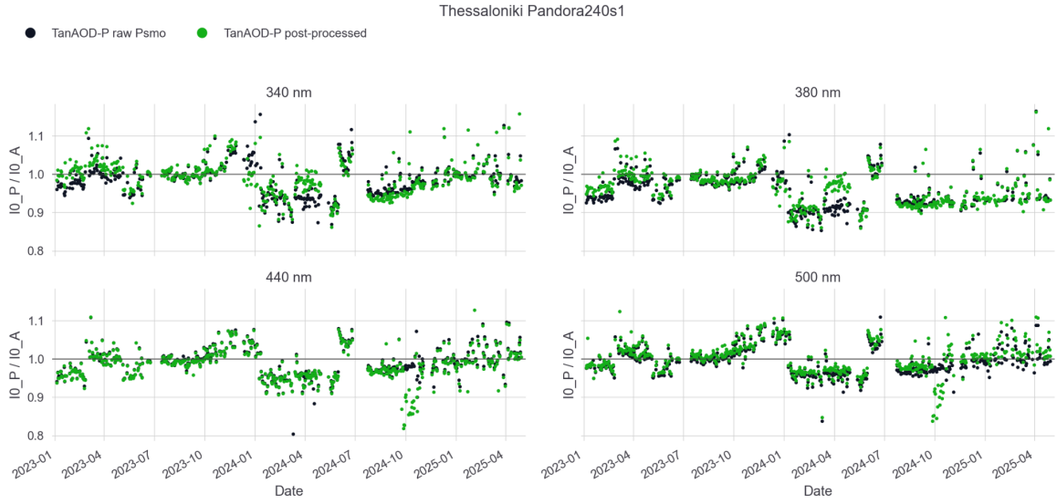


Figure 3.1-15: TanAOD-P application for Thessaloniki.

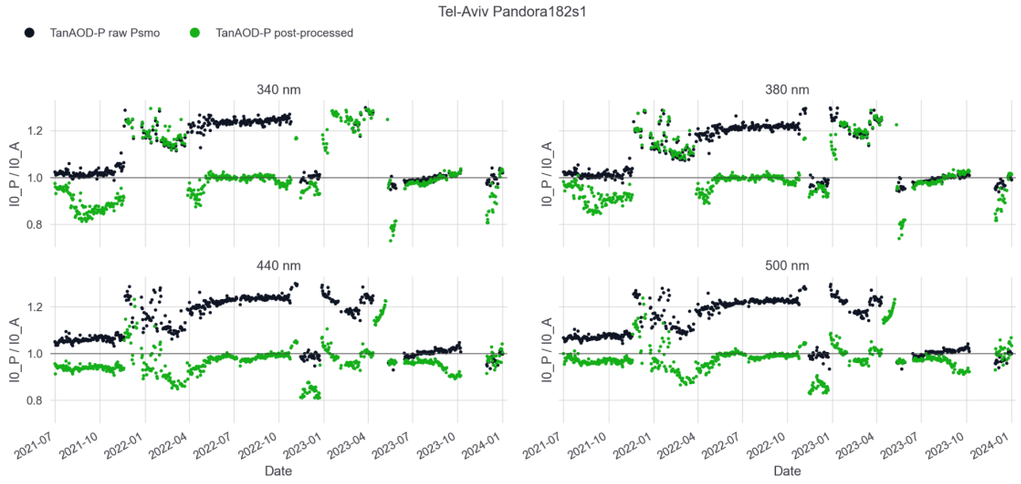


Figure 3.1-16: TanAOD-P application for Tel-Aviv.

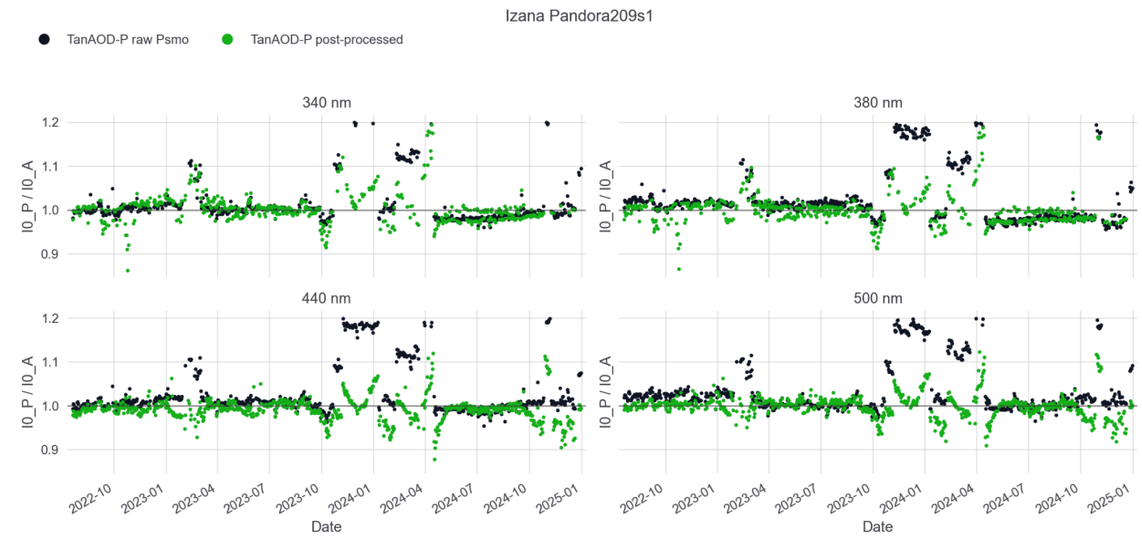


Figure 3.1-17: TanAOD-P application for Izaña.

3.1.2.3 Mulit-Year Application of TanAOD-P + TanAOD-L and Overall Performance Evaluation

Finally, TanAOD-P performance is evaluated on multi-year time scales for Thessaloniki, Figure 3.1-18, Tel Aviv, Figure 3.1-19, and Izaña, Figure 3.1-20, together with the previously shown TanAOD-L results. Green dots represent TanAOD-P data, and yellow dots represent TanAOD-L best estimates.

For reference, each method again also reports its raw form: TanAOD-L without Langley application, shown in gray, and TanAOD-P without P_{SM0} post-processing, shown in black.

We can summarize the performance as follows:

- The baseline difference between the two raw, untreated methods is remarkable and follows directly from the different algorithm concepts. TanAOD-L uses L1 data directly, whereas TanAOD-P benefits from the DOAS analysis, including a measured reference and additional closure polynomials in the fitting process.
- Both methods can recover AOD when instrument sensitivity changes at low frequency, correcting biases of up to 30 %.
- During periods with continuous or higher-frequency sensitivity changes, both methods produce stronger variability than their respective raw data. This is partly caused by the running application of the methods.
- Some shorter periods, characterized by extended bad-weather conditions, reduced data availability, and/or enhanced short-term stability fluctuations, cannot be recovered by either method.
- Performance at 500 nm is representative of the overall performance at other wavelengths. For wavelengths below 380 nm, Pandora bandpass-filter measurements clearly outperform regular measurements (not shown here).
- There are periods where one method outperforms the other. This indicates that an intelligent tandem operation should improve the quality achieved by either method alone.
- For these multi-year observations, the mean agreement with the Cimel reference data is about 5 % for Tel Aviv, 3.5 % for Thessaloniki, and 2 % for Izaña. This suggests that Pandora AOD based on TanAOD is useful for a wider range of applications, but is not yet accurate enough to serve as a reference dataset.

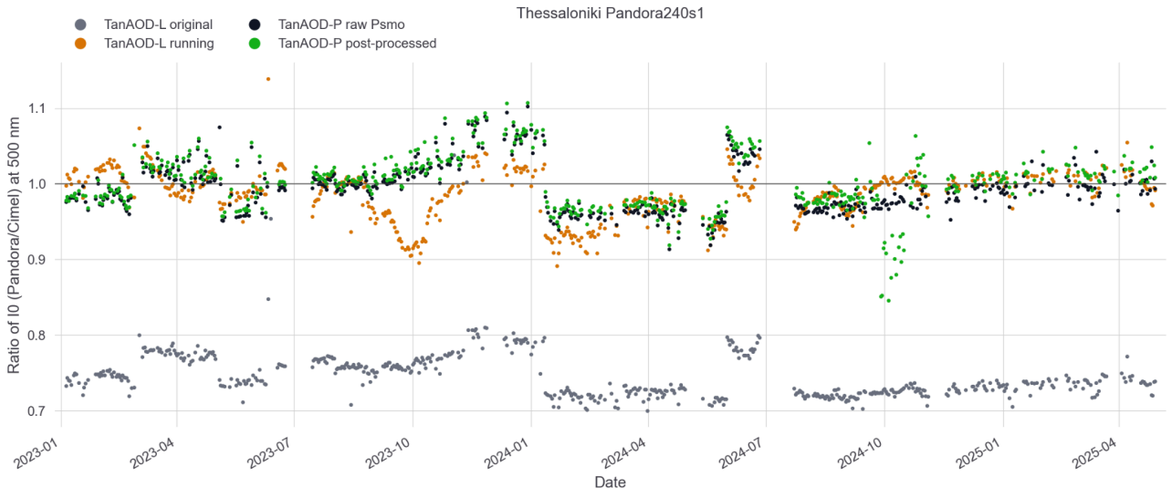


Figure 3.1-18: TanAOD application for Thessaloniki.

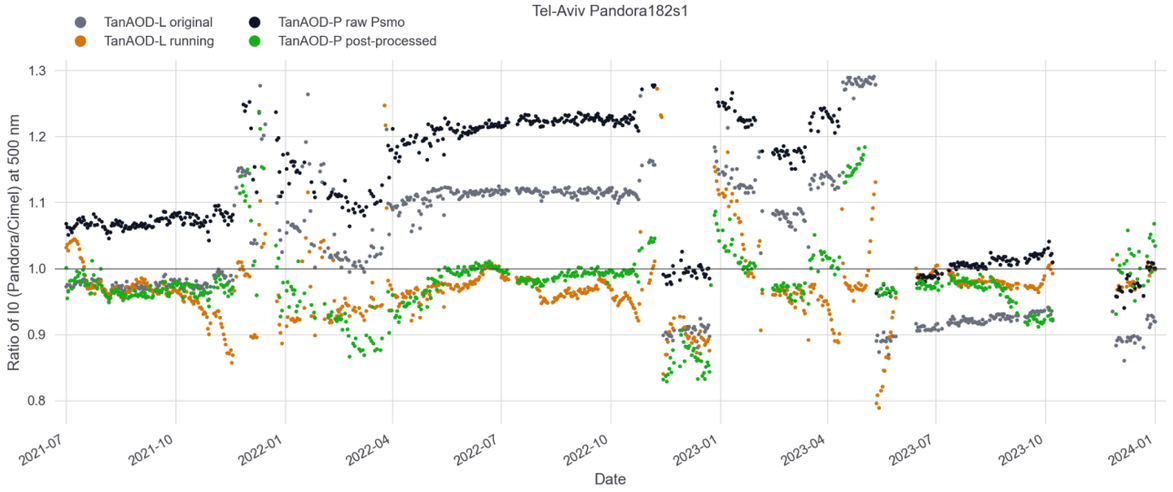


Figure 3.1-19: TanAOD application for Tel-Aviv.

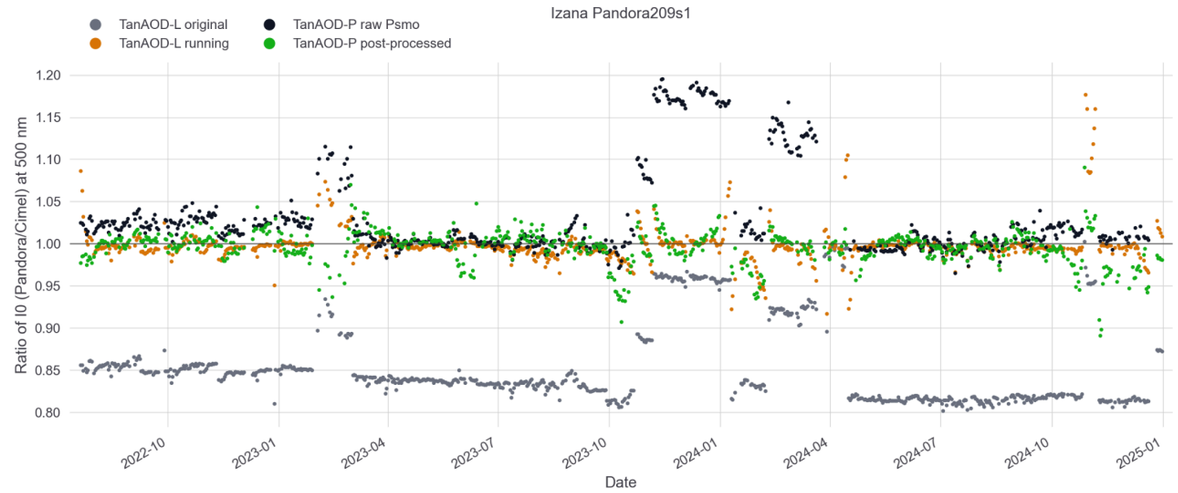


Figure 3.1-20: TanAOD application for Izaña.

3.1.2.4 Way Forward on the Spectral AOD Retrieval

Given the strong radiometric variability of the instrument, the largest challenge for both tandem methods was to tailor the algorithms to cover sudden jumps on different time scales, as well as drifts at different rates. As demonstrated, this has been achieved at a 5 % accuracy level, but further improvements are still possible.

For TanAOD-L:

- Further optimize smoothing lengths and data filtering to reduce overcompensation.
- Continue development of the sky-based tropospheric AOD product to further “reduce” the spectra before Langley extrapolation.

For TanAOD-P:

- Develop more suitable baseline spectral fitting setups, for example longer or overlapping fitting windows with better spectral anchors.

- Test “reduced” spectra in the fitting. This would allow longer windows to be used while constraining gas amounts from the respective best external fitting setup.
- Further investigate multi-day smoothing and weighting of the instrumental component.

For TanAOD as a combined concept:

- Improve understanding of the conditions under which each method performs best.
- Develop a concept for combining both tandem partners. Sequential application, including the use of one method as an a priori for the other, and weighted combination should both be tested.
- Develop a near-real-time implementation strategy. New instrumental correction terms require at least one full day to be retrieved, and potentially the full length of the rolling window. The best AOD should therefore be provided in staged processing levels.
- Prepare the processing infrastructure to handle and combine the two algorithm streams.

3.2 Scattered Light Correction Development (WP1.4)

The goal of WP1.4 is to develop a method to correct Pandora Level 1 data for spatial straylight (SLSPAT) effects. In this chapter, SLSPAT denotes the diffuse sky radiance entering the instrument FOV; the corresponding correction quantity is expressed separately through the diffuse correction factor (DIFF_CORR). This will help reduce errors in trace-gas and AOD measurements during high solar-zenith angles (SZA) and heavy aerosol loads. It's especially important for validating geostationary satellites like Sentinel-4, which frequently observe under these challenging conditions, whereas low-Earth-orbit missions tend to sample closer to solar noon.

SLSPAT often receives less attention than SLSPEC, yet it can introduce comparable biases in data products retrieved in the UV:

- Sky light entering the field-of-view "fills in" absorption features, causing e.g. ozone columns to drop off for high SZAs (and hence ozone slant column), even when data is already SLSPEC corrected (see Figure 3.2-1).

- SLSPAT grows from negligible (< 0.1 %) at low SZA to several percent at high SZA and elevated aerosol loads and affects shorter wavelengths most (e.g. ~4 % at 310 nm vs. ~1 % at 330 nm).

By quantifying and correcting SLSPAT, WP1.4 will help to improve the accuracy for Pandora retrievals in the UV for both morning/evening measurements and high-aerosol conditions.

3.2.1 Method Overview

The SLSPAT correction approach consists of the following steps:

I. Forward modelling of azimuthal radiance distributions using RTM simulations
II. Parametrisation of the radiance flanks and their spectral behaviour
III. Construction of a multi-dimensional LUT linking radiance features to aerosol conditions and DIFF_CORR
IV. Inversion of measured radiance features via LUT matching to retrieve aerosol types and DIFF_CORR
V. Validation via closure studies and comparison with measurements
VI. Extension towards aerosol type detection based on spectral radiance characteristics

3.2.2 Select PanSels and test On-Off-Sun measurement routines

First of all, in this initial task, I supported the PanSels process by providing the necessary site conditions for our instrument study, including the addition of so-called "On-Off-Sun" measurements to the schedule. Table PanSel candidates from WP1.6 gives an overview of all selected Pandora 2S-systems relevant to our WP1.4, chosen to meet the location type criteria: urban, rural, remote, and high latitude. These diverse site types are essential to capture the full range of aerosol types on a global scale. Additionally, the PanSels need to be co-located with an AERONET instrument.

To gain an initial impression of diffuse light behavior, I defined new "On-Off-Sun" measurement routines for different filter types (DIFF and OPEN). In these routines, scheduled after the direct sun observation, a sequence of multiple measurements is performed while scanning across the sun in the azimuthal direction using identical settings (filter wheel position, exposure time) (see Figure 3.2-2). To simplify the process, I modified the built-in *SunSearchRoutine* function to match the setup requirements by adjusting the horizontal increment and azimuth angle range. In the first test, I created new measurement routines FB and FE over an azimuthal range of −4° to +4° relative to the sun, using the Pandora 1-S system P106s1 in Innsbruck. This provided a preliminary overview of how the diffuse light behaves near the edges of the field of view (FOV) on both sides of the solar disc. The results for the Level 1 data within the azimuthal range of −4° to +4° show a significant decrease in the dynamics of the spectra already around −2° and +2° (Figure 3.2-3). We cannot use the DIFF (diffuser) aureole scans for our spatial-straylight study because the data are dominated by instrument-specific artifacts rather than true atmospheric scattering.

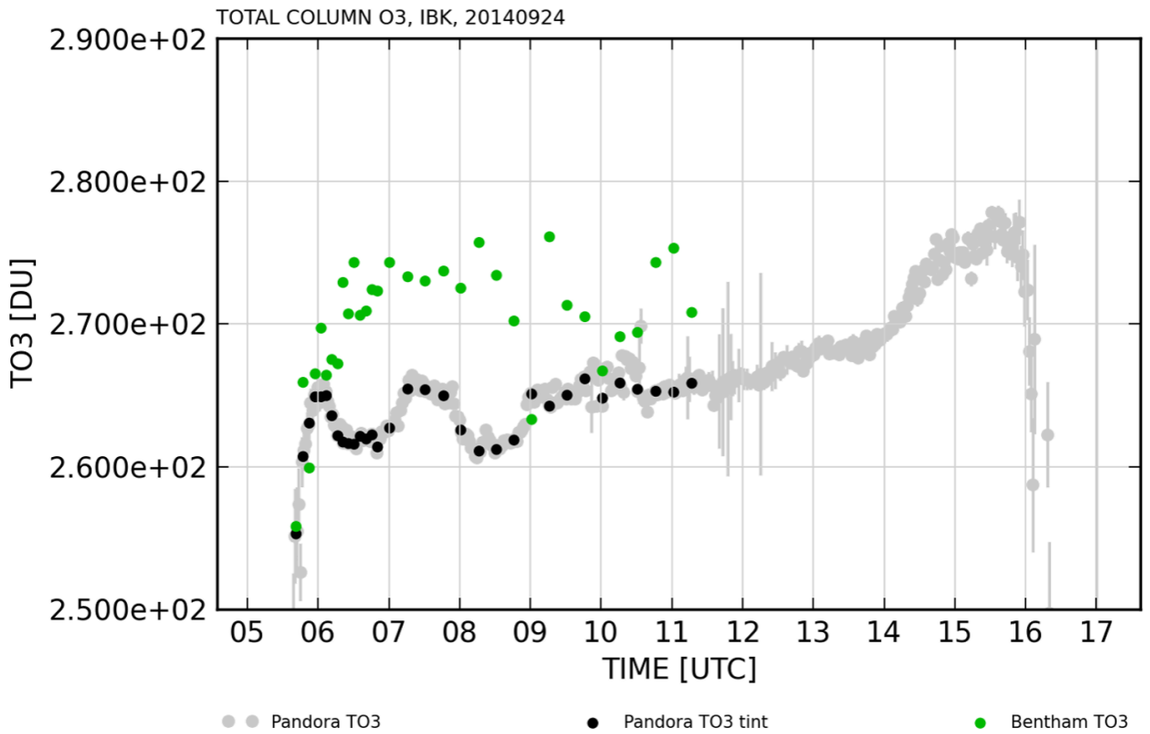


Figure 3.2-1: Comparison of total-ozone retrievals on 24 Sep 2014 at the Innsbruck Observatory. Gray circles show Pandora 106 measurements after spectral-straylight (SLSPEC) correction. Green triangles are ozone columns from the Bentham double-monochromator (entirely free of SLSPEC). Black dots interpolate the Pandora time series onto the Bentham sampling times, showing identical early-hour underestimation and confirming that both instruments are affected by SLSPAT (see Figure 3.3.3 in LuftBlick Proposal PanAMA-202410-Tech).

The pronounced “bumps” in the spectra arise from internal reflections in the spectrometer head or collimator tube, not from sky light. As a result, these measurements cannot be validly compared against RTM-generated lookup tables, which model atmospheric effects. By contrast, the OPEN “On-Off-Sun” scans show a smooth, symmetric fall-off toward the aureole edges. That uniform behavior is exactly what we are looking for to isolate true atmospheric scattering under different aerosol types, solar zenith angles, and wavelengths (Figure 3.2-4). Thus from now on we only use the On-Off sun measurements in combination with OPEN.

We improved the signal-to-noise ratio of the off-sun measurements by increasing the measurement time and demonstrated that an integration time of 2 seconds provides stable results. Furthermore, we performed a fine-tuning of the On–Off–Sun azimuthal scans and found that an azimuthal range of ±8° around the solar disk optimally represents the flanks, based on RTM simulations conducted with Eradiate.

We extended these On–Off–Sun measurements to a second instrument - a 2S instrument collocated with Pandora P106 in Innsbruck and used the same routine to observe a desert dust event with Pandora P209 in Izaña, Tenerife in September 2025.
Figures 3.2-5 and 3.2-6 show the normalized Level 1 On–Off–Sun radiance scans for two contrasting aerosol conditions:

- A moderately polluted, clear-sky urban day in Innsbruck (25 August 2025) with AOD = 0.1–0.3 (AERONET), and
- A desert dust event in Izaña (18 September 2025) with a higher AOD = 0.3–0.7 (AERONET).

The main difference between the two cases lies in the steepness of the radiance flanks outside the FOV. In Innsbruck, where the aerosol loading is lower and dominated by continental urban particles, the flanks are significantly steeper, indicating a sharper decay of radiance away from the solar disk. In contrast, during the Izaña dust event, the radiance decays more gradually, producing flatter flanks. This behavior reflects enhanced forward scattering and radiance redistribution caused by the presence of larger mineral dust particles.

These results confirm that the shape and curvature of the azimuthal flanks carry valuable information about spatial straylight and aerosol scattering properties under different aerosol optical depth and aerosol-type regimes.

3.2.3 Literature research and suitable RTM

On this topic, we reviewed comparable studies that used RTMs or fitting equations to determine the correction factors of the diffuse light for different aerosol types and FOVs. Russell et al. (2004) focused on dust and marine aerosols by analyzing sky radiance and solar beam transmission data from AERONET, along with in situ measurements of aerosol layer size distribution and chemical composition. The relatively large size of dust and marine aerosols results in a strongly forward-peaked scattering phase function, making them significant contributors to diffuse light errors in photometers and pyrheliometers. The study found that corrections for diffuse light (AOD correction factors) are smallest and almost negligible (~1%) for sun photometers with a narrow field of view (FOV) of less than 1° and increase for larger FOVs. For example, for pyrheliometers with a wider FOV of 2.85°, these corrections can be as large as 16% at 354 nm (Figure 3.2-7).

Russell et al. (2004) figured out that the increase in C_a with decreasing wavelength is caused by the increased forward peaking of phase functions with decreasing wavelength.

