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

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

Deliverable D7.3b

CINDI-3 semiblind inter-comparison Final report

Date: Friday 14th November, 2025
Version: 1.1

CINDI-3 semi-blind inter-comparison: Final report.

Martina M. Friedrich, Karin Kreher, Michel Van Roozendael,

FRM4DOAS-partners & CINDI-3 semi-blind inter-comparison participants:

André Achilli, Ramina Alwarda, Alkis Bais, Roberto Barragan, Darby Bates, Kirsten Blohm, Cedric Busschots, Elisa Castelli, Alexander Cede, Ka Lok Chan, Zeqing Chen, Liu Cheng, Emmanuel Dekemper, Erna Frins, Udo Frieß, Caroline Fayt, Robert Gilke, Pierre Gramme, Manuel Henning, Zhaokun Hu, Xiangguang Ji, Kevin Joshy, Dimitris Karagkiozidis, Kai Krause, Hyeong-Ahn Kwon, Kezia Lange, Hanlim Lee, Ang Li, Qidi Li, Chaonan Lv, Johannes Lampel, Yuhan Luo, Alexis Merlaud, Stefanie Morhenn, Monica Navarro, Dimitris Nikolis, Gyeong Park, Jong-Uk Park, Paolo Pettinari, Gaia Pinardi, Manuel Pinharanda, Ankie Piters, Olga Puentedura Rodriguez, Cristina Prados Roman, Denis Pöhler, Andreas Richter, Rob Ryan, André Seyler, Kimberly Strong, Martin Tiefengraber, Mark Wenig, Thomas Wagner, Pinhua Xie, Chengzhi Xing, Margarita Yela, Steffen Ziegler

Friday 14$^{\text{th}}$ November, 2025

Acronyms

σ standard deviation.

ABOM Australian Bureau of Meteorology.

ACTRIS Aerosol, Clouds and Trace Gases Research Infrastructure.

AIOFM Anhui Institute of Optics and Fine Mechanics.

AIRYX Airyx GmbH.

AMF air mass factor.

ATMO-ACCESS Sustainable Access to Atmospheric Research Facilities.

AUTH Aristotle University of Thessaloniki.

BIRA-IASB Belgian Institute for Space Aeronomy.

BK BK Scientific GmbH.

BVOC biogenic volatile organic compounds.

CESAR Cabauw Experimental Site for Atmospheric Research.

CINDI Cabauw Intercomparison of UV-Vis DOAS Instruments.

CNR-ISAC Consiglio Nazionale delle Ricerche - Istituto di Scienze dell'Atmosfera e del Clima.

CREGARS Centre for Reactive Trace Gases Remote Sensing.

D7.1 Deliverable 7.1: Campaign Planning Document.

D7.3.a Deliverable 7.3a: Final CINDI-3 campaign report.

DAILYREF the mean of the 10 daily references on each day.

DOAS Differential Optical Absorption Spectroscopy.

DOF Degrees of Freedom.

DSCD Differential Slant Column Density.

ESA European Space Agency.

EVDC ESA Validation Data Center.

FING Facultad de Ingeniería of Universidad de la República, Uruguay.

FIXREF the mean of the first 9 references measurements from June, 6th 2024.

FOV Field of View.

FRM4DOAS Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations.

FWHM Full Width at Half Maximum.

INTA Instituto Nacional de Técnica Aeroespacial.

IUPB Institut für Umweltphysik, Universität Bremen.

IUPHD Institut für Umweltphysik, Universität Heidelberg.

KNMI Koninklijk Nederlands Meteorologisch Instituut.

L1 level 1 (i.e. corrected for dark current, offset, as well as possibly straylight) spectra.

LATMOS Laboratory for Atmospheres, Observations, and Space.

LIDAR light detection and ranging.

LMU Ludwig Maximilian University of Munich.

los line of sight.

LT Local time.

MAX-DOAS Multi-Axis DOAS.

MPIC Max Planck Institute for Chemistry.

NASA National Aeronautics and Space Administration.

NDACC Network for the Detection of Atmospheric Composition Changes.

NIER The National Institute of Environmental Research of South Korea.

PI Principle investigator.

PKNU Pukyong National University.

RAL Rutherford Appleton Laboratory.

RMS Root mean square.

RSS Remote Sensing Site.

RT Room Temperature.

SAOZ Système d'Analyse par Observation Zénitale: System for Analysis by Zenithal Observation.

SEQREF the zenith measurement after each elevation scan.

SNR signal-to-noise ratio.

SUWON University of Suwon.

SZA Solar zenith angle.

UMelb University of Melbourne.

USTC University of Science and Technology of China.

UTC Coordinated Universal Time.

UToronto University of Toronto.

UV ultraviolet.

VEA Viewing Elevation Angle.

WS work shop.

Contents

MAY 2024

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
19	20	21	22	23	24	25
Whit Sunday	Whit Monday	Build-up/preparation	Build-up/preparation	Build-up/preparation	Build-up/preparation	Build-up/preparation
26	27	28	29	30	31	1
Build-up/preparation	Dry run phase	Dry run phase	Dry run phase	Dry run phase	Dry run phase	

JUNE 2024

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
26	27	28	29	30	31	1
						Dry run phase
2	3	4	5	6	7	8
Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase
9	10	11	12	13	14	15
Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase	Intensive phase
16	17	18	19	20	21	22
Intensive phase	Intensive phase	Intensive phase	Intensive phase	Optional extension	Optional extension	Site cleanup
23	24	25	26	27	28	29
Site cleanup	Site cleanup	Site cleanup	Site cleanup	Site cleanup	Site cleanup	

Figure 1: Final Campaign plan with the official inter-comparison period comprising 18 days highlighted in dark green.