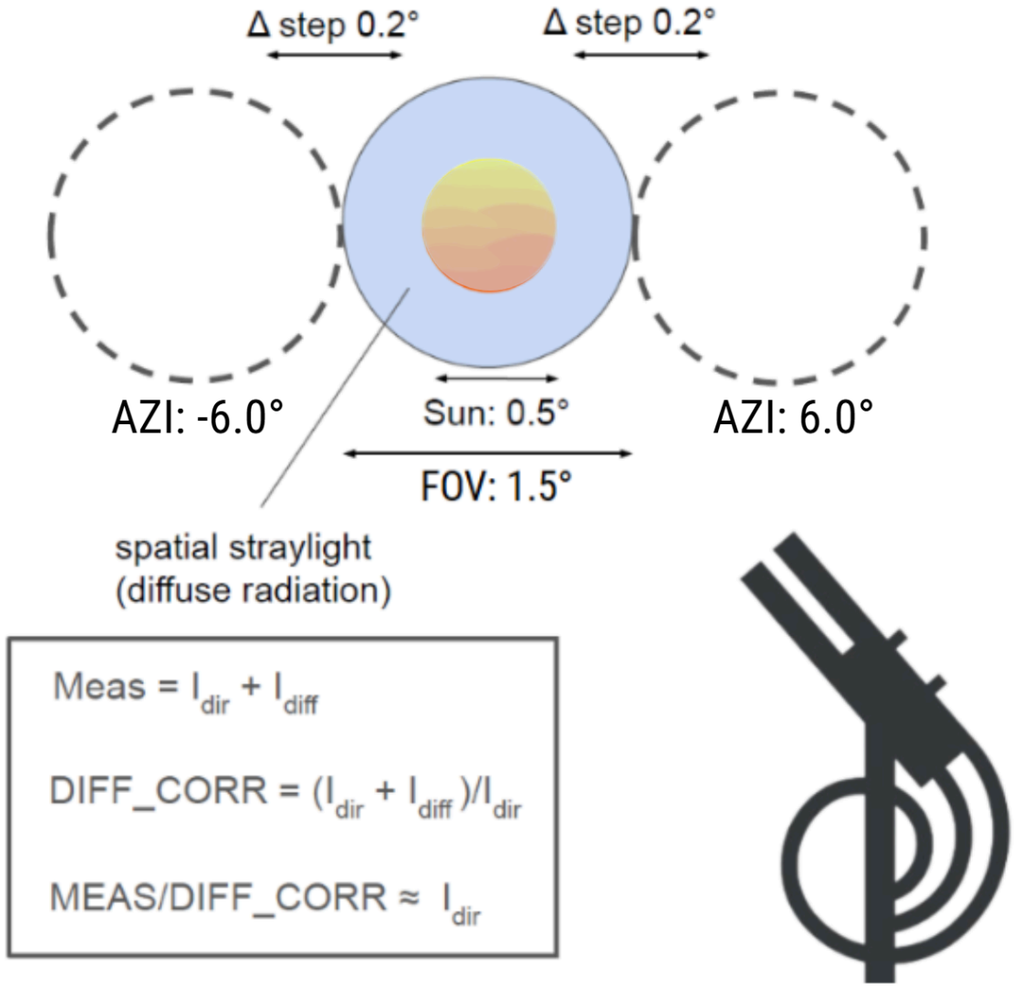


Figure 3.2-2: "On-Off-Sun" measurement routine to measure diffuse light.

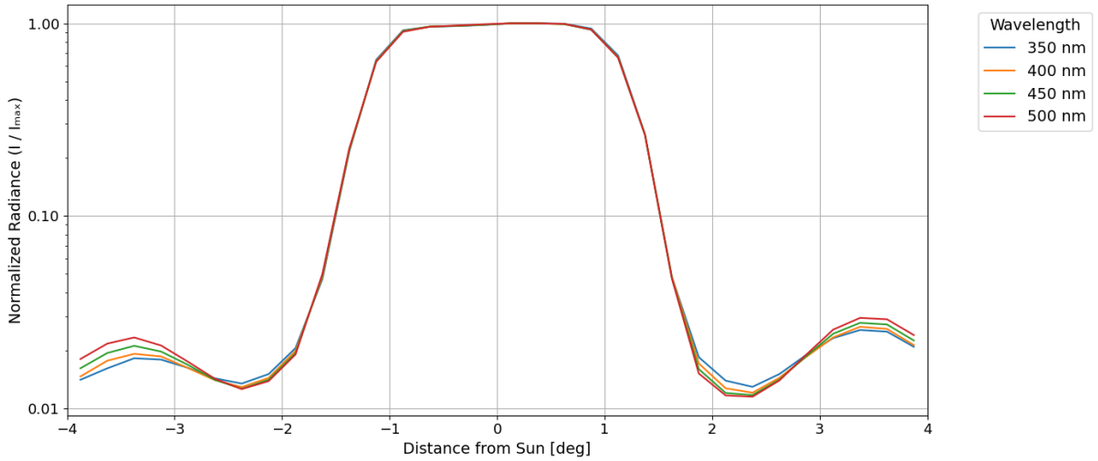


Figure 3.2-3: Normalized L1 "On-Off-Sun" spectra (DIFF) for four wavelengths, each scaled to its own maximum value, showing direct-sun and diffuse-light contributions versus distance from sun in degrees.

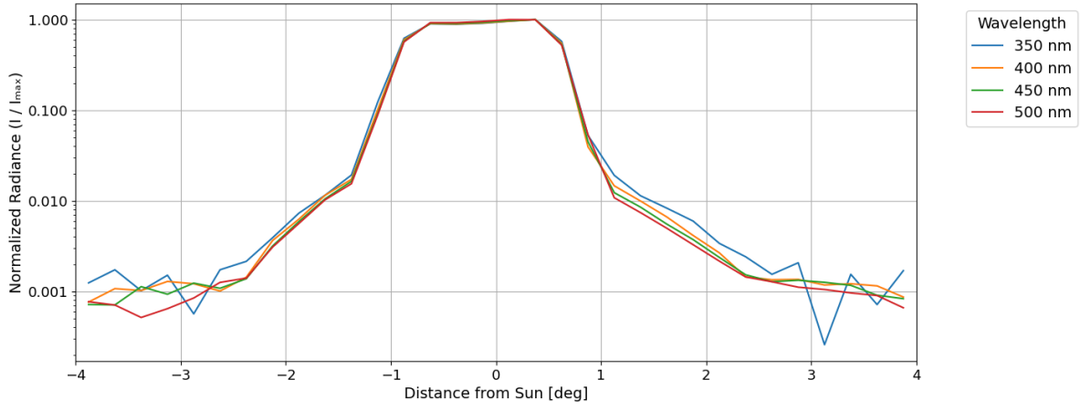


Figure 3.2-4: Normalized L1 "On-Off-Sun" spectra (OPEN) for four wavelengths, each scaled to its own maximum value, showing direct-sun and diffuse-light contributions versus distance from sun in degrees.

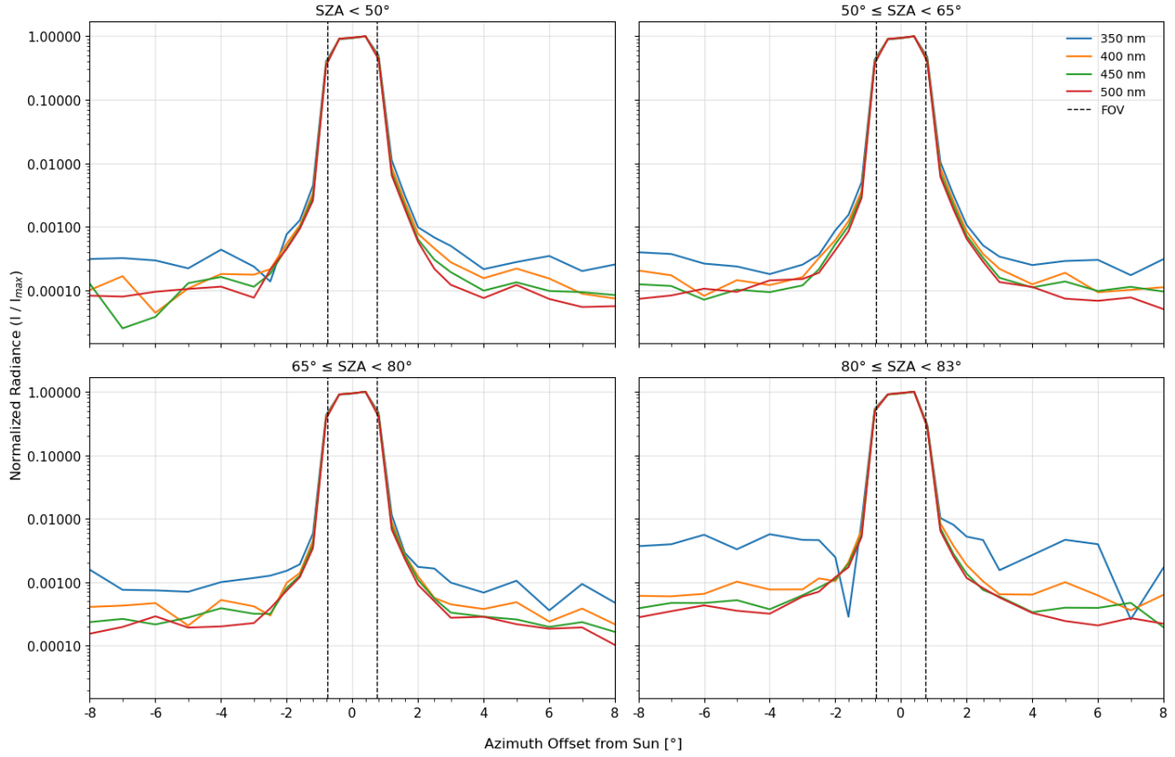


Figure 3.2-5: Normalized L1 "On-Off-Sun" spectra for four wavelengths in P106s1 Innsbruck, each scaled to its own maximum value, showing direct-sun and diffuse-light contributions versus the azimuth offset from the sun degrees for different SZA ranges measured at the 25.08.2025. The dashed line shows the FOV of 1.5° of the Pandora instrument with OPEN.

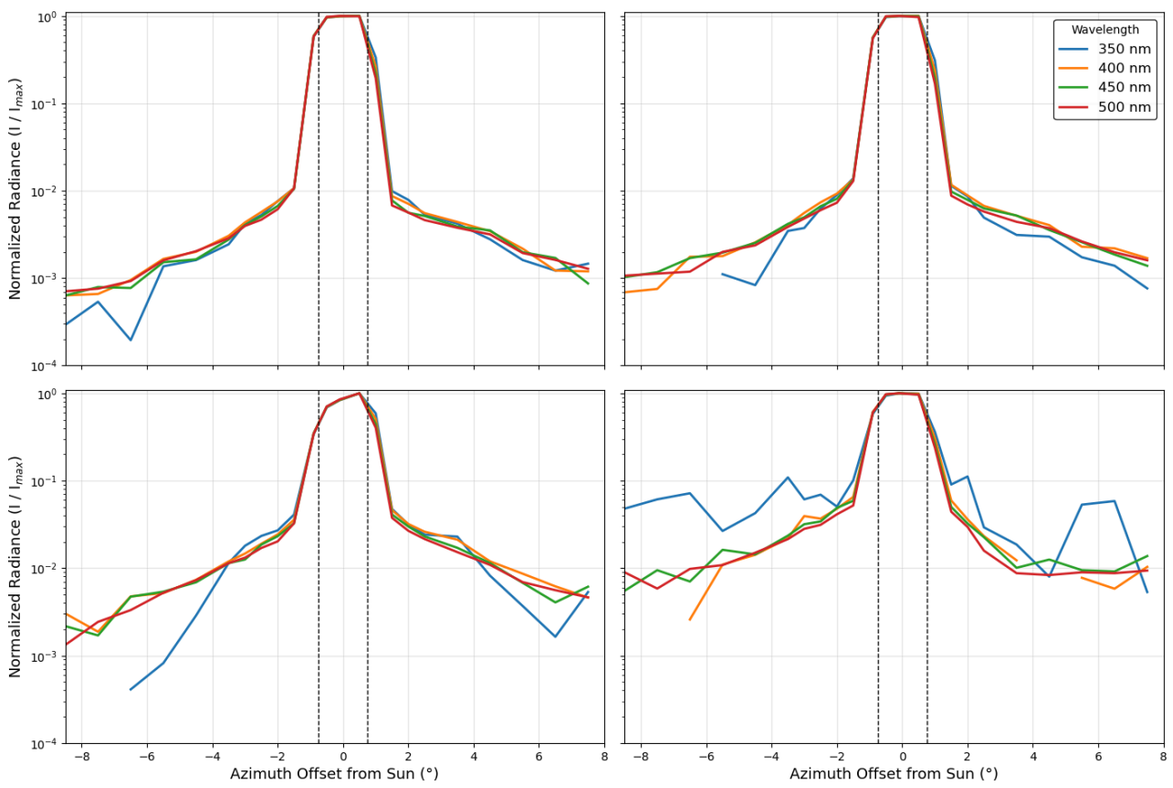


Figure 3.2-6: Normalized L1 "On-Off-Sun" spectra for four wavelengths in P209s1 Izaña, each scaled to its own maximum value, showing direct-sun and diffuse-light contributions versus the azimuth offset from the sun degrees for different SZA ranges measured at the 19.09.2025. The dashed line shows the FOV of 1.5° of the Pandora instrument with OPEN.

Furthermore, Russell et al. (2004) showed that AOD correction factors are well correlated with the wavelength dependence of AOD, as indicated by the Ångström exponent. They provided best-fit equations for determining correction factors from the Ångström exponent of uncorrected AOD spectra and demonstrated their application to vertical profiles of multiwavelength AOD. The limitation of this study, as well as similar studies that parameterised the correction of photometers for diffuse light, is that the aerosol type and AOD were known in those cases. Since this is not the case for our Pandora 2S system, and we cannot retrieve the AOD directly from our data yet, this approach is not applicable to our instrument for now.

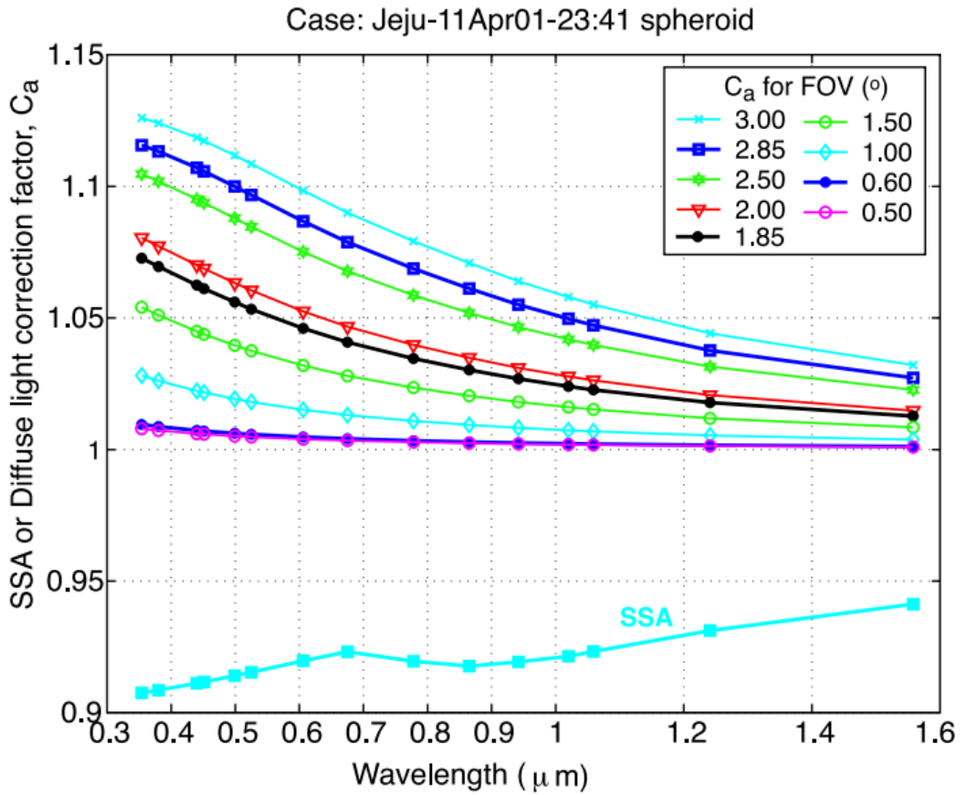
However, we can still learn from these studies. They suggest that the diffuse light contribution for our Pandora instrument, with its 1.5° FOV, is likely associated with relatively large correction factors for diffuse light (or spatial straylight).

Due to the strong dependency of SLSPAT on the aerosol type and load, and since these parameters cannot be directly determined from Pandora data yet, we need to perform radiative transfer model (RTM) simulations to create look-up tables to derive correction techniques. These will help us to translate the measured spectral information, along with the FOV, solar zenith angle (SZA), and assumed aerosol type (size distribution) and load, into simulated On-Off-Sun measurements.

After some research and discussions, we identified four potential RTM candidates, which are shown in detail by Table WP1.6. First, **libRadtran**, developed by the University of Munich, whose further development has largely stopped and is now less actively maintained (Table WP1.6). Second, **(V)Lidort** by Rob Spurr, which is no longer maintained and not freely available. The third most

promising option is the state-of-the-art 3D RTM **Eradiate**, developed by **Rayference**, which is actively maintained, regularly updated, and conveniently available as a Python package.

Figure 3.2-7: Single-scattering albedo (SSA) and diffuse light correction factors C_a for the aerosol size distribution



retrieved from AERONET Sun-sky measurements at Jeju, Korea, 11 April 2001, 23:41 UT (Figure and caption taken from Russell et al. (2004)).

3.2.4 Perform RTM calculations of the sky contribution for different aerosol conditions

3.2.4.1 Motivation and first steps

Because the aerosol type and aerosol optical depth (AOD) are not known a-priori for our Pandora measurements, we have two practical pathways for estimating the diffuse-light (SLSPAT) correction factor:

1. **Direct retrieval from aureole scans**: Identify the characteristic spectral-shape signatures associated with different aerosol types and infer the diffuse-light contribution out of this (see section 3.2.1 Select PanSels and test On-Off-Sun measurement routines).

2. **Look-up table:** Populate a look-up table with diffuse-factor values generated from radiative-transfer-model (RTM) simulations that span a representative range of aerosol types, size distributions, AODs, solar-zenith angles (SZA), field-of-views (FOV) and wavelengths. The goal is to develop a SLSPAT correction algorithm that is based on the simulated data from the look-up table to obtain the respective correction factor.

To implement the second approach, we installed the Eradiate RTM (a Python module) and carried out an initial set of radiance simulations without considering the FOV yet. These simulations covered multiple aerosol types (desert dust, continental, biomass-burning, and urban), a range of particle size distributions and shapes, for different AODs. We then analysed the simulated radiances as a function of azimuth for different SZAs, AODs, aerosol types, and wavelengths. Figure 3.2-8 shows the results of our first RTM simulations with Eradiate with respect to different parameters. For instance, the azimuthal radiance pattern varies significantly between aerosol types such as desert dust showing an almost flat profile, whereas variations in AOD, wavelength, and SZA primarily affect the overall radiance magnitude with comparable shape. These promising results suggest that Eradiate is a suitable RTM, capable of providing the basis for the SLSPAT correction algorithm to be integrated into the Blick software suite for routine Pandora data processing. Furthermore, this demonstrates that the

analysis provides a basis for constraining aerosol type differences, as these features exhibit a stronger variability across aerosol types than across other parameters such as AOD or wavelength.

3.2.4.2 Model set up

After the first aureole scans, a phase of constant testing and communication with the Eradiate team followed. The goal was to identify the most important functionalities and virtual instrument features required to simulate an azimuthal scan with a Pandora instrument for multiple SZAs in atmospheres with different aerosol conditions, such as varying AOD and aerosol type.

In the following, we describe the model setup and the parameters used in Eradiate to include a ground-based, Pandora-like instrument observing the solar disk and sky with a field of view (FOV) of 1.5°. Table 3.2-1 gives an overview of the most important model parameters for each simulation set.

We defined different model setups for five aerosol conditions: aerosol-free, and the aerosol types continental, desert extrapolated, urban, and biomass burning. These simulations were designed as a controlled sensitivity study; therefore, simplified assumptions were adopted, including spheroidal particles, horizontally and vertically homogeneous aerosol layers, and a fixed 2 km layer thickness.

The model atmosphere uses the AFGL (1986) U.S. Standard profile on a regular altitude grid from 0 to 120 km with a 1 km step.

Eradiate was applied in monochromatic-polarized mode, which means that a low-resolution monochromatic absorption dataset covering the wavelength range 250–3125 nm was selected. It includes a simplified subset of only H_2O, CO_2, and O_3 as air constituents and allows for the simulation of the polarization state of light during scattering.

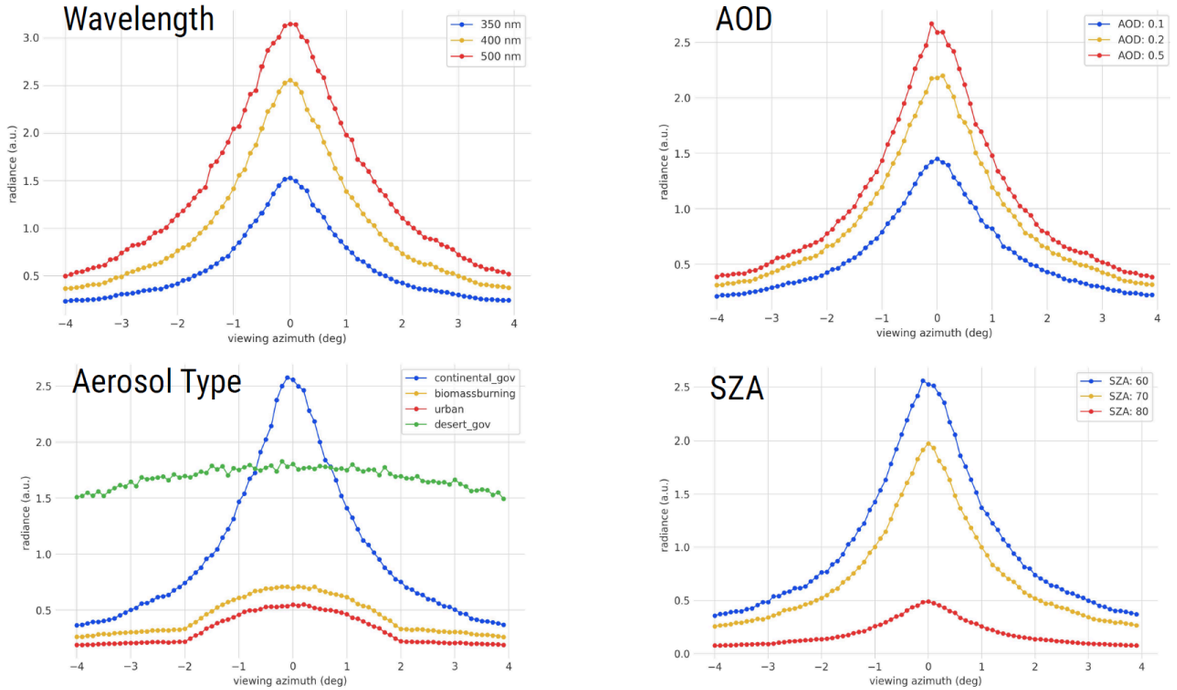


Figure 3.2-8: Radiance simulations with the Eradiate RTM for different wavelengths, AODs, aerosol types and SZAs dependent on the azimuth angle presenting an overview of the content of our planned look-up table. Reference values used: SZA = 60.0°, WL = 400 nm, AOD = 0.3 , Aerosol Type: continental.

The applied Rayleigh depolarization factor (δ = 0.0279) accounts for molecular anisotropy, slightly reducing the degree of polarization and ensuring realistic modeling of Rayleigh-scattered light in the atmosphere. Absorption and Rayleigh scattering were activated in all simulations.

All simulations were configured in principal plane geometry, where the viewing azimuth varies symmetrically around the solar azimuth, and the viewing zenith equals the solar zenith angle (VZA =

SZA). This setup isolates the scattering behavior as a function of relative azimuth, enabling the analysis of angular radiance under well-defined illumination conditions.

Sim ID	Aerosol Type	SZAs (deg)	AOD	Wavelength (nm)	Azimuthal Offset (deg)
REF	Aerosol-free	30–80	0.0	300–500	-6 to 6 with 35 steps.
CON	Continental	30–80	0.05–1.0	300–500	-6 to 6 with 35 steps.
DES	Desert	30–80	0.05–1.0	300–500	-6 to 6 with 35 steps.
URB	Urban	30–80	0.05–1.0	300–500	-6 to 6 with 35 steps.
BB	Biomass-Burning	30–80	0.05–1.0	300–500	-6 to 6 with 35 steps.

Table 3.2-1: Overview of the different model setups used for the Eradiate RTM simulations.

To determine the optimal simulation configuration, we performed multiple simulations with varying settings. First, we varied the number of samples per pixel (photons) and found that a value of 100,000 provides sufficiently stable results for subsequent analysis and fitting procedures. In addition, we observed that representing our 1.5° FOV with a single square pixel was not ideal. Therefore, we defined a circular FOV with a resolution of 100 × 100 pixels and used the mean radiance over the entire FOV for further analysis.

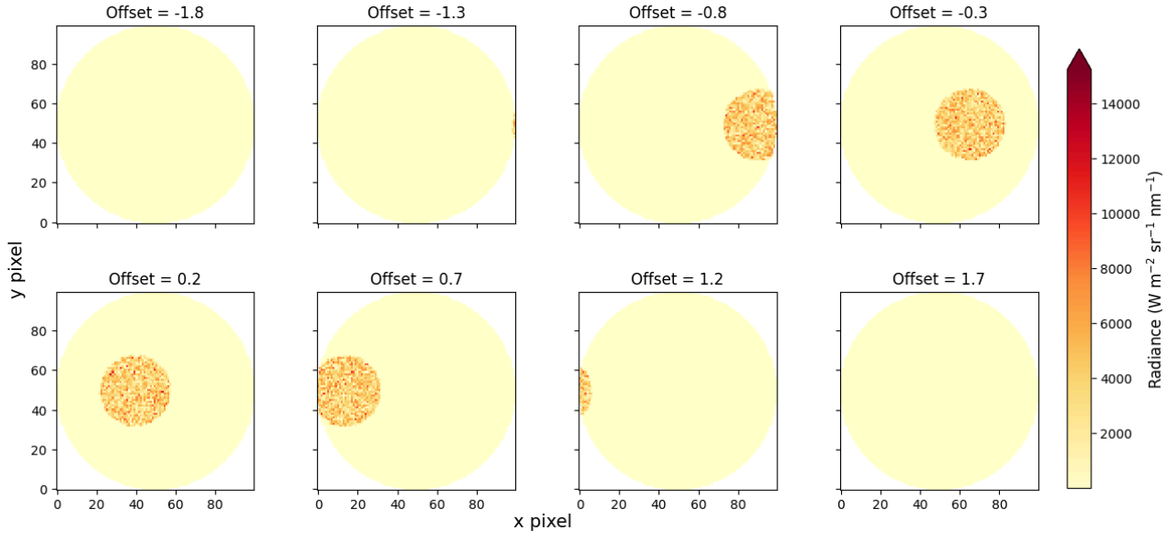


Figure 3.2-9: Circular FOV (1.5°) visualization of the modeled direct-sun measurements here shown as the radiance simulated with Eradiate at a resolution of 100 × 100 pixels using the built-in perspective camera. The simulation was performed for AOD = 0.2, SZA = 50°, and a wavelength of 380 nm.

3.2.4.3 Results

Figure 3.2-9 showcases an azimuthal scan across the solar disk for a moderately loaded desert dust atmosphere (AOD = 0.2). The figure presents the simulated radiance measured by the perspective camera, showing only a subset (0.5° instead of 0.1°) of the simulated azimuthal offsets to maintain visual clarity. It is apparent that the AstroObjectIllumination function, which defines the solar disk, actually represents a heterogeneous solar surface with significant radiance variations.

Figure 3.2-10 shows the normalized radiance of an azimuthal scan across the solar disk for a dust-desert-loaded atmosphere (AOD=0.3) for a spectral range of 350-500 nm. The left and right flanks represent the modeled spatial straylight arising from forward scattering by aerosols, which

must be parameterized in the next step to properly describe this effect in real Pandora measurements.

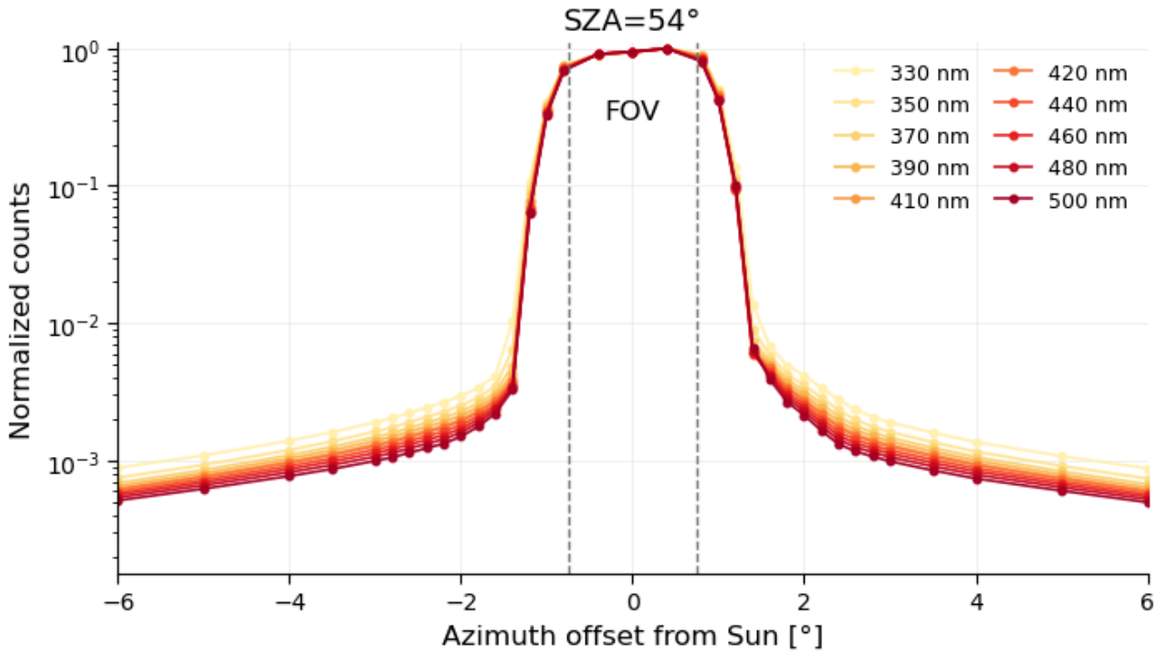


Figure 3.2-10: Logarithmic normalized radiance of an azimuth scan for a SZA of 55° of a moderately loaded desert dust atmosphere (AOD = 0.3) for the spectral range 320-500 nm with a stepsize of 10 nm. The vertical dashed lines represent the FOV (1.5°) of the simulated instrument.

3.2.5 Calculation of modeled SLSPAT and creation of LUT

RTM simulations define the forward problem, while the LUT-based matching of measured parameters represents the inverse step of the retrieval.

3.2.5.1 Parametrisation of radiance flanks and its spectra

The overall concept of our sensitivity study using Eradiate RTM simulations is to establish a reference aerosol-free atmosphere (simulation REF), representing an azimuthal scan under nearly absorption-free conditions, including only molecular absorption by H_2O, CO_2, and O_3 from the AFGL (1986) U.S. Standard atmosphere. This simulation provides the direct solar radiance (I_{dir}) within the instrument's FOV, serving as the undisturbed reference against which the aerosol-perturbed cases representing the direct and diffuse solar radiance ($I_{dir+diff}$) are compared to quantify the effects of spatial straylight caused by aerosols. The relative spatial straylight SLSPAT is defined by the following equation:

$$SL_{\mathrm{SPAT,rel}} = \left(\frac{\int_{\mathrm{FOV}} I_{\mathrm{dir+diff}}(\theta)\, d\theta}{\int_{\mathrm{FOV}} I_{\mathrm{dir}}(\theta)\, d\theta} - 1 \right) \times 100\%$$

where

- Θ is the azimuthal viewing offset angle,
- $I_{dir+diff}(\theta)$ is the simulated radiance including aerosols,
- $I_{dir}(\theta)$ is the aerosol-free reference radiance, and
- the integration is performed over the instrument's FOV of 1.5°.

The diffuse correction factor (DIFF_CORR), which is applied on the measured data by division is defined as the following:

$$\mathrm{DIFF\_CORR} = \frac{\int_{\mathrm{FOV}} I_{\mathrm{dir+diff}}(\theta)\, d\theta}{\int_{\mathrm{FOV}} I_{\mathrm{dir}}(\theta)\, d\theta}$$

and therefore the relative spatial straylight can be expressed as:

$$SL_{\mathrm{SPAT,rel}} = \left(\mathrm{DIFF\_CORR} - 1\right) \times 100\% \,.$$

Thus, SLSPAT and DIFF_CORR describe the same physical effect, expressed either as a relative straylight contribution or as a correction factor. In this framework, SLSPAT quantifies the relative contribution of diffuse radiance within the FOV, while DIFF_CORR provides the multiplicative factor required to recover the direct-beam radiance from the measured signal.

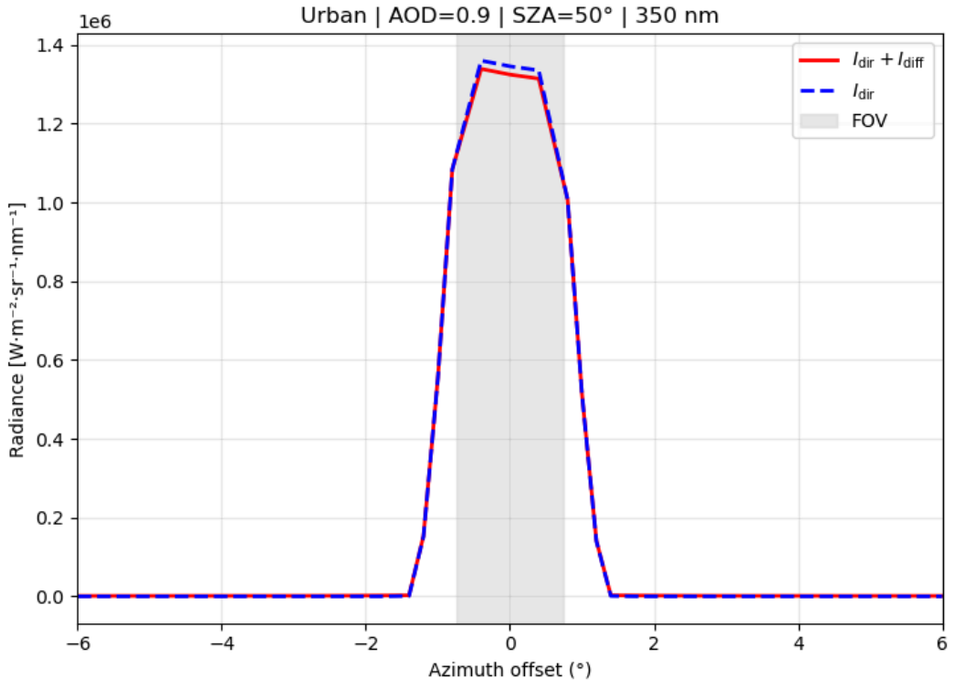


Figure 3.2-11: The direct solar radiance (I_{dir}) of Eradiate azimuth scans (blue) and $I_{dir+diff}$ for a medium loaded atmosphere (red; urban) from simulation URB at 430 nm, SZA of 50° and an AOD = 0.9. The grey area illustrates the instrument's FOV.

Figure 3.2-11 shows a comparison of the normalised radiance for two modeled azimuth scans: one for an aerosol-free case (blue), showing only the direct solar radiance (I_{dir}), and one for an urban aerosol-loaded atmosphere (red), where spatial straylight leads to a redistribution of radiance, reducing the direct-beam contribution within the FOV and increasing the relative contribution of diffuse radiance ($I_{dir+diff}$).

Figure 3.2-12 showcases spectrally the diffuse correction factor DIFF_CORR for a SZA of 60° for an atmosphere loaded with different aerosol types and for varying AOD conditions. As we would expect,

the DIFF_CORR decreases with increasing AODs and smoothly rises towards larger wavelengths. For example an AOD of 0.5 in an urban atmosphere causes around 3 % SLSPAT at 350 nm, at 500 nm this decreases down to 1 %. The DIFF_CORR factor gives us the spectral correction, which needs to be applied onto the measured L1 data, in order to correct for spatial straylight. In a final step this whole calculation of DIFF_CORR was applied onto the whole set of all 22344 variations (CON, URB, DES and BB), representing each variation of SZA, AOD, wavelength and aerosol type, always using REF as the reference simulation.

The final step is to identify the most consistent atmospheric regime within the LUT space in which the instrument is currently measuring to apply the correct DIFF_CORR term, because for Pandora measurements the AOD and the aerosol type are unknown. This inverse step yields the most probable aerosol regime and corresponding diffuse correction factor consistent with the observed radiance features. The next subsection describes the parametrisation of the flanks, which provide diagnostic information that helps constrain the current aerosol regime and identify the most suitable LUT case. The exponential form is empirically motivated and reflects the approximately exponential decay of forward-scattered radiance with increasing angular distance from the solar disk. Figure 3.2-13 demonstrates that an exponential function fitted to the flanks of the maximum-scaled logarithmic radiance data of a simulated azimuth scan by *Eradiate* provides the best description of the data. Out of the left and right tail data a symmetric mean is calculated (yellow curve) and asymmetric data with more than 10% difference are flagged out function as a cloud mask (Roman et al., 2022). The azimuthal radiance distribution, given as logarithmically normalised radiance as a function of azimuthal offset, was fitted using a three-parameter exponential model:

$$I(\theta) = A\, e^{-B|\theta|} + C,$$

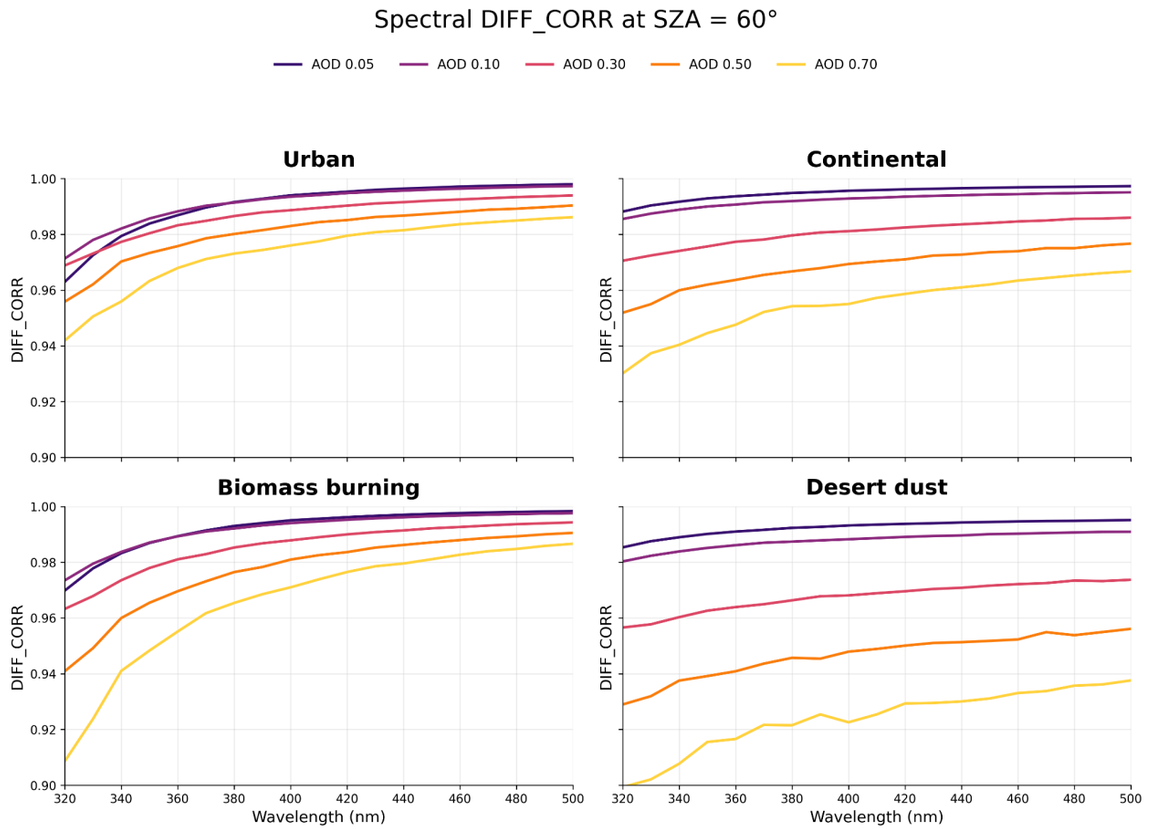


Figure 3.2-12: The upper panel shows the spectral contribution of the modeled diffuse correction factor (CORR_DIFF) an atmosphere loaded with urban, continental, biomass burning and desert dust aerosols for a SZA of 60° and for different AODs (colour lines).

where, *A* represents the amplitude parameter of the fitted radiance distribution, related to the direct-beam contribution, *B* is the exponential decay constant describing the angular spread of scattered light, and *C* represents the asymptotic diffuse background level. Figure 3.2-14 shows the *A*, *B*, and *C* parameters describing each flank fit, averaged over the left and right flanks across the entire spectral range for multiple AOD conditions for the continental aerosol type. These results demonstrate a clear dependence of the parameters A and C on both aerosol type and AOD, only B depends mostly only on the aerosol type and is almost independent of the AOD. In the final step, the spectral variations of the exponential parameters are parameterised by a quadratic polynomial fit, which represents the spectral dependence using the coefficients *a*, *b*, and *c* for each *A(λ)*, *B(λ)*, and *C(λ)*, where each is a function of wavelength (λ). Figure 3.2-15 shows the results of the spectral polynomial fits to A, B, and C, illustrating how their spectral behaviour is described by the polynomial parameters a, b, and c. After implementing this for all possible variations, a clear representation of each individual case, defined by a unique SZA, AOD, and aerosol type, was obtained. Finally, the entire procedure was applied to all simulations, and the results were stored in a multi-dimensional lookup table, which also includes the DIFF_CORR term for each case.

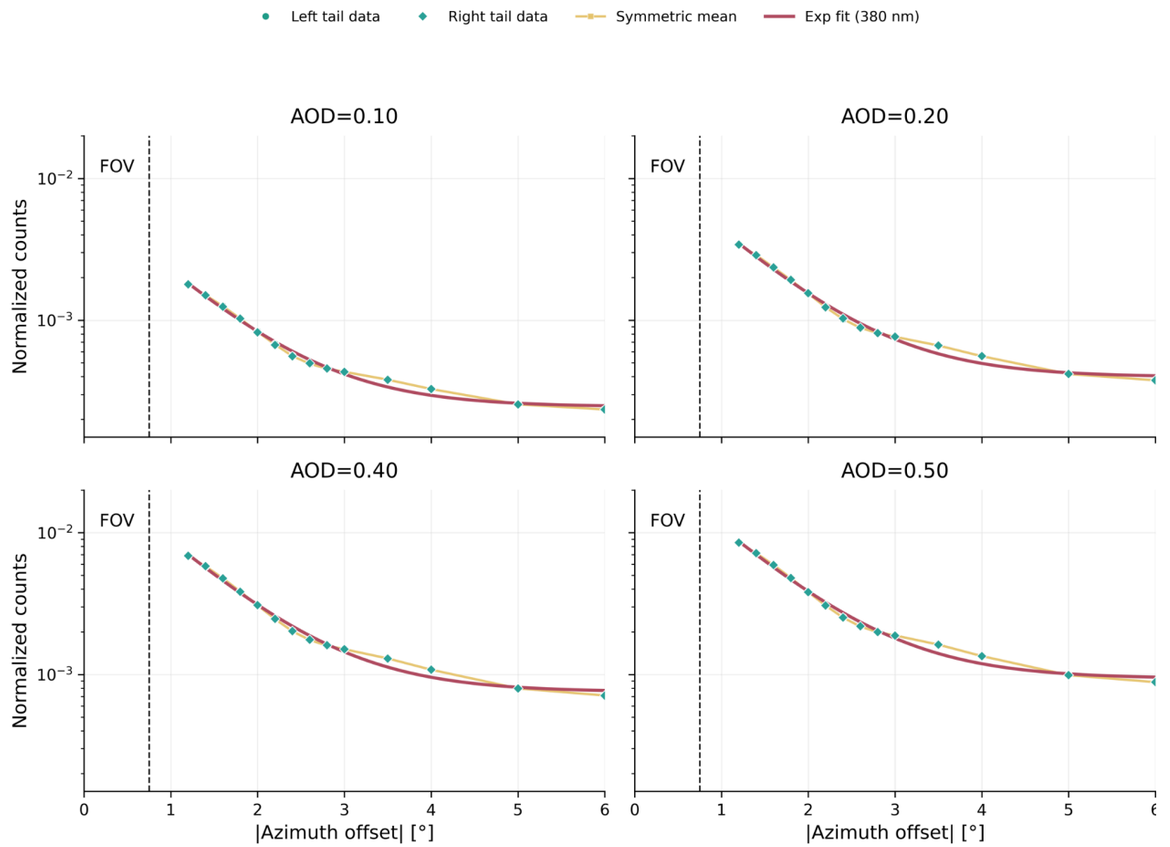


Figure 3.2-13: The log-scaled normalized radiance of Eradiate azimuth scans of simulation CON for a varying loaded atmosphere with AODs=0.1-0.5 for the wavelengths 380 nm (dots) and constant SZA=60°. The lines show the respective exponential fit for multiple AOD cases.

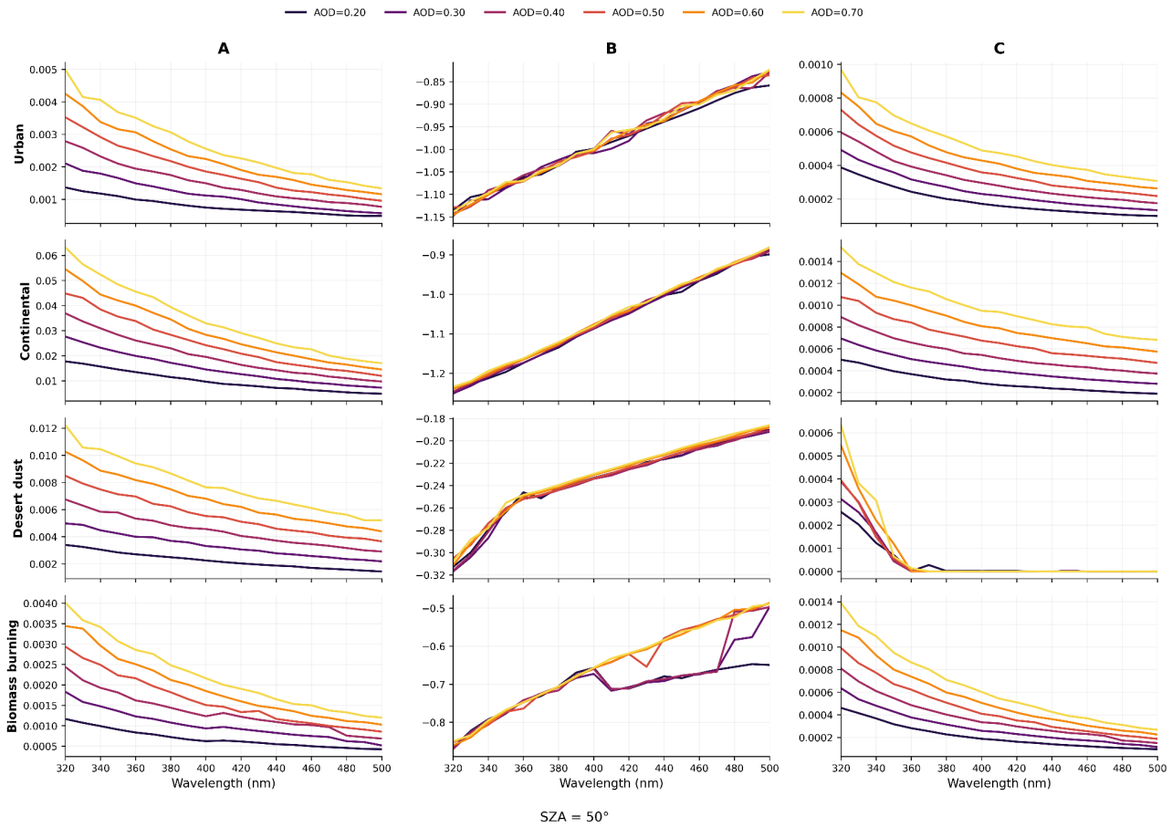


Figure 3.2-14: Exponential A,B, and C parameters over the spectral wavelength range (nm) all simulations, SZA = 50° and different AODs.

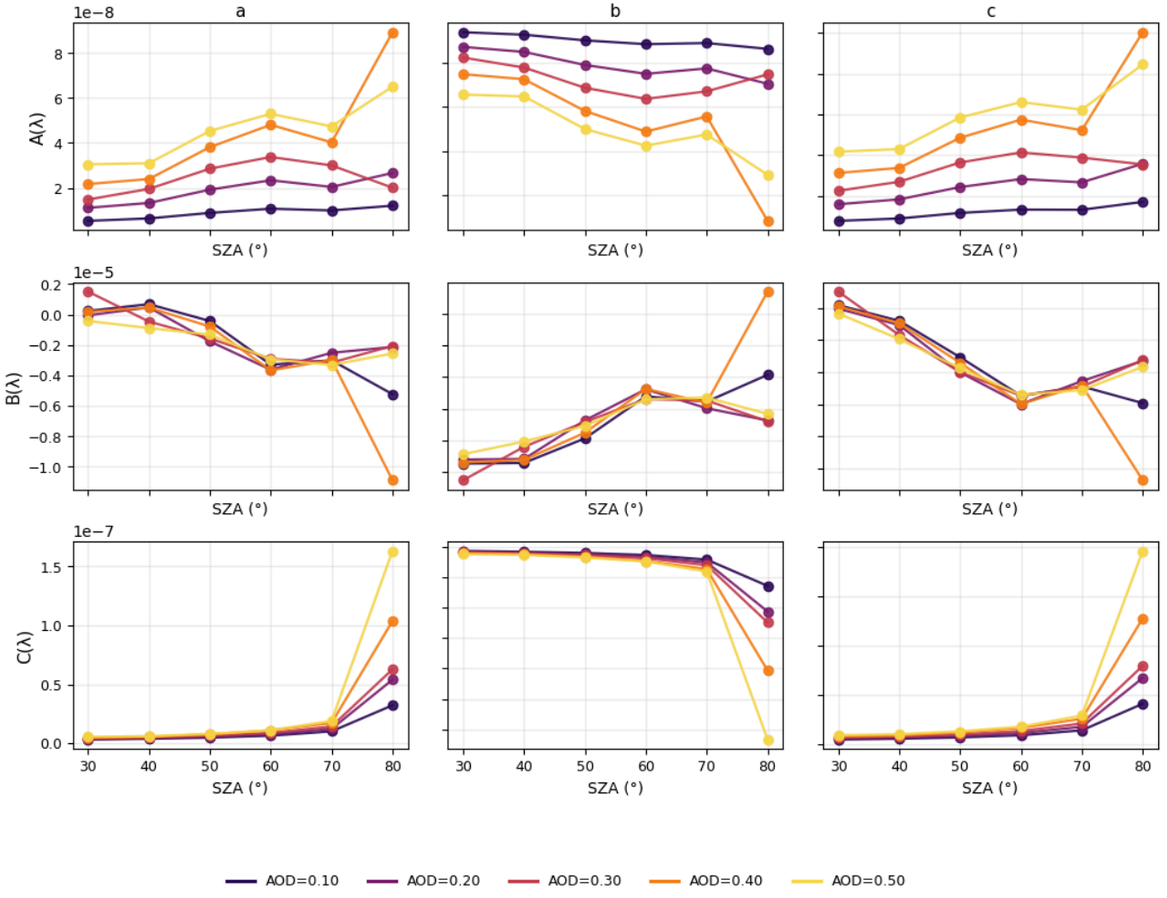
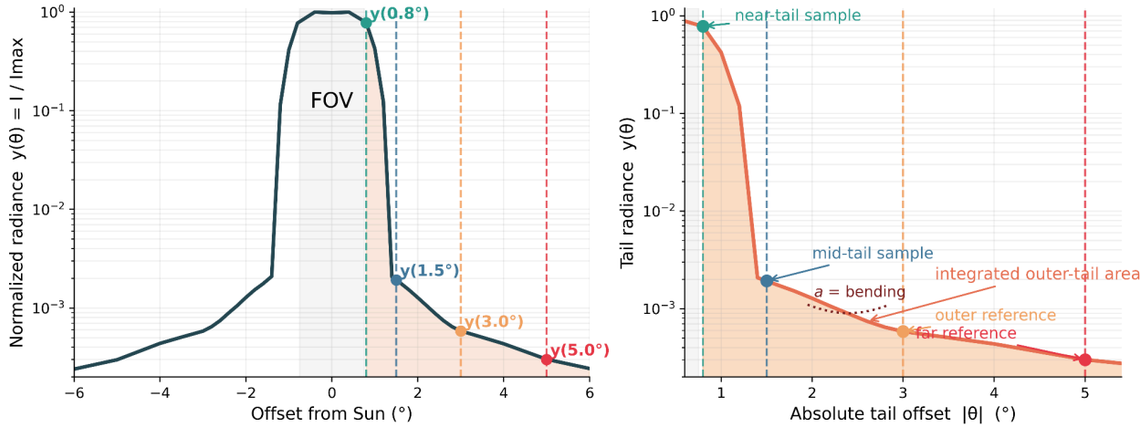


Figure 3.2-15: Spectral polynomial parametrisation (quadratic) for the simulation CON of the exponential parameters A,B and C with $A(\lambda)=a\lambda^2 + b\lambda + c$, showing the results for each SZA and for multiple AOD conditions (colours).

3.2.5.2 Spectral Parametrisation of the tail steepness, outer area and curvature

In addition, we parameterised not only the flanks, but also focused on a spectral parametrisation of the steepness, outer area, and curvature of the flank tail, which considerably stabilised our spatial straylight algorithm. For this, we calculated the steepness using the inner (y(0.8°)) and outer (y(3.0°)) tail ratio, as well as the ratio for the far tail (1.5° and 5.0°) (see Figure 3.2-16). The integrated outer-tail area, as well as the curvature, were also calculated and subsequently spectrally parameterised. These values were then added to the LUT to enable a more robust linear interpolation within the detection algorithm. Figure 3.2-16 shows that the direct tail slope, the tail area, and the curvature of the radiance curves exhibit a clear, mostly linear behaviour both spectrally and as a function of SZA.



Tail steepness

$$R_{\text{in/out}} = \frac{y(0.8^\circ)}{y(3.0^\circ)}$$

$$R_{\text{mid/far}} = \frac{y(1.5^\circ)}{y(5.0^\circ)}$$

Outer-tail area

$$\text{TailStrength} = \int_{\theta_{\text{FOV}}}^{6^\circ} y(\theta)\, d\theta$$

Tail curvature

$$\text{TailCurv} = a, \quad y(|\theta|) \approx a|\theta|^2 + b|\theta| + c$$

Figure 3.2-16: Tail steepness, outer-tail area and curvature, are spectrally parametrised and written into the LUT.

3.2.5.3 Populate the LUT with all coefficients and parameters

The final step consisted of populating the LUT with all derived parameters. These include the spectral coefficients of the flank fitting, as well as additional spectral shape descriptors of the tail, such as its steepness, integrated outer area, and curvature, for each simulated aerosol type, wavelength, SZA, and AOD combination.

Figure 3.2-17 illustrates a subset of the resulting LUT (blue) and provides an overview of the parameters included. In the application of the spatial straylight algorithm to corrected radiance L1 data from Pandora measurements, the same set of parameters is calculated directly from the observational data. The multi-dimensional LUT then enables a linear interpolation within the parameter space, allowing the measured parameter set to be mapped onto the LUT.

As a result, the combined set of fitted coefficients and tail descriptors can be used to infer the best-matching aerosol class and AOD regime within the LUT framework.

3.2.6 Development of the SLSPAT correction algorithm: Closure Study

At this stage, the workflow transitions from forward modelling and parametrisation to the inverse application, where measured radiance features are mapped onto the LUT to retrieve the most consistent aerosol regime. This section briefly describes the concept and implementation strategy of the **SLSPAT correction algorithm** within PM3. Figure 3.2-17 provides an overview of the planned SLSPAT correction workflow, which is tested below for a specific case.

At this stage, it is important to evaluate whether the procedure can reconstruct the direct radiance (I_{dir}) from the original measurement data ($I_{dir + diff}$) using LUT only. The goal is to start with a modeled radiance with an urban aerosol type for SZA of 50° and an AOD of 0.9.

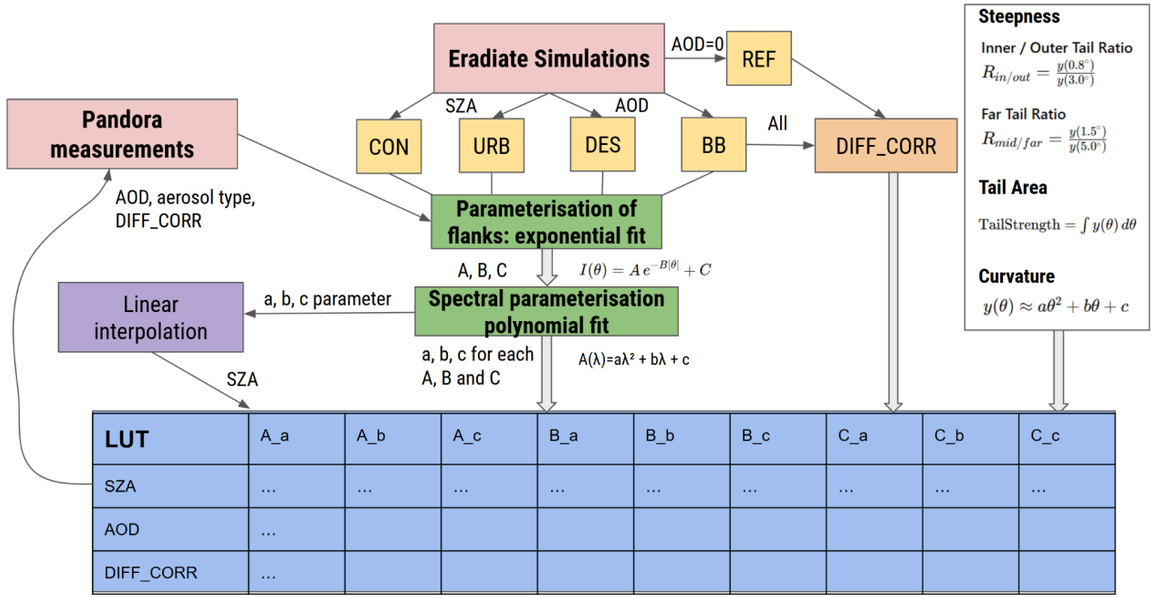


Figure 3.2-17: Spectral polynomial parametrisation of the exponential parameters A,B and C with A(λ)=aλ² + bλ + c, showing the results for each SZA and for all simulated AOD conditions (colours).

We assume no prior knowledge of aerosol type or AOD, and out of this we determine the correct DIFF_CORR value for each wavelength based on the previously generated LUT derived from RTM simulations.

The process begins by fitting the logarithmically normalized flanks of the modeled azimuthal radiance scan with an exponential function. This fitting yields the parameters **A**, **B**, and **C**, which are then spectrally parameterized through the coefficients **a**, **b**, and **c**. Additionally the spectral parametrisation of the tail metrics is also performed. These coefficients and parameters, together with the known SZA, serve as input to a linear interpolation within the LUT, in which the most consistent aerosol regime is identified by minimising the spectral residuals between measured and simulated spectral parameters.

In this example, the linear interpolation yields the most consistent aerosol regime within the LUT space, corresponding to a dominant urban aerosol type.

The identified SLSPAT regime then provides the DIFF_CORR factor for each wavelength (mean: 0.94), which is applied to the FOV by division. In this case, the spatial straylight correction increases the radiance within the FOV by approximately 2-7 %, depending on the spectral range. Figure 3.2-18 shows in the upper panel that the spatial SLSPAT correction algorithm performs robustly across the entire FOV. The residual closure error ranges between 0.1 % and 0.6 %, depending on the wavelength (lower panel), with a spectral average of 0.3 %.

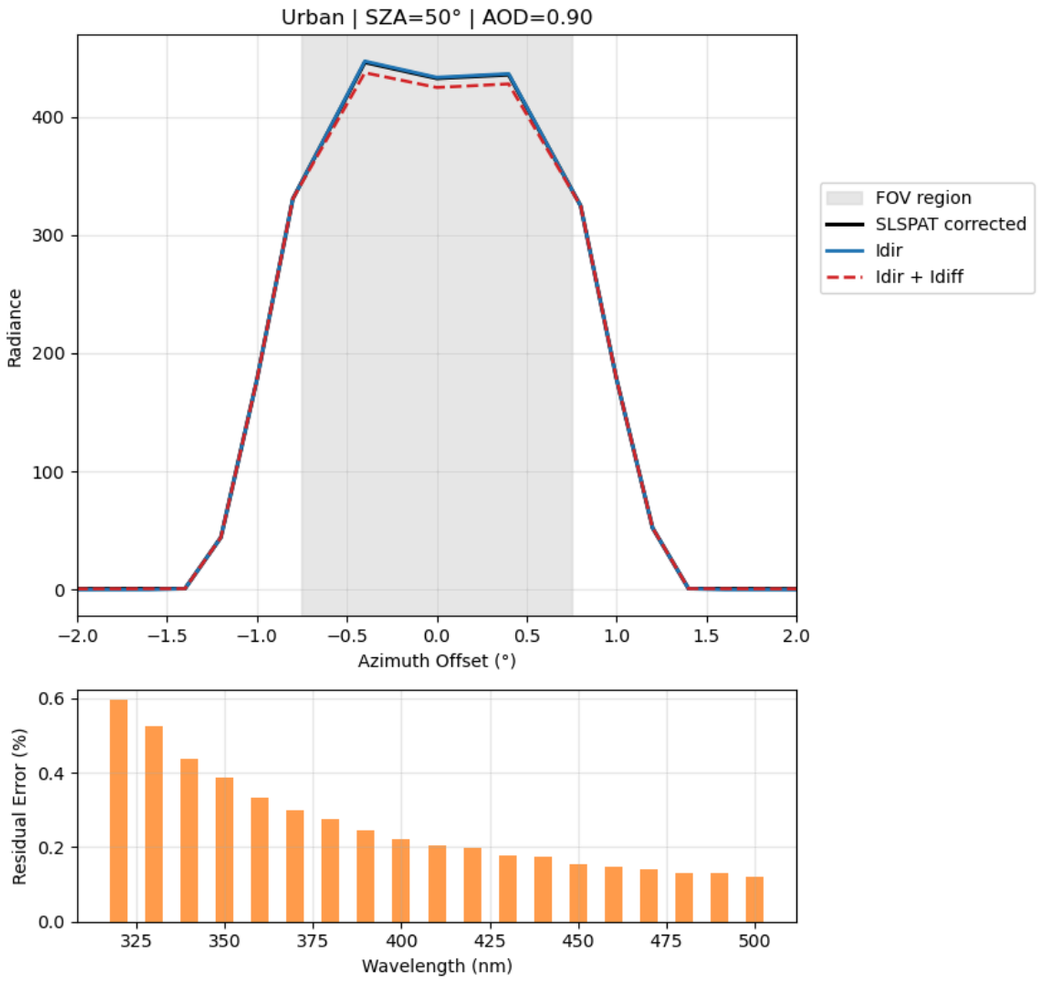


Figure 3.2-18: Modeled azimuthal radiance scans for the aerosol loaded $I_{dir} + I_{diff}$ (red), SLSPAT-corrected (black), and I_{dir} (blue) cases (upper panel). Spectral distribution of residual closure errors across wavelengths in percent (bars, lower panel).

3.2.7 Compare the modeled SLSPAT with the results from the On-Off-Sun measurements

Figure 3.2-21 compares measured L1 azimuthal radiance scans from Pandora P106s1 with simulations from the Eradiate RTM under comparable atmospheric conditions. Both datasets represent a moderately dust-loaded atmosphere with similar AOD.

The radiance flanks of the modeled and measured datasets are of the same order of magnitude. The measured flanks exhibit a smooth and gradual decrease in normalized radiance near the FOV edges, whereas the simulations show a sharper radiance gradient at the FOV edges. This difference can be attributed to the representation of the solar disk in Eradiate and the discretised field of view (100 × 100 pixel grid) used in the simulations, as well as potential atmospheric variability and residual instrumental effects, since Eradiate assumes an idealised instrument.

At SZA = 67°, the measured scans are smooth and stable across wavelengths, allowing for a consistent comparison of the flank behaviour.

These results show that the spectral shape and azimuthal behaviour of the radiance flanks are consistently represented, supporting their use for further analysis.

3.2.7.1 Apply SLSPAT correction on L1 measurements of Pandora 106s1

This section presents a showcase of the SLSPAT correction applied to Level 1 (L1) spectra measured by Pandora instrument 106s1 in Innsbruck on 13 March 2026. The aim is to illustrate both the spectral behavior of the correction parameters and their impact on the measured radiances for different solar zenith angles (SZA = 50°, 55°, 63°, and 72°).

Figure 3.2-19 summarizes the wavelength dependence of the derived **DIFF_CORR** and **SLSPAT (%)** parameters based on the detected aerosol type. The DIFF_CORR factor shows a smooth increase with wavelength, with slightly lower values at higher SZA, indicating a systematic dependence on atmospheric path length. In contrast, SLSPAT exhibits a strong spectral gradient, with highest values

in the UV region and a clear increase with SZA. At 320 nm, SLSPAT reaches values above 6% for SZA = 72°, while decreasing towards longer wavelengths for all viewing conditions.

The effect of the correction on the measured spectra is illustrated in Figure 3.2-20. The SLSPAT-corrected spectra closely follow the structure of the raw L1 counts, including absorption features, while introducing a consistent adjustment in intensity. This effect is most pronounced at shorter wavelengths and higher SZA, in agreement with the behavior of the SLSPAT parameter. The correction primarily modifies the broadband level of the spectra without altering their fine spectral structure.

Overall, this showcase highlights that the SLSPAT correction introduces wavelength- and SZA-dependent adjustments that are physically consistent with increased scattering effects at shorter wavelengths and longer atmospheric paths.

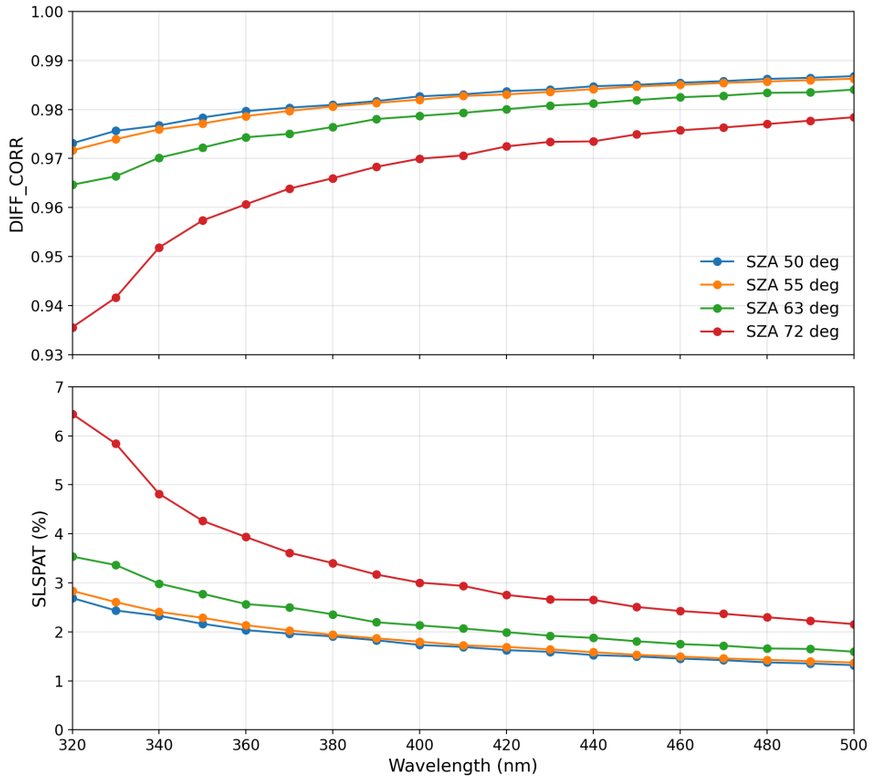


*Figure 3.2-19: Spectral dependence of the detected **DIFF_CORR** (top) and **SLSPAT (%)** (bottom) from Level 1 (L1) Pandora measurements in Innsbruck on 13 March 2026. Results are shown for different solar zenith angles (SZA = 50°, 55°, 63°, and 72°).*

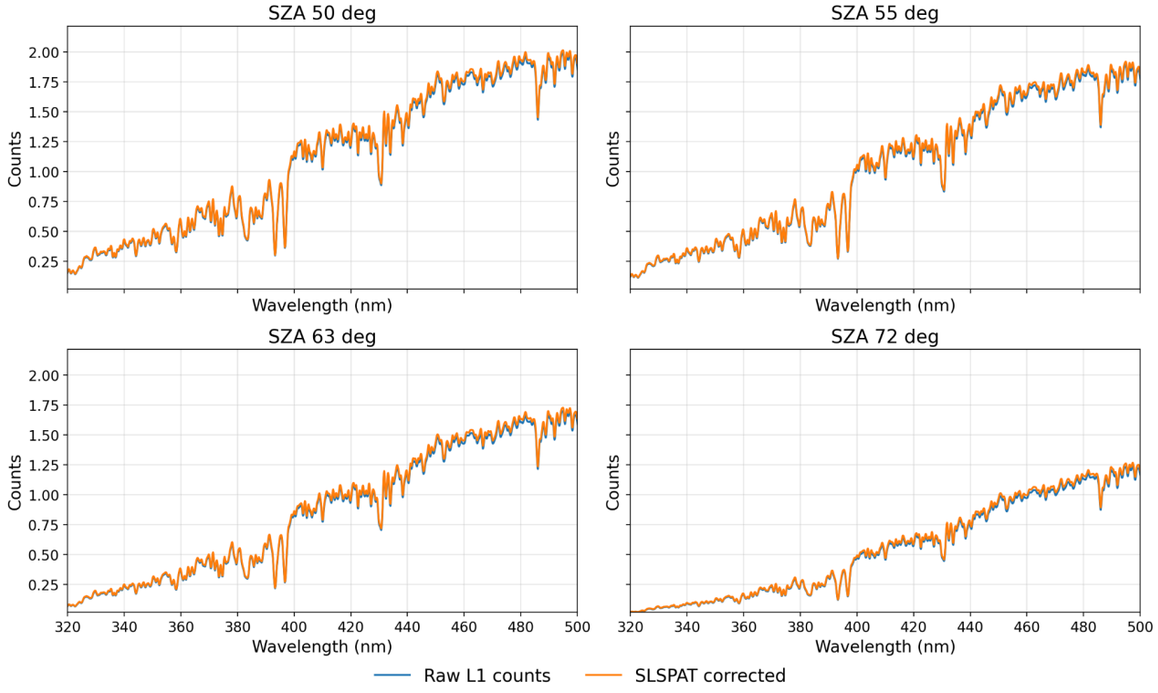


*Figure 3.2-20: **SLSPAT correction** on L1 Pandora spectra measured in Innsbruck on 13 March 2026. Raw L1 counts (blue) and SLSPAT-corrected spectra (orange) are shown for different solar zenith angles (SZA = 50°, 55°, 63°, and 72°).*

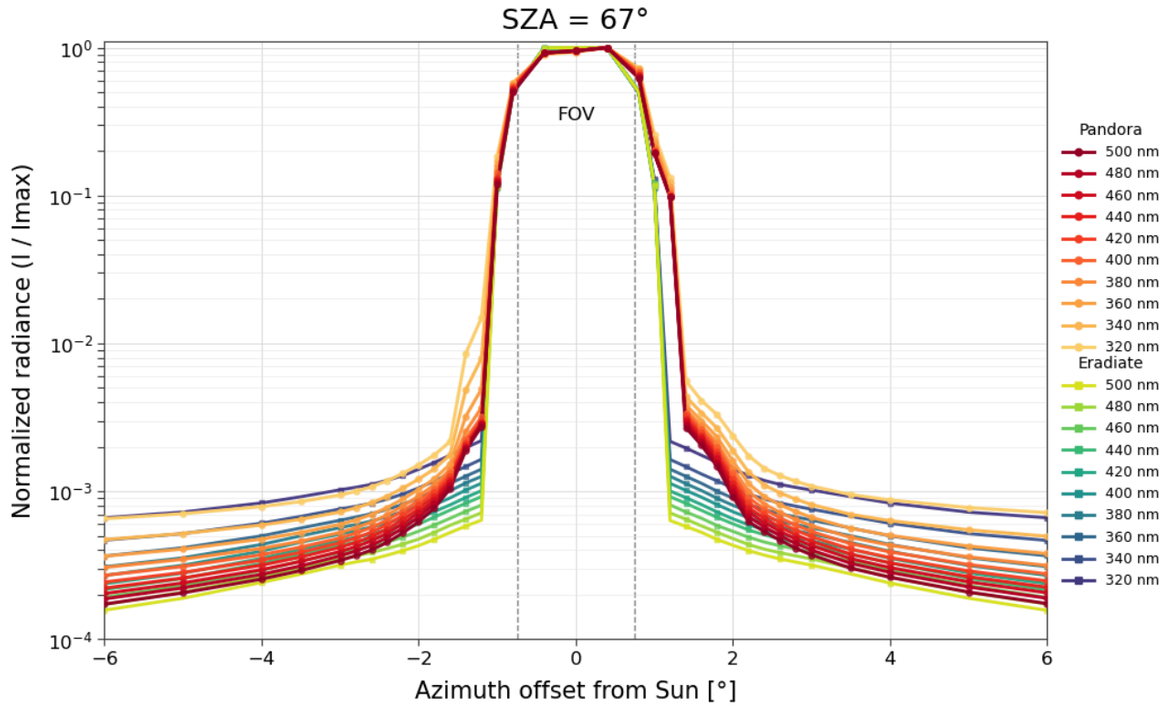


Figure 3.2-21: Logarithmic normalized radiance of an azimuth scan for a SZA of 67° of a moderately dust desert aerosol loaded atmosphere for the spectral range 320-500 nm with a stepsize of 20 nm. The yellow to red colorbar shows the L1 data measured by Pandora P106s1 for the 05.03.2026 in Innsbruck. The blue to yellow colorbar shows the simulated for a comparable case in terms of the AOD and the aerosol type.

3.2.8 From Spatial Straylight Correction to Aerosol Type Retrieval

The focus on aerosol type detection was motivated by the additional information contained in the spectral behaviour of the spatial straylight analysed during the SLSPAT algorithm development. Figure 3.2-14 shows that while all exponential fit parameters (A, B, and C) are used in the spectral parametrisation and subsequent interpolation, the curvature parameter B shows only weak sensitivity to AOD but a clear and systematic dependence on aerosol type. In contrast, A and C mainly describe the magnitude and scaling of the signal and are therefore more influenced by aerosol loading.

This behaviour indicates that the spectral shape is largely driven by B and contains information on aerosol optical properties rather than just concentration. The additional spectral parametrisation (e.g. via A_a and A_b) further captures systematic wavelength dependencies.

Given that aerosol type information is still limited at the global scale and often associated with large uncertainties, this represents a clear added value. The straylight signal therefore contains information relevant for characterising aerosol types. This information is exploited within the LUT-based framework, where the spectral parameters derived from the radiance flanks are mapped onto simulated aerosol conditions.

Furthermore, improvements in the aerosol type detection algorithm are expected to also enhance the SLSPAT correction indirectly. A better constraint on aerosol type or aerosol mixtures for a given observation allows for a more accurate determination of the corresponding DIFF_CORR. This introduces a beneficial feedback effect, where improved aerosol typing supports more accurate straylight correction.

Therefore, further tuning and optimisation of the aerosol type detection or retrieval algorithm should remain a key focus in the next phase of WP2.

3.2.9 Apply Aerosol Type Detection Algorithm on Pandora L1 data

3.2.9.1 CAMS Aerosol Mapping

To evaluate the aerosol-type detection results against CAMS, the CAMS aerosol composition fields were mapped onto the same four aerosol classes used in the Eradiate radiative-transfer simulations, namely Urban, Continental, Desert dust, and Biomass burning (Figure 3.2-22). This harmonization step was necessary because CAMS provides chemically resolved aerosol components, whereas the RTM simulations were performed for a limited set of aerosol types for which suitable optical properties were available. In this mapping, desert dust and biomass burning were represented directly by the CAMS dust and wildfire particulate matter fields, while the Urban and Continental classes were constructed from combinations of elemental carbon, secondary inorganic aerosol, and total organic matter. The partitioning of organic matter between the Urban (30 %) and Continental (70 %) types was introduced as a pragmatic approximation, motivated by the EC-rich combustion character of urban aerosol and the strong secondary and regional contribution to atmospheric organic aerosol described in the literature (Sillanpää et al., 2005; Jimenez et al., 2009). This procedure provides a consistent framework for comparing the detected aerosol-type fractions and their diurnal evolution with corresponding CAMS-based aerosol fractions.

3.2.9.2 Case Study desert dust 05.03.-06.03.2026 Innsbruck

A case study of a moderate desert dust event in Innsbruck (Pandora 106, 05–06 March 2026) was performed to assess the capability of the aerosol type detection algorithm. Figure 3.2-23 shows that the desert dust episode reached its peak on 06.03.2026, following an accumulation phase that began on 05.03.2026. This is reflected in increasing AOD values (AERONET) and a decrease in the Ångström exponent, indicating a shift towards coarse particles. In addition, CAMS dust concentrations also increase, particularly towards 06.03.2026. Figure 3.2-24 and Figure 3.2-25 illustrate the fitting of the radiance flanks and the spectral parameterisation, which provide the coefficients required for the linear interpolation in LUT space. Based on these coefficients, the aerosol type can then be determined. Figure 3.2-26 further shows that the detected aerosol type captures the main temporal evolution observed in CAMS data over the course of the daily cycle, while also revealing that the

algorithm captures finer temporal variations than the European CAMS reanalysis air quality data. On a daily average, the evaluation indicates a slight overestimation of the desert aerosol fraction on 06.03.2026, which is predominantly compensated by the continental aerosol contribution. Differences in temporal resolution and model assumptions likely contribute to the remaining discrepancies. The radiance flanks derived from Pandora L1 data can occasionally exhibit instabilities, which propagate into the LUT-based linear interpolation and may lead to increased uncertainty in the retrieved aerosol regime.

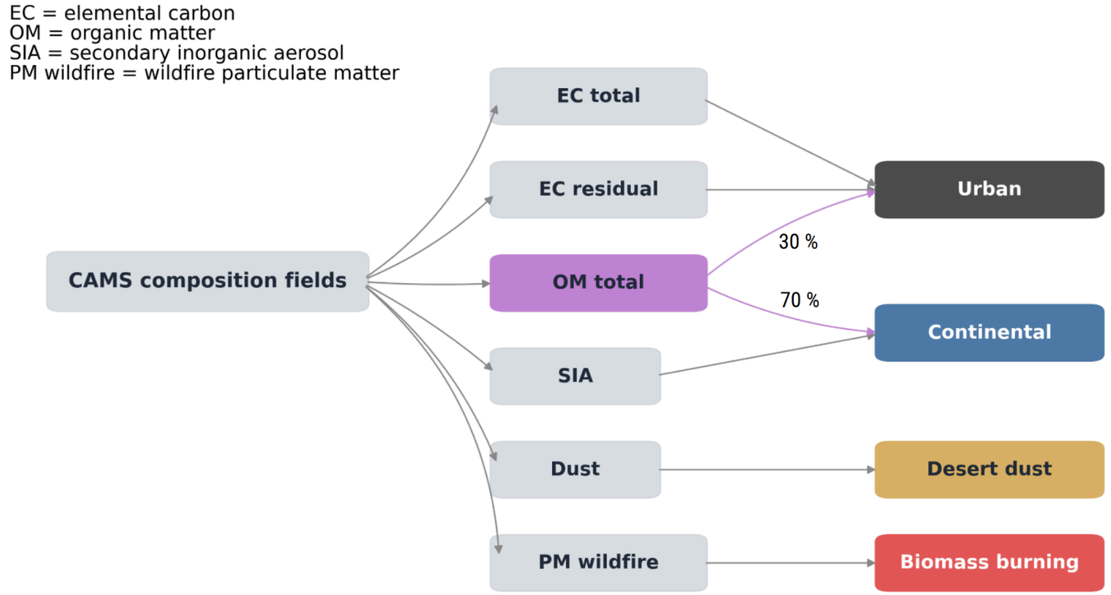


Figure 3.2-22: Mapping of CAMS aerosol composition fields onto the four aerosol types used in the Eradiate simulations: Urban, Continental, Desert dust, and Biomass burning. Total organic matter is partitioned pragmatically between the Urban (30 %) and Continental (70 %) classes following the literature.

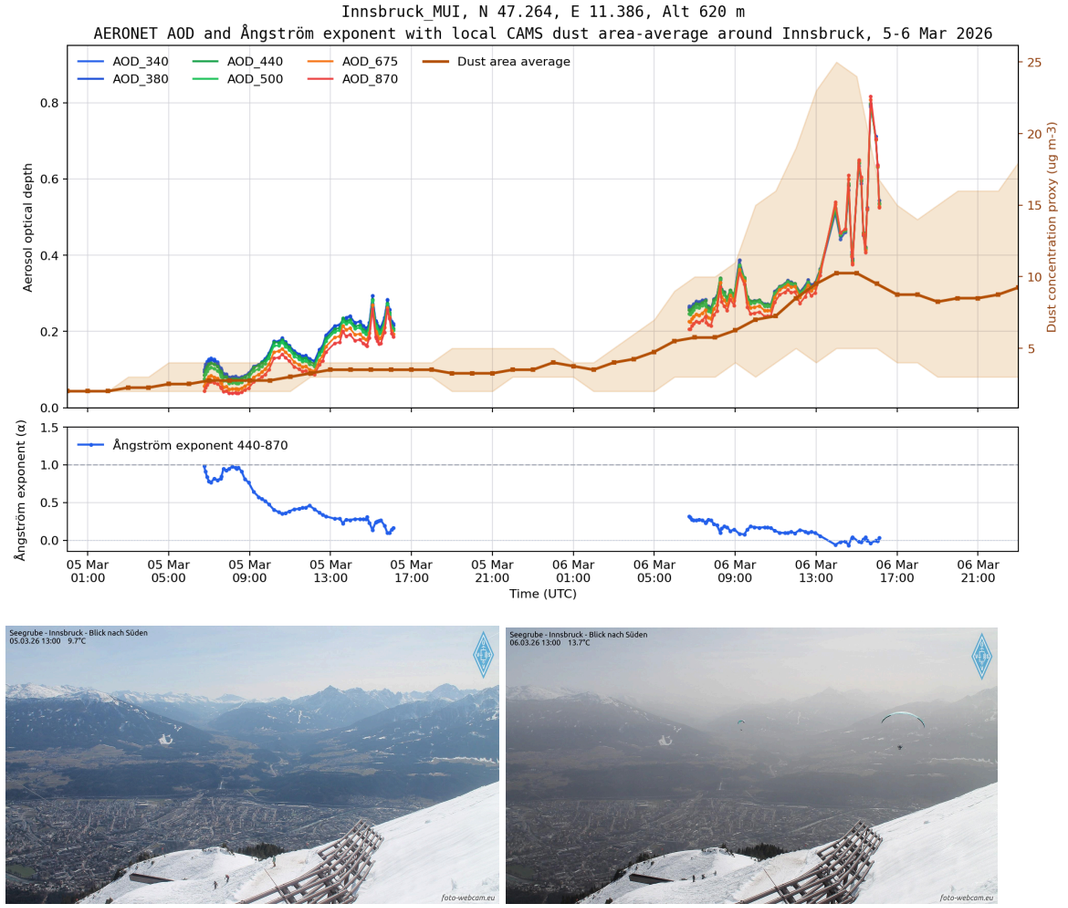


Figure 3.2-23: Time series from 05.03.-07.03.2026 of AOD AERONET data for Innsbruck for different wavelengths (colours) and the dust concentration from the CAMS European air quality analyses with a resolution of (0.1°, approx. 10km). The lower panel shows the Ångström exponent (α). Webcam images taken from the Seegrube of Innsbruck at the rather clear sky day 05.03. (left) and the dust loaded day at 06.03.2026 (right) (www.foto-webcam.eu: accessed at 21.04.2026).

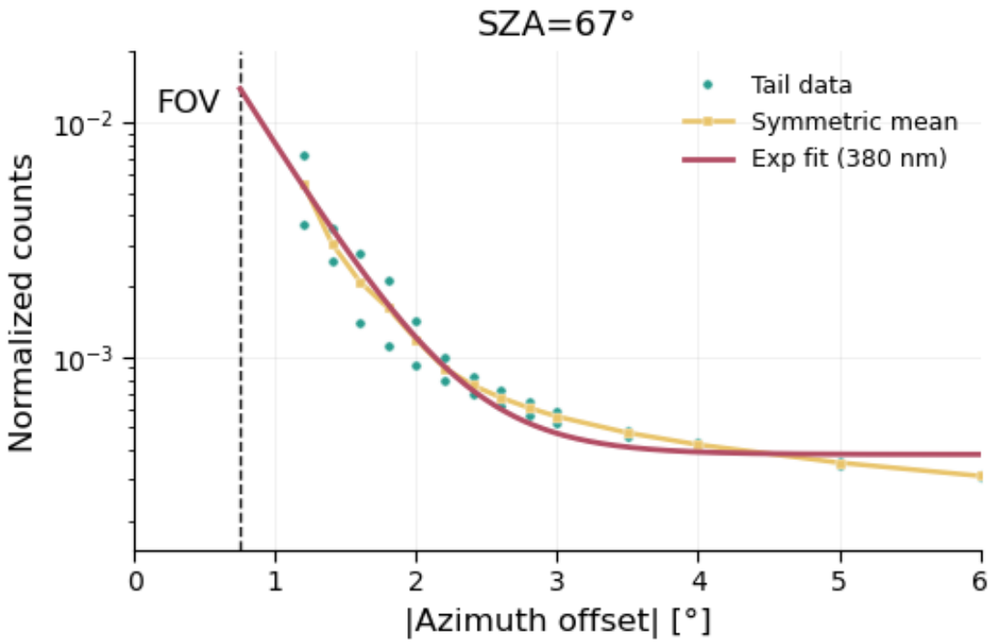


Figure 3.2-24: The log-scaled normalized radiance for the 05.03.2026 at SZA=67°. The tail data show the actual measurements of the left and right flank, the yellow line, the symmetric mean and the red line the respective exponential fit, shown for 380 nm.

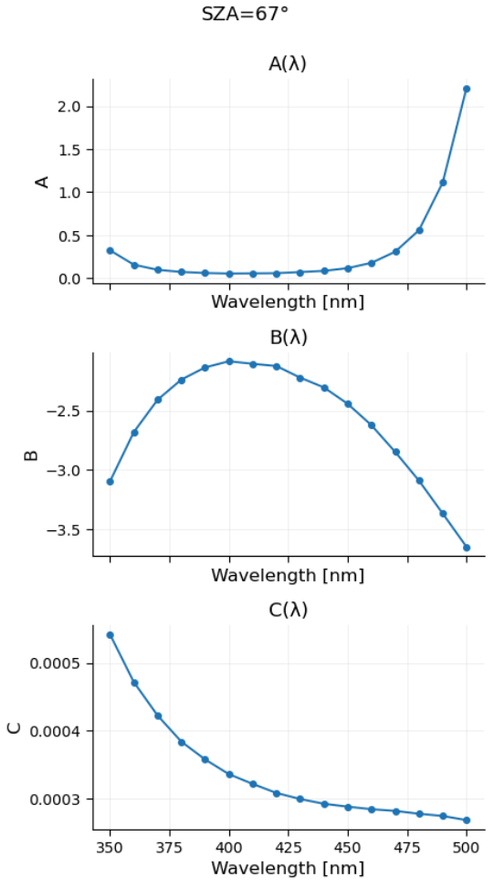


Figure 3.2-25: Exponential A,B, and C parameters for Pandora L1 data at 05.03.2026 over the spectral wavelength range (nm).

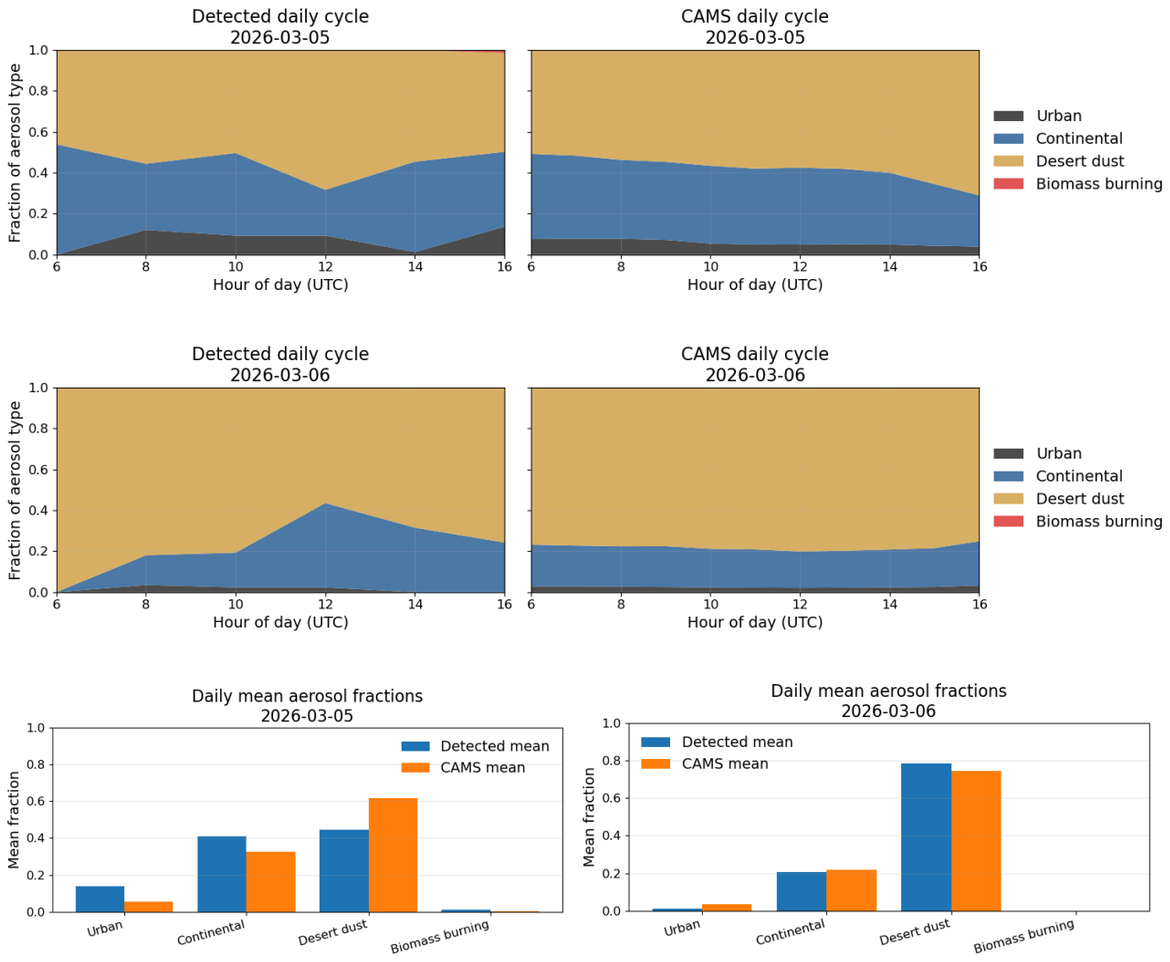


Figure 3.2-26: Detected daily cycle of the aerosol type (upper left panel) and daily cycle of aerosol type from CAMS (right upper panel) in Innsbruck at 05.03.- 06.03.2026, respectively. The lower panel shows the daily mean fraction at 11.03.2026 for 4 aerosol types: urban, continental, desert dust and biomass burning, respectively.

3.2.9.3 Case Study biomass burning 11.03.2026 Thessaloniki

We also performed a case study of a biomass-burning event in Thessaloniki on 11.03.2026 as a proof of concept to assess whether the aerosol type detection algorithm is able to capture biomass burning events. Figure 3.2-27 shows that the biomass-burning concentration from the CAMS European air-quality analyses increased over the course of the day in the region of Thessaloniki, with the strongest signal occurring in the afternoon and evening. The satellite-based fire detections from various satellite instruments further support the presence of active fire sources in the surrounding area.

Figure 3.2-28 further shows that the detected aerosol type agrees overall well with the CAMS daily cycle. In both datasets, the aerosol composition is dominated by continental aerosol, while urban and desert dust contributions remain secondary. At the same time, the retrieval indicates a small but distinct biomass-burning fraction, which is also reflected in the CAMS data, although with weaker temporal variability. On a daily mean basis, the biomass-burning contribution remains comparatively small, but its occurrence is consistent with the regional fire activity and the enhanced CAMS biomass burning concentration.

Conclusion

Overall, the SLSPAT framework provides a consistent approach to derive DIFF_CORR from measured radiance features, while simultaneously constraining aerosol properties within a physically based LUT inversion.

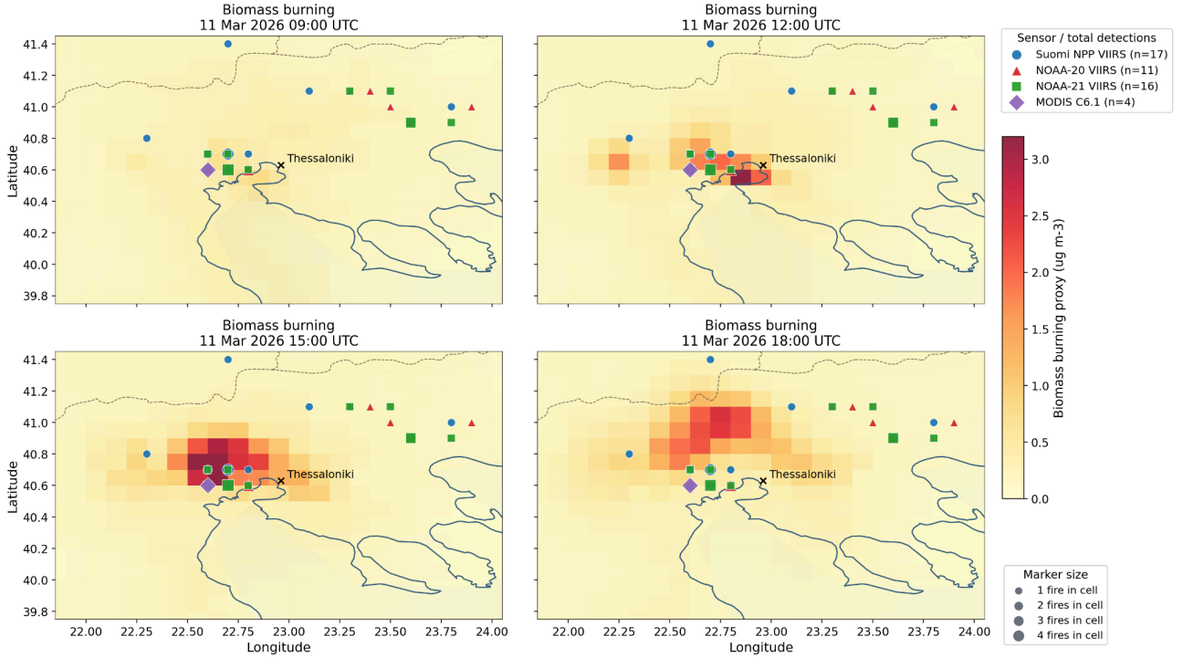


Figure 3.2-27: Concentration of biomass burning CAMS European air quality analyses with a resolution of (0.1 degrees, approx. 10km) (colours) at 11.03.2026 in the region of Thessaloniki. Markers represent fire activity measured by the satellite instruments Suomi NPP, NOAA-20, NOAA-21 and MODIS C6.1, marker sizes are a proxy for the number of fires per grid cell.

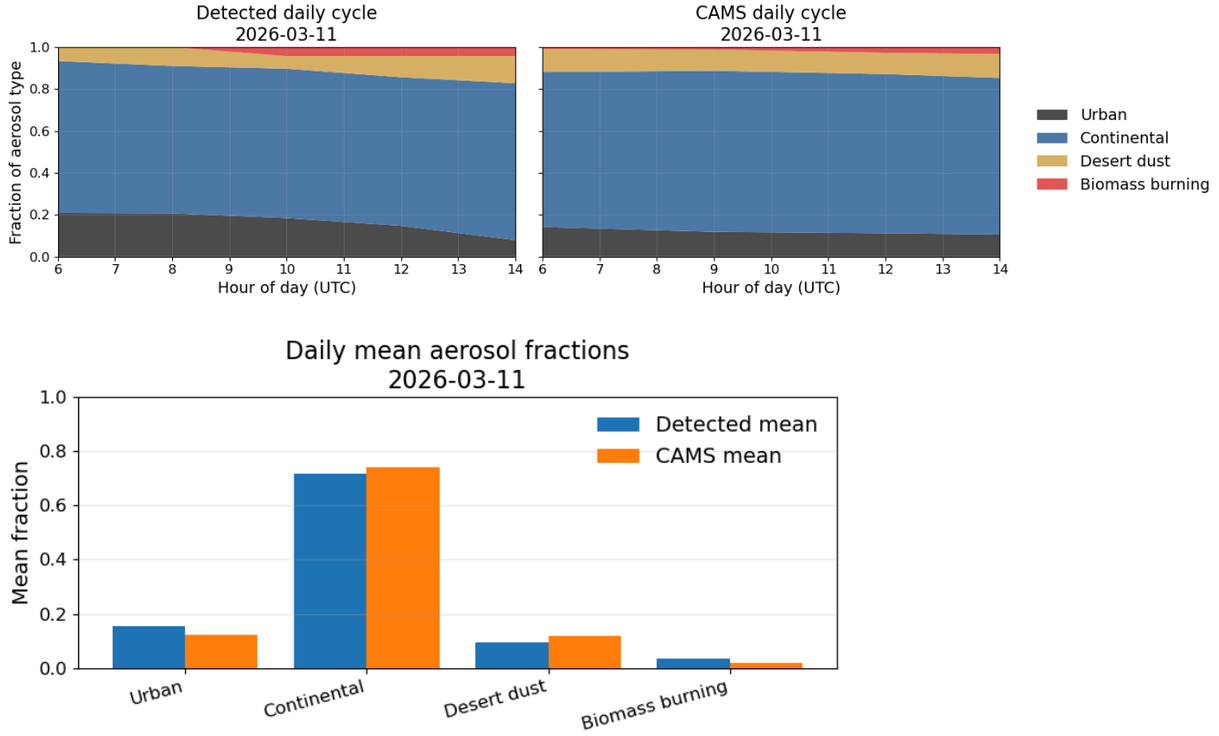


Figure 3.2-28: Detected daily cycle of the aerosol type (upper left panel) and daily cycle of aerosol type from CAMS (right upper panel) in Thessaloniki. The lower panel shows the daily mean fraction at 11.03.2026 for 4 aerosol types: urban, continental, desert dust and biomass burning.

4 Validation Studies

4.1 Validation Plan (WP1.5)

4.1.1 Introduction

Accurate validation of satellite-derived atmospheric constituents is critical to ensure the reliability of satellite products for scientific and regulatory applications. Ground-based measurements from the Pandora Global Network (PGN) can serve as high-quality reference data for total column retrievals of species such as ozone (O3), nitrogen dioxide (NO2), and formaldehyde (HCHO), sulfur dioxide (SO2) and water vapor (H2O).

Traditionally, validation efforts have focused on direct 1:1 comparisons between satellite and PGN total column values at individual sites. While this provides baseline performance metrics (e.g., bias, correlation, root mean squared error and standard deviations), it fails to account for systematic errors that emerge due to changes in the retrieval settings, calibration status or instrumental stability for instance. The goal of this sub-WP is to develop a modern, interpretable validation framework that:

- Integrates a range of feature variables to explain potential differences (e.g., SZA, aerosols, retrieved resolution changes, calibration file, interactions with other retrieval products),
- Models the temporal and spatial patterns,
- Is generic as possible to be used for various parameters of interest,
- Moves beyond simple station-wise comparisons,
- Distinguishes between instrument-specific / network effects versus the satellite product.

To achieve this, Generalized Additive Models (GAMs) (Hastie and Tibshirani, 1986) are chosen as a best-practice starting point, which provide a flexible and transparent statistical foundation for multivariate and nonlinear modeling. This approach depicts an explainable, and extensible framework for assessing satellite-PGN differences. By integrating calibration file identifiers, observational geometry, seasonal cycles, geographical context, and internal PGN diagnostics such as quality codes and effective temperature proxies, we can move beyond correlation statistics to a mechanism-oriented understanding of bias drivers.

This allows for interpretable diagnostics for individual instruments, network-wide validation consistency checks, explicit modeling of influential effects by feature variables, capacity to detect both instrument drift and satellite degradation, inclusion of advanced quality indicators.

GAMs are chosen as a first approach in particular due to their flexibility and interpretability. Unlike black-box methods such as deep neural networks, GAMs provide direct insight into the functional relationships between covariates and residuals, making them ideal for the validation purpose.

Nevertheless, to further enhance this framework, several other techniques may be considered along the project duration:

- Principal Component Analysis (PCA): Useful for reducing high-dimensional input (e.g., vertical profiles,spectral information from multiple wavelengths) into orthogonal components if a huge range of feature variables are available.
- Distributional Regression with GAMLSS2: A recent advancement allowing not only the location (mean) but also scale (sigma) and shape (asymmetry) parameters to be modeled as smooth functions. This is particularly relevant for satellite validation, where heteroscedasticity or skewed residual distributions may occur. The GAMLSS2 framework also enables boosting for variable selection.

This modular and expandable structure ensures the validation framework remains relevant as data volumes grow, retrieval algorithms evolve, and new instrumentation is added to the validation framework.

4.1.2 Overall Strategy and Modeling Concept

4.1.2.1 Core Concept

The primary validation metric *Y* is the difference between the satellite and the ground based instrument measurements:

$$Y_{i,t} = Satellite_{i,t} - PGN_{i,t}$$

Where i denotes a specific PGN location and t is the time index. On how to select both accordingly, we follow the guidelines for temporal and spatial co-location criteria as already applied the S5P Mission Performance Center (Verhoelst et al. (2015),Garane et al. (2019)):

The GAM model used within the validation plan aims to explain this difference using both smooth functions of continuous features and categorical feature variables that account for instrument states.

Based on a baseline validation framework (VF) with baseline features, a continuous evaluation of the primary validation metric is performed to identify new features. This automatically triggers new spin-off studies to improve the validation metric and features selection. Ultimately, new features can be incorporated into the validation framework which therefore can grow over time, as illustrated in Figure 4.1-1.

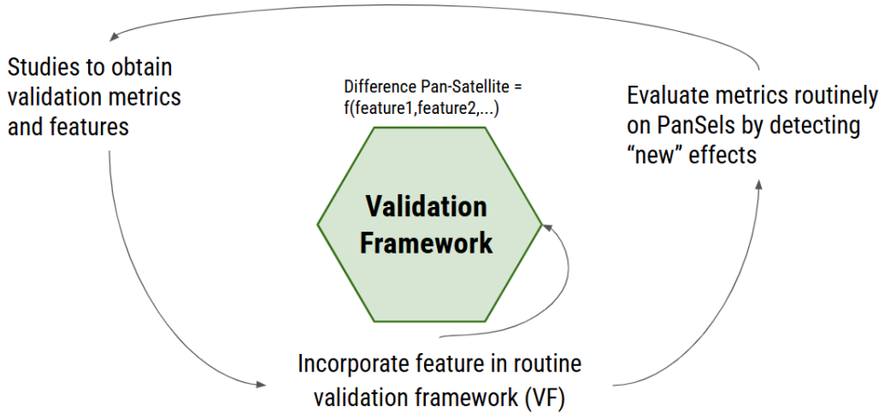


Figure 4.1-1: Validation concept illustration.

4.1.2.2 Key Modeling and Validation Stages

Stage 0:Instrument-Specific Modeling

GAMs are first fit individually to each PGN instrument to intrinsically check the consistency of a data product (e.g. direct sun total column O_3 amounts). This is based on de-seasonalized timeseries:

- Diagnose local patterns seasonalities
- Attribute differences to feature variables like solar zenith angle (SZA), day of year (DOY) and others as listed in Table: Smooth Features and Table: Categorical Features.
- Assess the impact of different calibrationfiles used for the retrieval.

Stage 1: Instrument-Specific Difference Modeling

Similarly to stage 0, GAMs are fitted individually to each PGN instrument, but using the primary validation metric *Y* as response variable instead of a data product itself.

Stage 2: Network-Wide Difference Modeling

A global GAM is constructed using data from all available PGN sites to:

- Quantify spatial and regional effects,
- Evaluate whether the satellite product exhibits regional bias or temporal drift.

This three-stage structure balances sensitivity to individual instrument behavior with the ability to detect systematic network or satellite issues.

4.1.3 Feature variables

The following Tables provide an overview of smooth and categorical features variables, which can grow over the project duration. SZAand DOYwill serve as baseline feature variables, as well as the calibrationfile.

Table: Smooth Features

Variable	Description	Purpose
SZA	Solar zenith angle	Observational geometry effect
DOY	Day of year	Seasonal pattern (modeled cyclically)
O2-O2 column	Slant column of O2-O2	Proxy for cloud contamination
Altitude	Station elevation	Influence on air column measurements
Latitude, Longitude	Geographic location	Detect regional differences
Ozone effective temperature	Retrieved gas temperature proxy	Sensitive to instrumental changes and retrieval bias
Quality Code	Internal PGN quality indicator	Encodes resolution changes for different spectral regions

Table: Categorical Features

Variable	Description	Role in Model
Calibrationfile	Identifier for retrieval version	Factor for detecting calibration-induced bias
PanSel group	Identifier for instrument group	Factor for detecting biases related to PanSel group, such as remote or urban sites

In particular, the quality code displays shifts in instrumental resolution from laboratory to field conditions. These internal PGN indicators serve as crucial input features for understanding variations in the satellite–PGN difference, particularly under changing operational conditions.

In addition to total columns, direct sun products also provide gas effective temperatures, which can also be used as proxies for retrieval sensitivity. These are particularly useful in diagnosing instrumental characterization errors or subtle retrieval biases. As illustrated in Figure 4.2-2, O_3 effective temperatures as part of the PGN data product, will serve as an explanatory feature variable, in particular for O_3 difference modelling in Stage 1.

4.2 Validation Framework (WP1.5)

4.2.1 Backend design for comparison data

In order to perform an efficient validation with external data, such as satellite data, the backend design and workflows on how to access these auxiliary data, is crucial. Ready to use tools for accessing satellite data, together with pre-processing towards Pandora specific locations, already exist, e.g. HARP Atmospheric toolbox, and have been investigated and tested. However, it does not avoid harvesting the needed orbit files and transforming them into the HARP compliant format, which

is the basic requirement. Moreover, it is limited to a subgroup of satellite data, and not even including other types of dataset families, such as reanalysis datasets (MERRA-2, ERA-5).

Therefore, the envisioned validation framework includes a customized system prone towards the validation use-case. It includes a scalable auxiliary-data backend designed to generate Pandora-site-linked parameters that can be compared with Pandora measurements. The system is intentionally organized as a generic pipeline rather than as a product-specific script: source data are discovered, downloaded only when needed, spatially reduced to site-relevant subsets, transformed into harmonized observations, and then stored in a query-efficient database layout. Although Sentinel-5P is a key example, the same architecture also supports other families such as MERRA-2 and AERONET, which highlights the extensibility of the approach.

A central design choice is the separation between transient source data and persistent analysis-ready products. The representative example is the processing of one Sentinel-5P overpass for a single Pandora site. The backend does not directly transform the full satellite swath into the final comparison product. Instead, it applies a staged spatial-reduction workflow. First, the raw overpass is cropped to a circular region around the target site, producing a compact site-specific subset, as illustrated in Figure 4.2-1 Sentinel 5P cropping of multiple sites and Figure 4.2-2 Sentinel 5P subset crop for Innsbruck. Second, the subset is processed to identify the final pixel or observation used for the site-level record. This two-step design is important for scalability: the expensive handling of the large original swath is performed only once, while the smaller subset can be reused for repeated analyses, testing, and visualization.

For long-term storage, the backend uses Parquet as the main format for curated daily observations. This is a strong architectural decision because Parquet is columnar, compact, and highly scalable for analytical workloads. The data are partitioned by dataset family, site, processing configuration (e.g. nearest pixel or pixel averages), and date, which enables fast selective reads without scanning the full archive. DuckDB complements this by providing lightweight metadata management and SQL-based access to the Parquet holdings. Together, Parquet and DuckDB create a backend that is both storage-efficient and fast to query, while remaining simple to deploy and easy to extend.

Another important strength is harmonization. The backend converts family-specific outputs into a shared auxiliary-observation structure with stable site identifiers, processing metadata, and provenance links. This makes the system suitable not only for one-off comparisons with Pandora, but also for broader validation studies. The same storage and query model can therefore support different data products and families without changing the overall backend design.

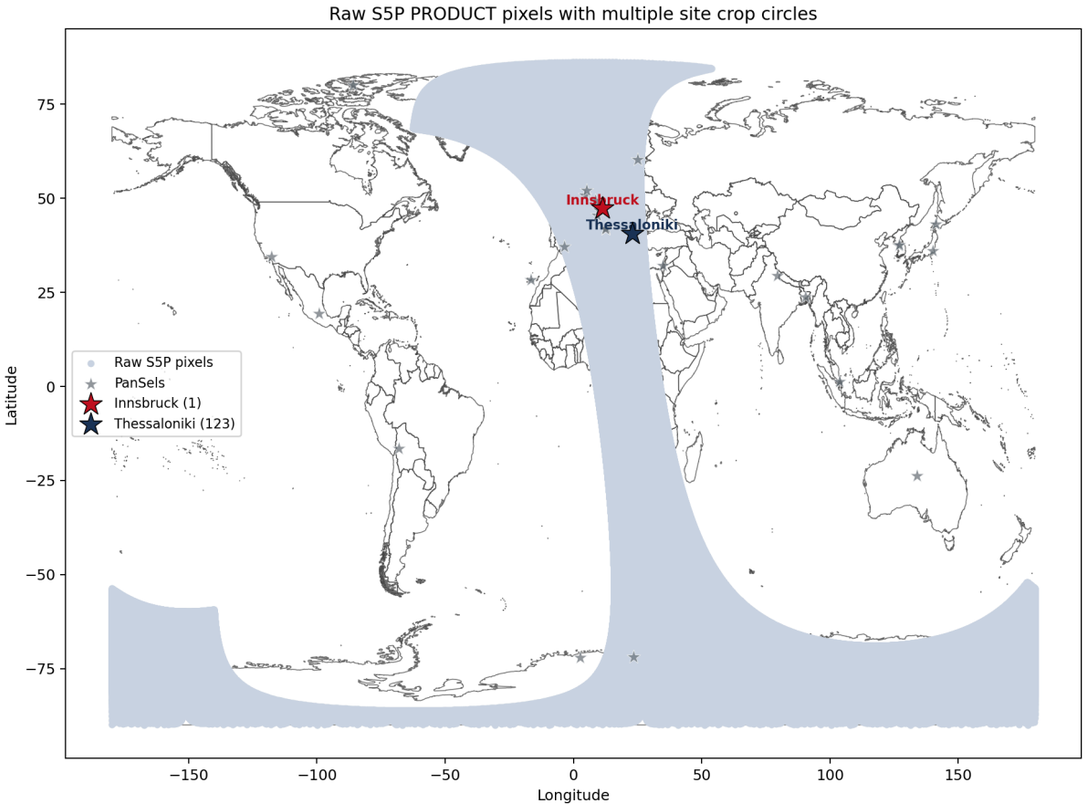


Figure 4.2-1 Sentinel-5P cropping of multiple sites

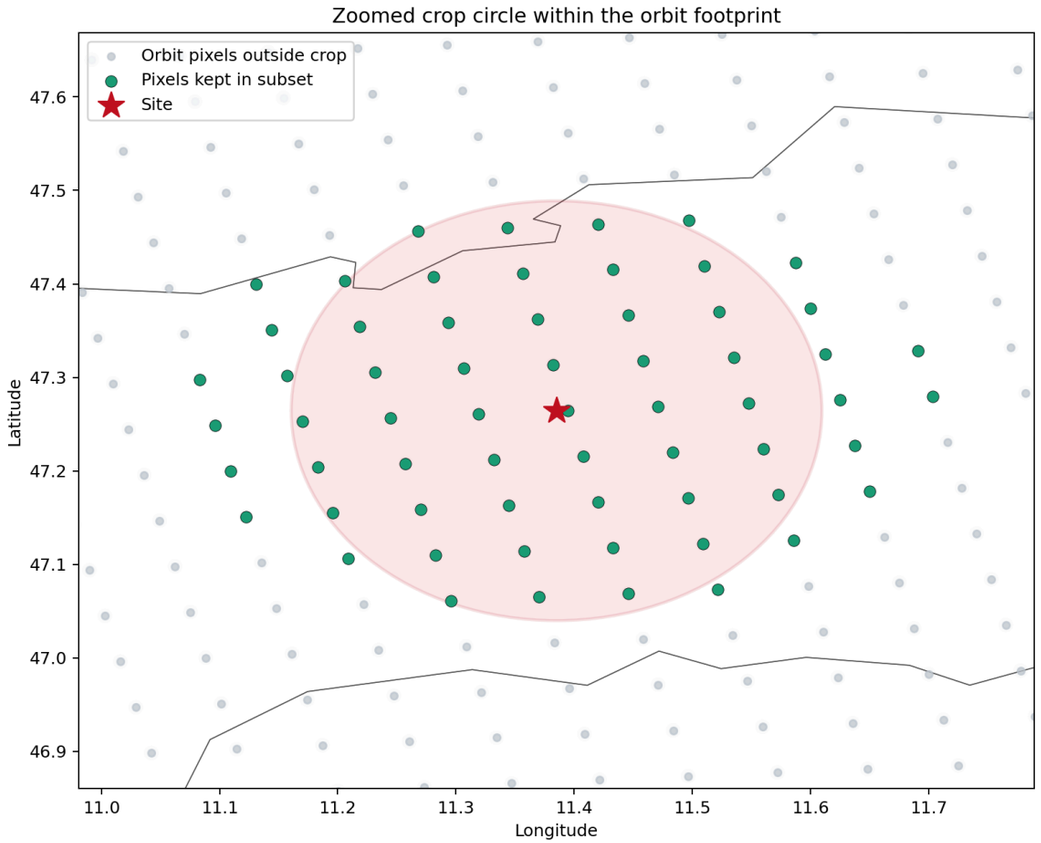


Figure 4.2-2 Sentinel-5P subset crop for Innsbruck using at 25km circle for reusable analysis.

4.2.2 Use Cases and Diagnostics

This section is summarizing diagnostics of the validation framework, illustrated exemplarily using research datasets for total column O_3. These datasets have been created as part of a pilot study to test the applicability of the GAM approach within the validation context.

4.2.2.1 Instrument Outlier Detection

Mission requirements and target performances, such as 3% level for Sentinel 4 O_3 column, can already be used to detect individual outliers. As illustrated in Figure 4.2-3, the period after November 2024 is running out of the 3% level. While standard correlation metrics as the R^2 value can indicate a good agreement, standard linear regression coefficients such as bias and slope cannot give deeper insight into the actual problem. In the case study for P196s1, the difference arises from a climatological O_3 temperature used in the retrieval setting. The impact of this, and the corresponding effect, can be captured by a GAM, as illustrated in Figure 4.2-4.

- Compare station-specific coefficients and residual patterns.
- Highlight instruments or periods with persistently high residuals, which fall outside reported uncertainties or mission requirements

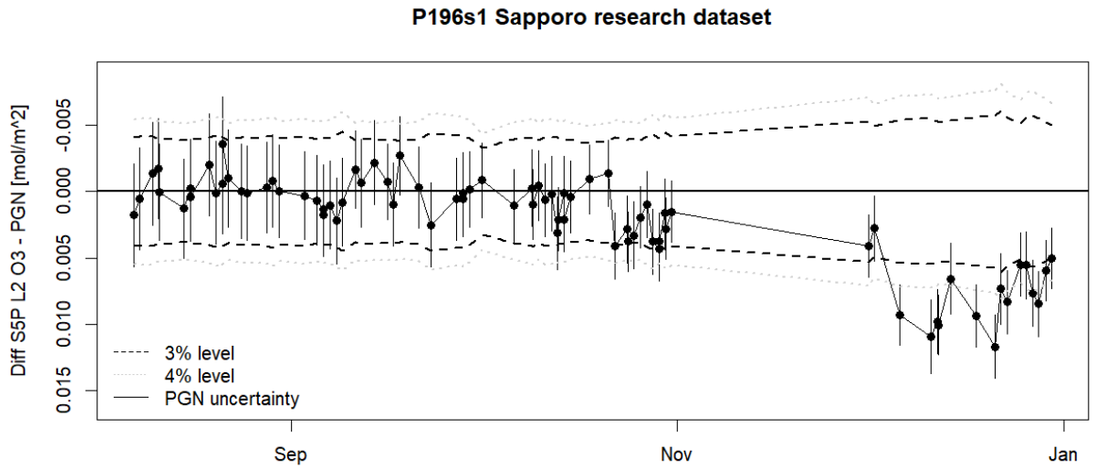


Figure 4.2-3: O_3 difference between S5P and P196s1 for research dataset at Sapporo.

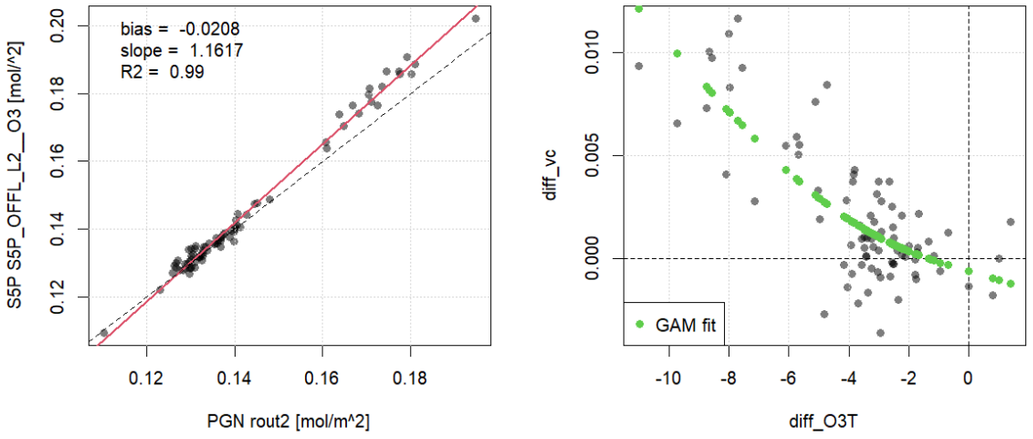


Figure 4.2-4: Scatterplot TotO_3 comparison (left) and effect of O_3 temperature difference (right).

4.2.2.2 Calibration Impact and Temporal Analysis

Update in an instrument calibration can cause unforeseen structural changes in the time series of a data-product. While this does not necessarily change the overall seasonal pattern, systematic jumps in the seasonality can show up as well.

By de-seasonalizing a variable, the residuals should reveal no temporal changes or drifts. This is a first consistency check as part of Stage0, which is illustrated in Figure 4.2-6. The corresponding seasonality is expressed by a GAM fit into the day of year in Figure 4.2-5. Jumps can be detected by a breakpoint analysis, as used within the PGN QAQC checks (see technical note of ESA project FRM4AQ). However, included as a feature variable within the GAM framework, such a coefficient would directly quantify and report the source of the problem.

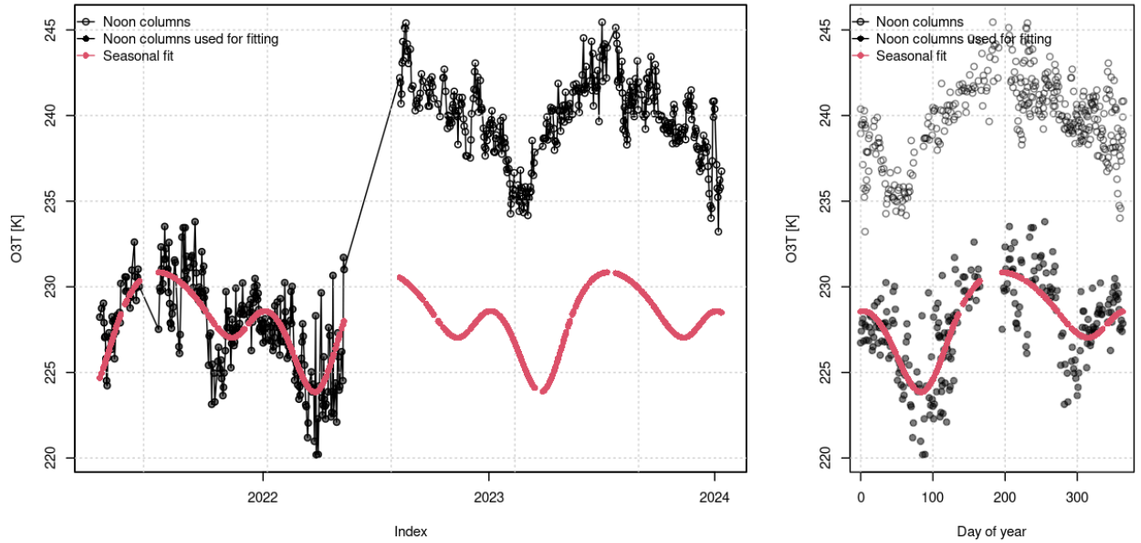


Figure 4.2-5: O_3 temperature time series at P193s1 Tsukuba (left) and the seasonality expressed by GAM fit on the day of year (right).

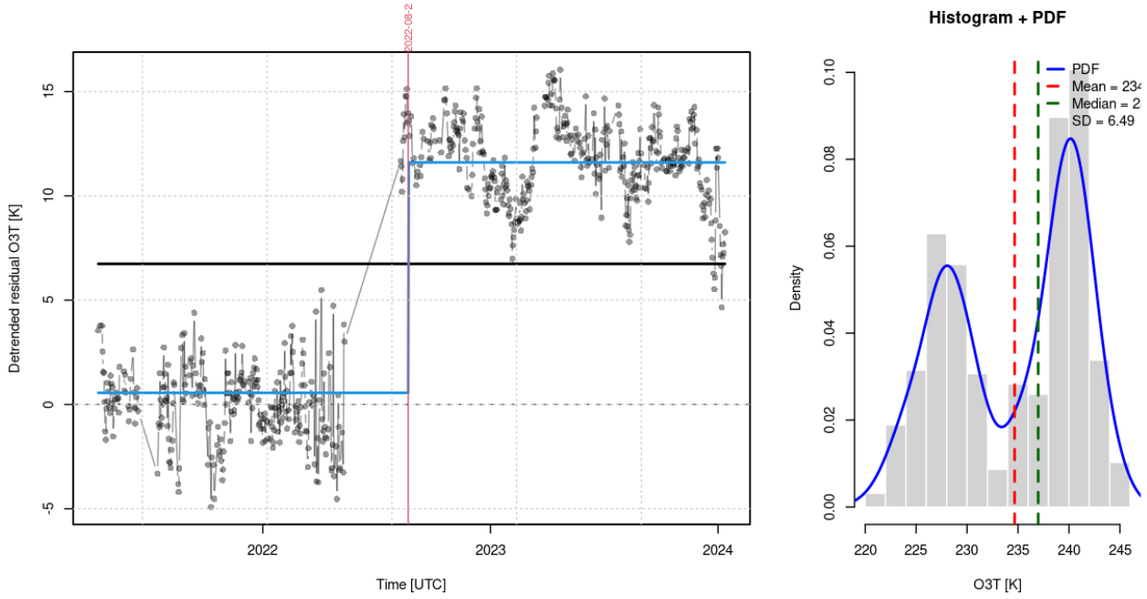


Figure 4.2-6: De-seasonalized O_3 temperature time series at P193s1 Tsukuba plus an added breakpoint analysis (left); Histogram of O_3 temperature (right) highlighting the two regimes.

As part of Stage1, targeted variables such as Sentinel-5P total column O_3 can be validated with instrument-specific timeseries, as illustrated in Figure 4.2-7 Comparison of TotO3 for P182s1. The same core methods used in Stage0 to detect structural changes, such as breakpoint analysis, can be applied to these time series to identify instrumental drifts or sudden shifts. In addition, explanatory feature variables, such as day of year and solar zenith angle, can already help to highlight systematic differences. Observed patterns, such as the seasonal wave visible in Figure 4.2-7 Comparison of TotO3 for P182s1 (topright) can trigger deeper re-investigations of retrieval setups on both sides, the satellite and the ground-based. In particular, if observed differences leave the envisioned mission requirements.

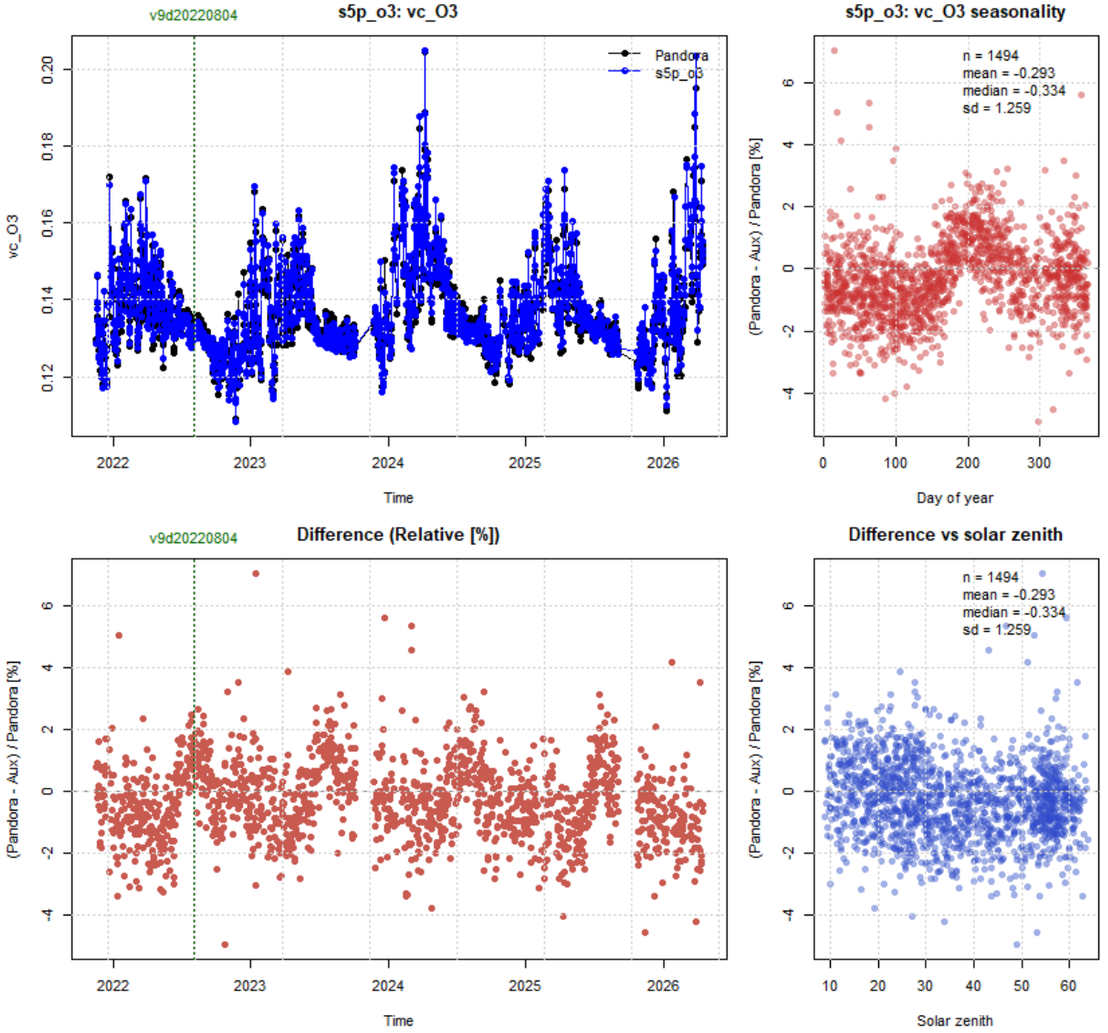


Figure 4.2-7 Comparison of TotO3 for P182s1 at Tel-Aviv with S5P: timeseries for Pandora and S5P (topleft), relative difference as a function of day of year (topright), time (bottom left), and solar zenith angle.

4.2.2.3 Grouped Pattern Detection

The selected PanSels represent specific site groups, including pollution-related categories such as remote, urban, and suburban sites, longitude/latitude-based groups, and co-located sites. Co-located sites are particularly relevant because they enable intra-pixel comparisons, where co-located PGN instruments are expected to provide consistent data products. Based on the instrument-specific differences identified in Stage1, enhanced pattern-detection methods can be applied to derive overall and systematic differences to the network, while also identifying individual sites that may require further review regarding their applicability.

Figure 4.2-8, showing a sub-network example for $TotO_3$ validation of Sentinel-5P, presents an intermediate time series for Pandora 240 as part of an ongoing calibration process. In this example, clear deviations from the behaviour of the remaining sites can be observed. In addition, the difference between Pandora 240 and Sentinel-5P clearly exceeds the 3% band in the seasonal view, meaning that such a dataset would be immediately detected by the proposed approach. Most importantly, the differences to the network are well within the predefined boundaries in this example.

Once Stage2 is fully established and all related target variables are available and compared, the defined groups can be incorporated into the GAM approach as categorical features to quantify discrepancies among them.

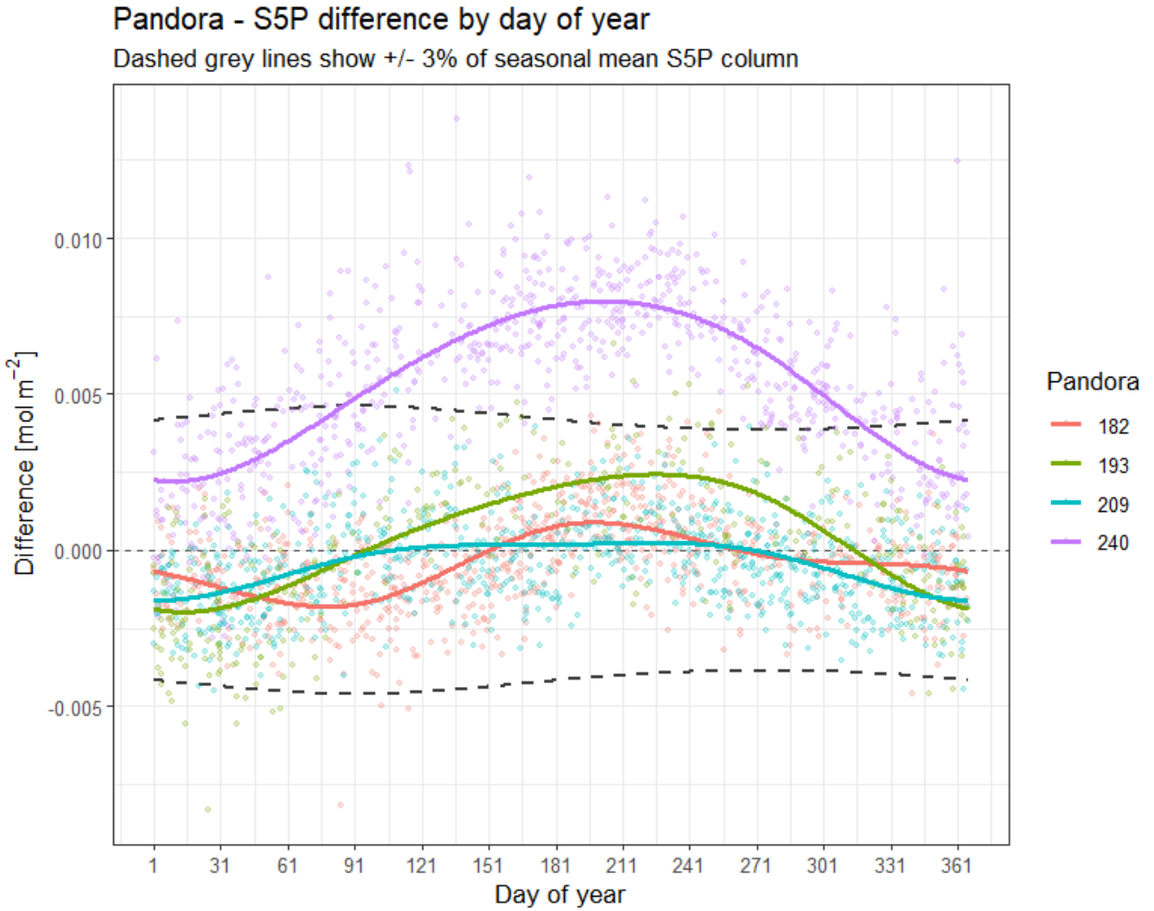


Figure 4.2-8 Sub-network example of $TotO_3$ compared to Sentinel-5P.

5 Appendix A Detailed Description of Reduced Running Langley concept for Spectral AOD Retrieval

The goal is to estimate a time-dependent extraterrestrial calibration constant I0(t) directly from Pandora clear-sky direct-sun measurements, and then use that time-varying calibration to retrieve aerosol optical depth (AOD). In this approach, AERONET is not used to derive the calibration itself. It is used only afterward as an external validation reference.

5.1 A.1 Physical Basis

Direct-sun measurement equation

The working measurement equation is:

I = I0 * f * exp(-(tau_aero * m_aero + tau_ray * m_ray + tau_gas * m_gas))

where:

- I is the measured direct-sun irradiance or signal
- I0 is the extraterrestrial calibration constant
- f is the Earth-Sun distance factor
- m_aero, m_ray, m_gas are aerosol, Rayleigh, and gas airmasses
- tau_aero, tau_ray, tau_gas are the corresponding optical depths

Rearranging gives the aerosol optical depth:

tau_aero = [-ln(I / (I0 * f)) - tau_ray * m_ray - tau_gas * m_gas] / m_aero

Langley transformation

For the Langley fit, the known extinction terms are moved back into the measured signal:

I_reduced = I * exp(tau_ray * m_ray + tau_gas * m_gas)

This gives the linear Langley relation:

ln(I_reduced / f) = ln(I0) - tau_aero * m_aero

The fitted quantities therefore have a direct physical meaning:

- x-axis: m_aero
- y-axis: ln(I_reduced / f)
- intercept: ln(I0)
- slope: -tau_aero
- therefore I0 = exp(intercept)

The key assumption is that aerosol optical depth is approximately constant during each selected clear-sky fitting segment.

5.2 A.2 Running Langley Concept

A classical Langley calibration estimates one I0 from one carefully selected calibration period. The running Langley method extends this idea in time:

1. perform repeated Langley fits on many clear AM and PM half-days
2. estimate one I0 value per half-day and wavelength
3. combine AM and PM information into daily values
4. smooth the resulting daily I0 series in time
5. interpolate the smoothed I0(t) onto the measurement timestamps
6. recompute AOD using the time-varying calibration

The result is a calibration timeline rather than a single fixed calibration constant.

5.3 A.3 Method Steps

Clear-sky selection

Before fitting, the L1 time series is cloud-screened. The implementation uses a clear-sky mask derived from the direct-sun signal, and only the retained subset enters the Langley analysis.

This step is essential because the running Langley method assumes that the fitted signal variations are dominated by air mass and not by cloud transients.

AM/PM segmentation

Each day is split into two half-days:

- AM
- PM

The split is defined by the daily minimum aerosol airmass, which acts as an effective noon. Measurements before that minimum belong to the AM segment and measurements after it belong to the PM segment.

Build the Langley signal

For each wavelength, the measured L1 signal is converted to the Langley fitting quantity:

ln(I_reduced / f)

with:

I_reduced = I * exp(tau_ray * m_ray + tau_gas * m_gas)

This step removes the known non-aerosol attenuation terms and isolates the part of the signal that should vary linearly with aerosol airmass.

Fit one Langley line per half-day and wavelength

For each half-day and wavelength, a linear regression is performed between:

- x = am_aero
- y = ln(I_reduced / f)

Only segments with enough points and enough airmass span are allowed to produce a valid fit.

Point filtering

Before the regression, candidate points are filtered by a set of practical conditions, for example:

- maximum allowed aerosol airmass
- minimum number of valid points
- minimum airmass span
- optional residual-based outlier rejection

The purpose of these filters is to prevent unstable intercept estimates from sparse, narrow, or visibly contaminated segments.

Half-day fit quality scoring

Each half-day fit is assigned a soft quality score between 0 and 1, plus hard rejection tests.

The soft score is:

quality_weight = w_points * w_span * w_r * w_err * w_tau

where:

- w_points rewards sufficient retained points
- w_span rewards sufficient airmass coverage
- w_r rewards strong linear correlation
- w_err penalizes large intercept uncertainty
- w_tau penalizes implausibly large fitted aerosol optical depth

Hard rejection can remove a fit completely if, for example:

- the regression correlation is too weak
- the intercept uncertainty is too large
- the fitted aerosol optical depth is outside the accepted range
- the segment already failed because too few points survived filtering

Combine AM and PM into a daily value

After quality scoring, AM and PM fits are merged into one daily value per wavelength.

The logic is:

1. if only one side is valid, keep that side
2. if both AM and PM are valid and consistent, combine them
3. if they disagree but one side is clearly better, keep the better side
4. otherwise reject the day

The AM/PM consistency is checked on the relative I0 difference:

rel_i0_diff = exp(|intercept_AM - intercept_PM|) - 1

Smooth the daily intercept series

The daily quantity that is smoothed is not I0 directly, but the intercept series:

intercept = ln(I0)

Smoothing on the log scale is natural because the intercept is the direct regression output. The smoothing stage suppresses noisy day-to-day variability while preserving the slower calibration drift.

The implementation supports several smoothing options, including a weighted local linear smoother. In the weighted case, days with better fit quality receive higher influence, and days supported by both consistent AM and PM fits can receive additional weight.

Optional controls also exist for:

- forcing the smoother to restart at known break points
- limiting the allowed day-to-day relative drift

Convert the smoothed intercepts to I0(t)

After smoothing, the daily intercepts are converted back to calibration constants:

I0_smoothed = exp(intercept_smoothed)

These daily I0 values are then interpolated onto the original measurement timestamps to produce a continuous time-dependent calibration series I0(t).

Recompute AOD with time-varying calibration

The final AOD product is computed with the same retrieval equation as before, but using the interpolated time-varying calibration instead of one fixed I0:

tau_aero(t) = [-ln(I(t) / (I0(t) * f(t))) - tau_ray * m_ray - tau_gas * m_gas] / m_aero

This produces the running-Langley AOD time series.

5.4 A.4 Assumptions And Caveats

- The method assumes approximately stable aerosol optical depth within each selected AM or PM fitting segment.
- Cloud screening quality is critical. Residual cloud contamination directly biases the intercept.
- The current gas treatment follows the same practical simplification used elsewhere in the project. It is internally consistent, but still simplified.
- The AM/PM split is defined by the daily minimum aerosol airmass, which is a practical approximation of noon.
- Quality thresholds are heuristic and may need retuning for different sites, seasons, or instruments.
- A smoothing window that is too wide can hide real calibration jumps.
- If known instrument changes occurred, smoothing should be interrupted across those break times.
- The summarized correction table is only a compact representation of the running result, not a replacement for the full I0(t) series.

6 Appendix B Detailed Description of the Spectral AOD Retrieval

6.1 B.1 Physical basis and forward model

For one day and one spectrometer wavelength range, the retrieval starts from a broadband DOAS quantity P_smo(t, lambda). Operationally this may be the evaluated smoothing polynomial alone or

the smoothing polynomial with a low-pass component of the spectral residual added. The central assumption is that the aerosol slant optical depth scales with aerosol air-mass factor m(t), whereas a spectrally smooth instrumental contribution is additive in the fitted optical-depth domain.

$$P\_eff(t, \lambda) = P\_smo(t, \lambda) + \mathcal{L}[r(t, \lambda)],$$

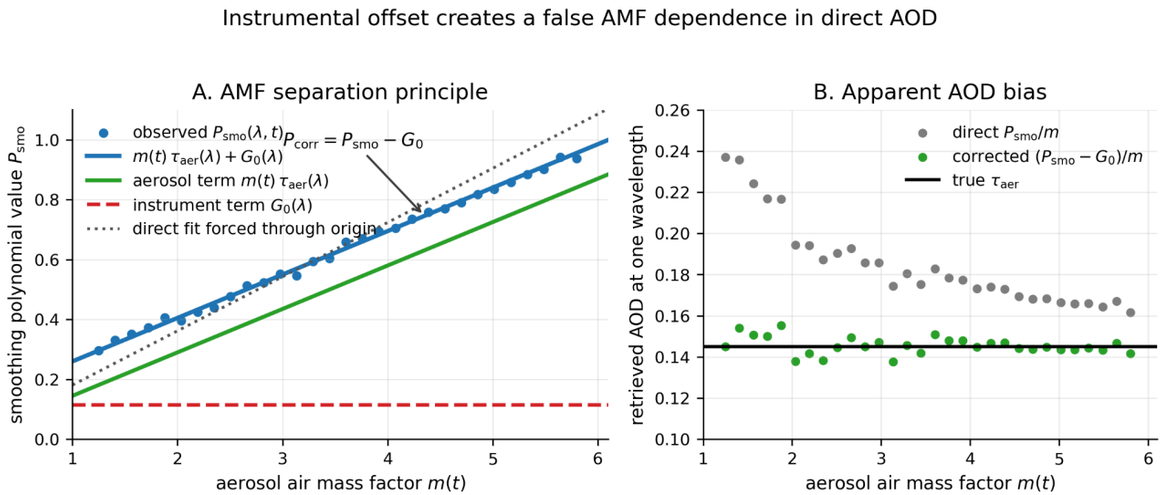
$$P\_eff(t, \lambda) = m(t) \cdot \tau\_aer(t, \lambda) + G\_inst(t, \lambda) + \varepsilon(t, \lambda),$$

The simplest instrument representation is a day-constant spectrum G0(lambda). A more flexible but still identifiable representation allows an additional grey temporal term g(t), constrained to zero mean over the day.

$$G\_inst(t, \lambda) = G_0(\lambda),$$

$$G\_inst(t, \lambda) = G_0(\lambda) + g(t), \quad with \quad \Sigma\_t\, g(t) = 0,$$

This anchoring is important: without the zero-mean constraint, a constant offset could be redistributed arbitrarily between G0(lambda) and g(t). The instrument term is not interpreted as aerosol; after division by air mass it would appear as an artificial dependence on 1/m.



6.2 B.2 Aerosol spectral parameterization

The aerosol spectrum is represented by a two- or three-parameter Angstrom-type expression. The three-parameter form includes curvature and is the practical default for the present retrieval configuration.

$$\tau\_aer(t, \lambda) = \exp[\beta_0(t) + \beta_1(t)\, L + \beta_2(t)\, (L^2 - \langle L^2 \rangle)], \quad L = \ln(\lambda/\lambda\_ref).$$

In the usual physical interpretation, β_1 is the negative Angstrom exponent and β_2 describes spectral curvature. The numerical fit uses a centred and scaled logarithmic wavelength coordinate to reduce collinearity between the linear and quadratic terms. This transformation does not change the physical model; it only improves conditioning.

Soft physical constraints are applied to keep the fitted AOD, Angstrom exponent, and curvature in a physically plausible range. Additional endpoint checks can be applied to the Angstrom exponent so that a quadratic spectrum cannot remain plausible near the reference wavelength while becoming unphysical at the edge of the fitted wavelength range.

6.3 B.3 Retrieval modes

- **Direct:** No explicit instrumental subtraction. The vertical AOD is obtained from the broadband DOAS term divided by aerosol air mass, followed by the constrained spectral aerosol fit. Appropriate when the instrument term is negligible or when the diagnostics do not provide reliable evidence for a separable instrumental contribution.
- **Polynomial-coefficient correction:** The DOAS smoothing polynomial coefficients are regressed against air mass. Their intercepts reconstruct a wavelength-dependent G0(lambda) before the aerosol fit. Most useful in the UV when the polynomial coefficients show a statistically coherent instrumental spectral shape.
- **Multiday G0:** Single-day polynomial G0 estimates from neighbouring days are combined with temporal and uncertainty weighting, then subtracted from the target day. Designed for persistent instrumental states where one day alone may be ambiguous but surrounding days provide a coherent estimate.

Direct retrieval

The direct mode assumes that any additive instrumental term is negligible or not reliably separable from the available measurements. The retrieval therefore fits the aerosol spectrum to the broadband term without an independent instrumental correction. In the limiting diagnostic view, the direct vertical AOD is simply P_smo divided by air mass.

$$\tau^*(t, \lambda) = P\_smo(t, \lambda) / m(t),$$

Direct retrieval is deliberately conservative. It is preferred when the instrument fingerprint is weak, when the direct slant relationship is already adequate, or when attempted separation would risk converting real aerosol temporal variability into an artificial instrument correction.

Polynomial-coefficient instrument correction

Pandora DOAS fits describe the broadband spectral term by smoothing-polynomial coefficients. The polynomial correction estimates the instrumental contribution in this coefficient space instead of fitting an independent intercept at every wavelength. If the evaluated smoothing polynomial can be written as a linear combination of spectral basis functions f_k(λ),

$$P\_smo(t, \lambda) = \Sigma\_k\ c\_k(t)\ f\_k(\lambda),$$

each coefficient is decomposed into an air-mass-scaling aerosol part and an air-mass-independent instrumental intercept.

$$c\_k(t) = A\_k\ m(t) + C\_k + \eta\_k(t).$$

The coefficient intercepts C_k are then mapped back into wavelength space as

$$G_0,poly(\lambda) = C_0 + \Sigma\_\{k=1\}^\{K'\}\ w\_k\ C\_k\ f\_k(\lambda),$$

The raw smoothing polynomial is corrected as

$$P\_smo,corr(t, \lambda) = P\_smo(t, \lambda) - G_0,poly(\lambda),$$

and the operational observable becomes

$$P\_eff,corr(t, \lambda) = P\_smo,corr(t, \lambda) + \mathcal{L}[r(t, \lambda)].$$

The key advantage is that $G_0(\lambda)$ is estimated as a spectrally coherent function constrained by the DOAS polynomial basis. This can recover wavelength-dependent instrumental structure that a set of independent per-wavelength intercepts would confound with the Angstrom aerosol shape.

The correction is stabilised by robust regression of the coefficient-air-mass relationships, by optional shrinkage of weak spectral-shape intercepts, and by limiting the polynomial order used to reconstruct $G_0(\lambda)$. When the robust regression indicates that the coefficient-air-mass model is not trustworthy, the correction can be suppressed rather than forcing an unstable subtraction.

Multiday G0 consensus

The multiday approach addresses days where an instrumental state is persistent but the single-day polynomial estimate is ambiguous. First, candidate G0_i(lambda) spectra are extracted for individual days using the polynomial-coefficient method. Only candidates passing feasibility and quality checks are admitted to the neighbour pool.

For a target day d0, admitted neighbouring days di are combined using a Gaussian temporal kernel and uncertainty weighting. The target day itself is excluded from the consensus estimate so that the correction is not evaluated against its own noisy estimate.

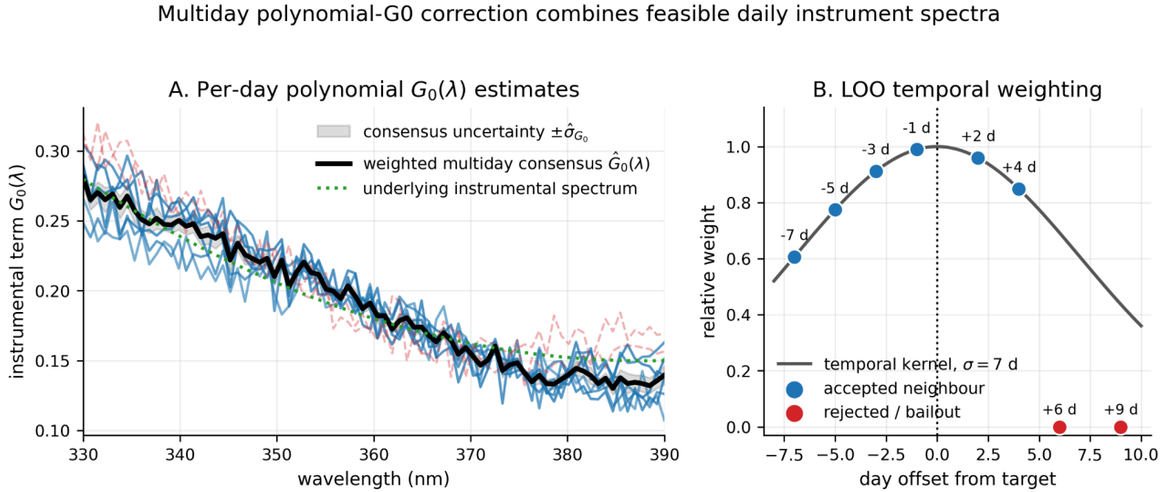
$$w\_i(\lambda)=exp(-0.5*((d\_i-d0) / \sigma\_d)^2)/ \sigma\_i(lambda)^2$$

$$G\^(\lambda) = \Sigma\_i\ w\_i(\lambda) * G_0\_i(\lambda)/ \Sigma\_i\ w\_i(\lambda)$$

The uncertainty used for the final decision is based on the weighted between-day spread of the admitted neighbour estimates, because day-to-day disagreement is the practical limitation on applying a consensus correction.

$$P\_eff,corr(t,\lambda) = P\_smo(t,\lambda)-G\^(\lambda)$$

A per-wavelength soft threshold may be applied to the consensus before subtraction. The correction is applied only when enough neighbours are available and the target day satisfies the relevant instrumental-evidence gates.



6.4 B.4 Feasibility and Quality Assessment

The retrieval does not assume that an instrumental correction is always beneficial. Each day and spectrometer wavelength range is first assessed for whether the information content is sufficient to separate aerosol and instrument. The most important diagnostic is the artificial dependence of direct AOD on inverse air mass, because an additive G0 term produces exactly this structure after conversion to vertical optical depth.

- **Air-mass imprint:** Correlation of direct AOD with inverse air mass. An additive instrument term in P_smo appears as an artificial 1/m dependence after division by air mass.
- **Direct-fit adequacy:** Through-origin fit P_smo proportional to m. High agreement indicates that the data are already consistent with an aerosol-only slant term.
- **Intercept evidence:** Improvement when an intercept is added to P_smo versus m. Quantifies whether an air-mass independent term is statistically useful.
- **Temporal stability:** AOD variability and morning-afternoon asymmetry. Guards against attributing real aerosol evolution or unstable sampling to the instrument.
- **Polynomial intercept significance:** Signal-to-noise of smoothing-coefficient intercepts. Tests whether the reconstructed spectral G0 is supported by the DOAS polynomial coefficients.
- **AOD and alpha after correction:** Residual dependence of AOD and Angstrom alpha on inverse air mass. A physically successful correction should reduce, not increase, air-mass dependence in both AOD and spectral slope.

In the UV, the polynomial correction is considered only when the air-mass fingerprint is present, the direct aerosol-only slant relationship is not already excellent, and the day is quiet enough for the coefficient regression to be meaningful. A high signal-to-noise ratio in the polynomial-coefficient intercept can provide independent evidence for a real spectral instrument term.

In the visible range, the polynomial correction is treated more conservatively because the wavelength interval and air-mass distribution can make the polynomial-coefficient regression less stable. Visible corrections are therefore normally controlled by additional gates based on the formal significance of the consensus G0, the target-day air-mass fingerprint, and the number of admitted neighbours.

For candidate G0 estimates, an additional physical consistency check evaluates whether subtracting the candidate correction reduces the air-mass dependence of both AOD and the two-wavelength Angstrom exponent. This is physically motivated: if the correction is removing an instrumental component, it should not make aerosol optical depth or spectral slope more dependent on viewing geometry.

6.5 B.5 Solver formulation

The retrieval is a separable nonlinear least-squares problem. The Angstrom parameters are nonlinear, while the instrument terms are linear for any fixed aerosol spectrum. This structure is exploited by solving the instrument part analytically inside the nonlinear fit.

The fast operational solver alternates between two steps. First, for the current aerosol spectrum it solves the linear instrument model analytically. Second, for the current instrument estimate it updates the Angstrom parameters by Gauss-Newton iterations in linear optical-depth space. This avoids the bias that can occur when negative or near-zero corrected optical depths are forced through a logarithm.

For one time step t, the Gauss-Newton residual is

$$r_t(\lambda) = P_{eff}(t, \lambda) - G_{inst}(t, \lambda) - m(t) \exp[\beta_t \cdot \varphi(\lambda)],$$

with Jacobian

$$J_t(\lambda, j) = -m(t)\, \tau_{aer}(t, \lambda)\, \varphi_j(\lambda).$$

This linear-space formulation avoids the log-space bias that would arise if negative or near-zero values of (P_eff - G_inst)/m had to be clipped.

The second solver is a full variable-projection formulation fit, in which the nonlinear Ångström parameters are optimized while the linear instrument terms are solved analytically at every cost-function evaluation. In compact form, the objective is

$$\min_\beta \; ||P_{eff} - m\, \tau_{aer}(\beta) - G_{inst}^*(\beta)||^2 + \lambda_{phys} J_{phys} + \lambda_s J_{smooth},$$

where G_inst*(β) denotes the analytic solution of the linear instrument subproblem for a given aerosol state. The physics penalty J_phys keeps AOD magnitude, Ångström exponent, and curvature within plausible bounds, and J_smooth denotes optional smoothness terms in time or instrument behaviour.

6.6 B.6 Practical interpretation and Limitations

- The separation is strongest when the day spans a useful range of aerosol air mass and the instrumental term is approximately additive in the DOAS broadband optical-depth domain.
- Aerosol temporal evolution can mimic part of an air-mass signature; therefore intercept-based corrections are accepted only when additional stability metrics support them.
- The polynomial correction is powerful when the smoothing coefficients contain a coherent instrumental spectral shape, but it can amplify noise if the coefficient-air-mass regression is poorly constrained.
- The multiday consensus improves robustness for persistent instrumental states, but it relies on neighbouring days being representative of the same instrumental behaviour.
- A physically credible correction should reduce artificial air-mass dependence in AOD and should not introduce an air-mass dependence in Angstrom alpha.

7 Appendix C Accompanying tasks and tools

7.1 C.1 Framework for running reduced Langley to on trace gas data

Within the PGN, linear atmospheric absorbers are calibrated using adjusted LE type techniques[1]. This involves selecting a measured spectrum as a reference and using adjusted LE to determine the magnitude of atmospheric absorption features in this reference. Depending on the absorber type, the analysis (=extrapolation to zero (=top of atmosphere), as a function of AMF) is performed either via irradiance (for AOD, see above Develop an LE method for L1 data) or slant columns (for linear trace gases like NO2, SO2 and HCHO). Ultimately, both techniques enable atmospheric absorption features (for AOD and gases) to be constrained in order to retrieve absolute data based on this reference spectrum.

For a perfectly stable instrument, the validity of this reference spectrum would remain valid: reapplying a LE to later data would produce the same results as initially found. However, in operational reality, instrumentation is subject to change, and consequently, the validity of reference spectra expires. If an LE were to be reapplied to data from a new instrumental state (with an unchanged reference spectrum), the results would be different. This would be a kind of translation of instrumental changes into changes in irradiances (AOD) or slant columns (for trace gases), respectively.

Due to the linear nature of the absorption processes considered here, the final results would be sufficiently similar if, for a changed instrument's state, a new reference spectrum were selected and calibrated, or if the new LE results were simply considered as a bias in the data products. This beneficial similarity paves the way for applying the LE in a kind of running mode (rLE), while keeping the actual reference spectrum the same.

[1] Due to the nature of the techniques it is most suitable for linear absorber. Non-linear absorbers like H2O, O2 and O3 use non-linearity aware calibration techniques.

rRedLE for linear trace gases The adjusted LE technique for trace gases is also a version of LE (RedLE) that reduces variability. In contrast to RedLE for L1 data (see Develop an LE method for L1 data), the trace gas RedLE reduces slant columns. In case of NO2 and HCHO, the tropospheric component[2] is reduced, and for NO2, also the stratospheric component[3]. No variability-reduction is applied for SO2.

In addition to the reduction (except for SO2), the LE is also applied to a larger number of days - usually a couple of weeks - to make the extrapolation more statistically robust. This modified reduced LE is therefore referred to as Red*M*LE. In running mode, it is referred to as rRedMLE.

Figure 3.1-4 shows two examples of RedMLE applications for Pandora 253 in Tucson, Arizona, for total column NO2. For this extreme example, the structural change at the beginning of 2025 can already be seen in the time series. The corresponding RedMLE plots show that the intercepts are almost zero for the first period but are biased low for the second. Please note that the time series expresses VCs, but the RedMLE data are given in SCs.

This is underlined by the full analysis of this dataset, which is shown in Figure 3.1-5 (note again that the intercepts are SCs). Here, only points that exceed the overlapping 1σ uncertainties from the initial calibration (gray bar) and the new estimation are considered significant. The intercept is always assigned to the last day of the window. The growing deviation from the initial calibration is evident for NO2 in the top figure panel. The rRedMLE for total HCHO (bottom panel) also reveals periods with biased data, but at different times.

As a first test application, we analyzed the rRedMLE for a subgroup of the PGN network for total column NO2, HCHO and SO2. To improve interpretability, we converted the SC into vertical columns (VC) using the minimum AMF for each day, then aggregated the data into quarterly mean values. The result is presented in Figure 3.1-6 and depicts the expected mean minimum change with correct

[2] The tropospheric columns are retrieved from a combination of sky radiance measurements at different elevations ("MAX-DOAS" mode).
[3] The stratospheric NO2 columns are estimated from a climatology put together by measurements from the OSIRIS instrument, a payload of the Odin satellite.

calibration. We consider this to be a potentially very useful overview for the data users, given the unavoidable delay in data quality assurance and related biases[4].

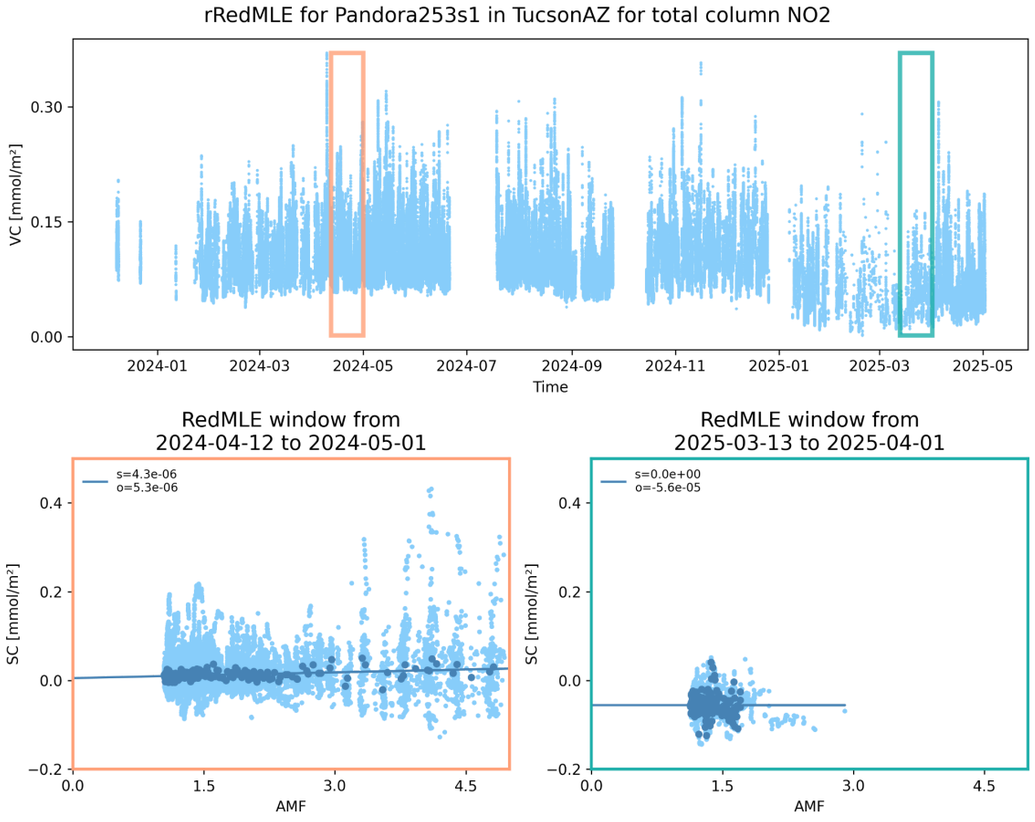


Figure 3.1-4: Two RedMLE examples for Pandora 253s1 in Tucson Arizona for direct sun total column NO2 in the bottom panel. The periods correspond to highlighted areas in the time series of VC in the top panel.

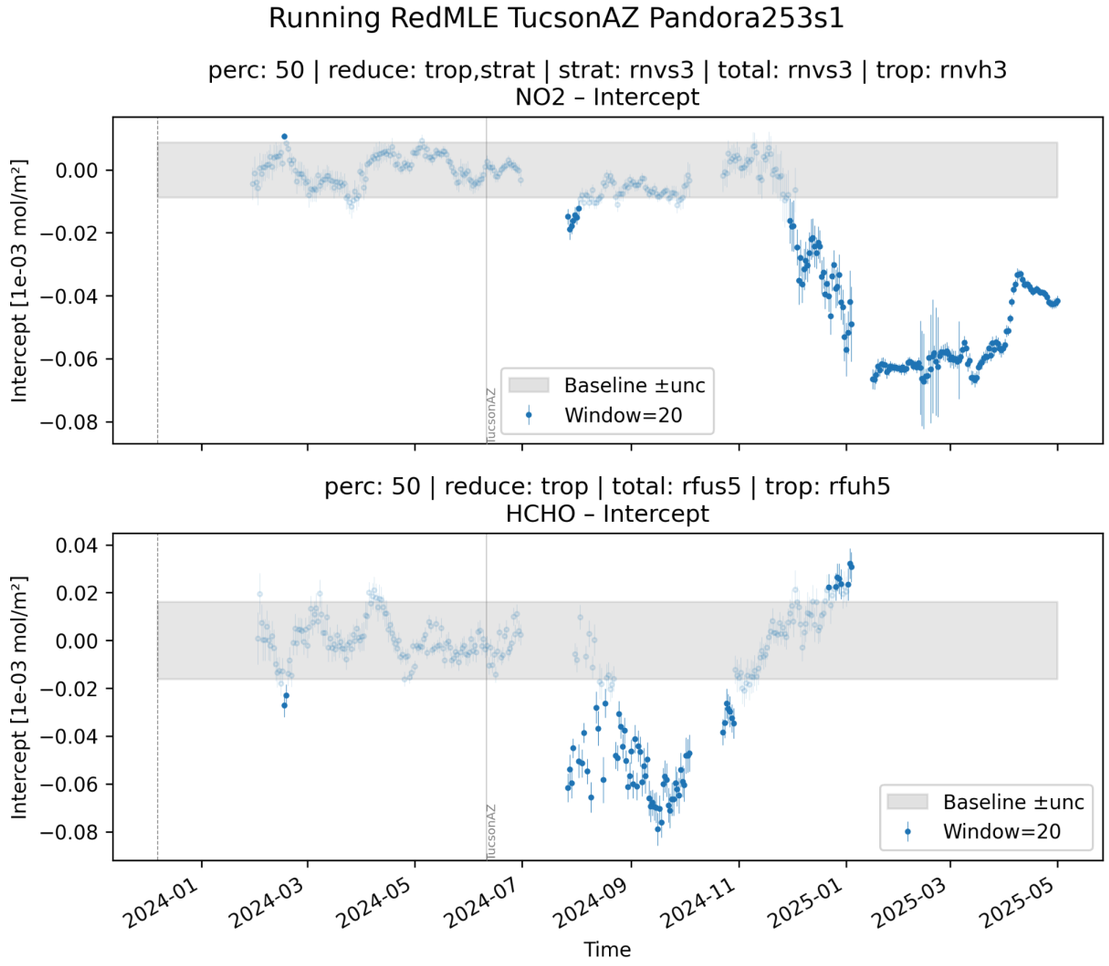


Figure 3.1-5: rRedMLE example application for Pandora 253s1 in Tucson Arizona for direct sun total column NO2 (top) and HCHO (bottom), using a 20 days running window. The gray shading reflects baseline uncertainty from the initial calibration.

[4] It should be noted that this is an expected result of the steep growing rate of the PGN network with basically constant calibration and QAQC staff.

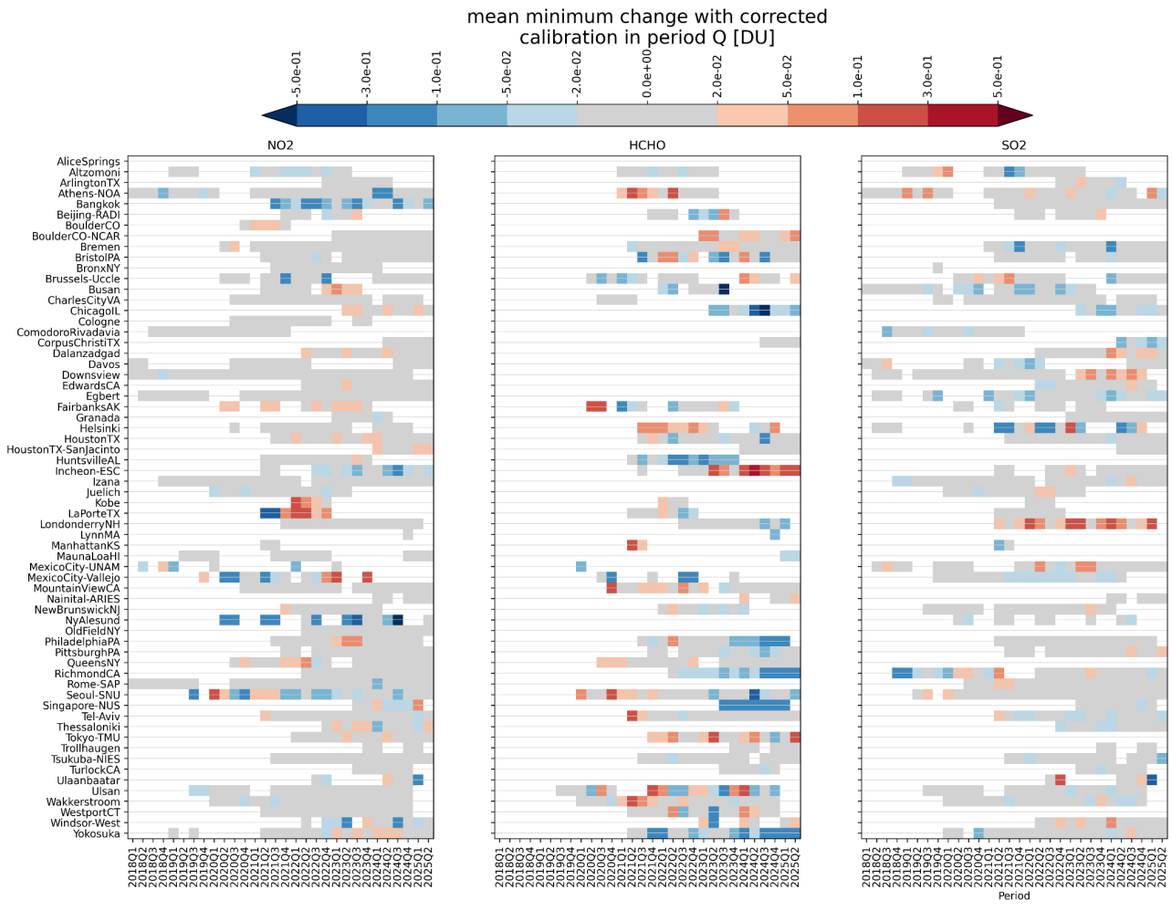


Figure 3.1-6: Application of rRedMLE to a subgroup of the PGN network (on EVDC) for total column NO2 (left), HCHO (center) and SO2 (right). For readability, the retrieved intercepts are aggregated into means per quarter.

7.2 C.2 Select PanSels for WP1.6

For this sub-WP we define a *PanSel* as a two channel (UV and VIS) Pandora-2S instrument that

- covers the spectral range from about 300 to 900 nm (considering both channels)
- operates side-by-side with an AERONET sun-photometer and, wherever feasible, an aerosol lidar
- samples different aerosol regimes from remote background atmospheres (e.g. high-altitude mountain or polar sites) to heavily polluted urban settings
- optionally delivers an already quality-assured multi-year data
- optionally has co-located Pandora-2S for internal cross-validation.

The list of PanSel candidates is as follows:

ID	Certified PGN location (lat °)	Location type	Available daily files*	QA %**	Main co-located instrumentation
121	Izaña Atmospheric Research Ctr (28.31 N)	Remote mountain	1 816	0.0	FTIR suite, Brewer, AERONET
209	Izaña Atmospheric Research Ctr (28.31 N)	Remote mountain	952	0.0	FTIR suite, Brewer, AERONET
182	Tel-Aviv University (32.11 N)	Urban coastal	1 255	13.5	AERONET
235	Korea University, Seoul (37.59 N)	Urban	328	0.0	AERONET
240	Aristotle Univ. Thessaloniki (40.63 N)	Urban coastal	873	0.0	UVVIS Brewer & DOAS, AERONET
115	CNR ISAC Rome (41.84 N)	Semi-rural	2 405	–	AERONET

117	Univ. Rome La Sapienza (41.90 N)	Urban	2 817	7.9	Aerosol lidars (532/1064 nm), SODAR, AERONET				
120	PMOD/WRC Davos (46.81 N)	Rural mountain	1 363	49.6	Brewer, AERONET				
133	Univ. Heidelberg (49.42 N)	Urban	252	0.0	AERONET				
162	Brussels-Uccle (50.80 N)	Semi-rural	1 416	7.3	Brewer, O_3/PTU sondes, DOAS, AERONET				
152	Ny-Ålesund, Svalbard (78.92 N)	Remote polar	1 574	29.8	FTIR suite, lidars, MWR, sondes, AERONET				

* *Available daily files* = number of Level-0 (raw) daily files already archived in the PGN.
** *QA %* = fraction of processed days that passed the PGN "quality-assured" filter.

7.3 C.3 Research about suitable RTM

The planned tropospheric AOD retrieval will be based on simplified geometrical considerations and will make use of pre-calculated look-up tables (see proposal for more details).

Also for the realization of other sub-WPs (e.g. Scattered Light Correction Development (WP1.4)) RTM calculations will be needed to model e.g. the forward scattering of different aerosol types and load which manifest in a distorting portion of radiance with the direct sun FOV.

Considering those use cases we have prescreened available RTMs and listed key aspects in the tabular view below.

Model	Functionality	Maintainability	Relevance
Eradiate *Rayference*	• 3D, Monte Carlo–based solver for fully arbitrary scenes • Spectral coverage: UV–VIS–NIR (280–2500 nm) • Satellite-image accuracy < 1	• Open-source (LGPL-3.0) on GitHub with active commits • Comprehensive docs & tutorials • Annual public workshops & tutorials	• Specifically built for calibration/validation of modern EO sensors • Selected for the Digital Twin Earth program • Designed to meet sub-percent accuracy requirements
libRadtran *LMU Munich*	• Primarily 1D plane-parallel (spherical shells limited) • Multiple solvers (DISORT, two-stream, MYSTIC Monte Carlo) • Polarization, Raman scattering, GUI	• 3D Monte Carlo not freely available • Maintained by Univ. of Munich team • libRadtran 2.0.1 overview paper (2016) + recent binary release ~6 months ago	• Widely used for 1D radiative studies in research and operational chains • Extensive user community and examples
(V)LIDORT *R. Spurr*	• 1D, pseudo-spherical discrete-ordinate (scalar LIDORT & vector VLIDORT) • On-the-fly linearization for sensitivity (retrievals) • Polarization support	• Not freely available • Developed starting 2006; vector version completed ~2010 • Last major public update (v2.8.3) in 2021	• Widely embedded in satellite retrieval algorithms (e.g. OMI/Aura) • Excellent for fast forward models in 1D

SCIATRAN *Rozanov et al.*	• Multilayer clouds, BRDF/albedo support • Scalar/vector UV–VIS–NIR–SWIR solver for limb, nadir, in-atmosphere modes [Rozanov et al., 2023]	• Freely available • Active development (v4.6 released 2023) with Python interface and updated aerosol databases [Rozanov et al., 2023]	• Adopted in Copernicus and major research campaigns • Key for trace-gas retrievals from limb-scanning sensors

While all listed options would in principle meet the requirements, we put special emphasis on integrability, maintenance actuality and support level.

7.4 C.4 Collect the Aeronet AODs

To automate the ingestion of column-integrated aerosol data for validation, we implemented a Python routine that queries NASA's AERONET Version 3 "print web data" interface for any site and user-defined period. The script issues a single HTTP request that already applies AERONET's built-in temporal averaging (e.g. 10-min means).

The returned ASCII table is parsed, invalid fill values (-999) are masked, and the spectrally resolved variables (AOD, triplet variability, and reported exact wavelengths) are reorganised into an xarray Dataset with a dedicated wavelength dimension alongside the time axis. All remaining scalar fields (e.g. solar geometry, cloud screening flags) are preserved as time-indexed variables. The result is a self-describing, analysis-ready dataset without further reformatting.

See below a Python application example

```python
# In[1] – set the query
from aeronet_helper import get_aeronet_aod   # assume the function lives here
site     = "Innsbruck_MUI"  # AERONET site name (any valid station label)
date0    = "2024-07-01"     # start of analysis period (YYYY-MM-DD)
date1    = "2024-07-03"     # end   of analysis period (YYYY-MM-DD)
level    = "AOD15"          # notquality-assured Level 1.5 data
avg_min  = 10               # built-in temporal averaging (minutes)
```

```python
# In[2] – download + structure into an xarray Dataset
ds = get_aeronet_aod(site, date0, date1, level=level, avg_interval=avg_min)
```

The output variable would look like shown below

```python
# In[3] – quick look at the result
print(ds)
<xarray.Dataset>
Dimensions:                 (time: 432, wavelength: 6)
Coordinates:
  * time                    (time) datetime64[ns] 2024-07-01T00:05:00 ... 2024-07-03T23:55:00
  * wavelength              (wavelength) int64 340 380 440 500 675 870
Data variables:
    AOD                     (wavelength, time) float64 0.108 0.102 ... 0.041
    Triplet_Variability     (wavelength, time) float64 0.003 0.003 ... 0.001
    Exact_Wavelengths_of_AOD (wavelength, time) float64 0.3401 0.3804 ... 0.8696
    Solar_Zenith_Angle(deg) (time) float64 68.2 67.4 ... 75.1
    Water_Vapor(cm)         (time) float64 1.46 1.48 ... 1.12
    ...
Attributes: (omitted)
```