1 Introduction

This document is the final report of the Differential Slant Column Density (DSCD) semi-blind intercomparison of Multi-Axis DOAS (MAX-DOAS)-type instrument part of the third Cabauw Intercomparison of UV-Vis DOAS Instruments (CINDI) which took place at the Cabauw Experimental Site for Atmospheric Research (CESAR) from 21 May until 24 June 2024, see Fig. 1. The final report about the other parts of the campaign (e.g. car-Differential Optical Absorption Spectroscopy (DOAS), airborne measurements, light detection and ranging (LIDAR) measurements etc.) are reported on in the accompanying document, Deliverable 7.3a: Final CINDI-3 campaign report (D7.3.a).

CESAR is one of the main facilities of the Dutch Ruisdael Observatory (https://ruisdael-observatory.nl/) located midway in between the cities of Rotterdam and Utrecht, and managed by Koninklijk Nederlands Meteorologisch Instituut (KNMI) at 51.968° N, 4.929° E. Over 100 researchers from 18 countries participated in CINDI-3, operating 33 UV-Vis DOAS instruments[1]. Table 1 contains a list of institutes/ companies participating in the semi-blind inter-comparison and the Principle investigators (PIs) and other involved persons. Table 2 gives a brief overview of the participating instruments, their spectral range, Field of View (FOV), the processing software used to fit DSCDs to the spectra, as well as the detector temperature. Here, Room Temperature (RT) denotes that the detector is neither cooled nor stabilized. Note that the color coding used in Table 2 will be used in the instrument labels to indicate the instrument type, where only blue (Skyspec), red (Pandora) and green (SAOZ) are truly describing the type of instrument, where orange indicates that the instruments had imaging capability (however, also custom build and hence not "a instrument type") and gray indicates an instrument that did not fit in any of the other categories, mostly meaning that it is a custom build instrument (with the exception of #4).

CINDI-3 is the third inter-comparison campaign of this kind, following CINDI in 2009 (Piters et al., 2012; Roscoe et al., 2010) and CINDI-2 in 2016 (Kreher et al., 2020). The overall campaign planning and coordination was carried out by a steering committee, led by BIRA-IASB, coordinating the CINDI-3 inter-comparison schedule and all additional activities such as the mobile and aircraft observations which are reported on in detail in D7.3.a.

CINDI-3 was organized as part of the Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS)-Centre for Reactive Trace Gases Remote Sensing (CREGARS) Topical Center for trace gas remote sensing, within the framework of the collaboration between ACTRIS and Network for the Detection of Atmospheric Composition Changes (NDACC), and conducted with additional support from European Space Agency (ESA), National Aeronautics and Space Administration (NASA) and from the EU Horizon2020 Sustainable Access to Atmospheric Research Facilities (ATMO-ACCESS) project (https://www.atmo-access.eu/). Additionally, the campaign was supported by KNMI.

[1] all but 3 had MAX-DOAS capability. The 3 Système d'Analyse par Observation Zénitale: System for Analysis by Zenithal Observation (SAOZ) instruments were zenith sky looking only

Table 1: Institutes and companies and the respective PIs and other involved persons that either participated with at least 1 instrument to the semi-blind inter-comparison during CINDI-3 or contributed directly to the data processing

Australian Bureau of Meteorology (ABOM) and University of Melbourne (UMelb)	Rob Ryan
Anhui Institute of Optics and Fine Mechanics (AIOFM)	Zhaokun Hu, Ang Li, Yuhan Luo, Chaonan Lv, Pinhua Xie, Qidi Li
Airyx GmbH (AIRYX)	Johannes Lampel, Denis Pöhler
Aristotle University of Thessaloniki (AUTH)	Alkis Bais, Dimitris Karagkiozidis, Dimitris Nikolis
Belgian Institute for Space Aeronomy (BIRA-IASB)	Michel Van Roozendael, Cedric Busschots, Emmanuel Dekemper, Caroline Fayt, Martina M. Friedrich, Pierre Gramme, Alexis Merlaud, Gaia Pinardi.
BK Scientific GmbH (BK)	Karin Kreher
Consiglio Nazionale delle Ricerche - Istituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC)	André Achilli, Elisa Castelli, Paolo Pettinari.
Facultad de Ingeniería of Universidad de la República, Uruguay (FING)	Roberto Barragan[2], Erna Frins.
Instituto Nacional de Técnica Aeroespacial (INTA)	Monica Navarro, Olga Puentedura Rodriguez, Cristina Prados Roman, Margarita Yela
Institut für Umweltphysik, Universität Heidelberg (IUPHD)	Udo Frieß, Kirsten Blohm
Institut für Umweltphysik, Universität Bremen (IUPB)	Kai Krause, Kezia Lange, Andreas Richter, André Seyler
KNMI	Ankie Piters
Laboratory for Atmospheres, Observations, and Space (LATMOS)	Manuel Pinharanda
Ludwig Maximilian University of Munich (LMU)	Zeqing Chen, Manuel Henning, Mark Wenig
LuftBlick	Stefanie Morhenn, Martin Tiefengraber, Alexander Cede
Max Planck Institute for Chemistry (MPIC)	Robert Gilke, Thomas Wagner, Steffen Ziegler.
Pukyong National University (PKNU)-The National Institute of Environmental Research of South Korea (NIER)	Hanlim Lee, Gyeong Park
Rutherford Appleton Laboratory (RAL)	Ka Lok Chan
University of Suwon (SUWON)	Hyeong-Ahn Kwon, Jong-Uk Park
University of Science and Technology of China (USTC)	Liu Cheng, Xiangguang Ji, Chengzhi Xing
University of Toronto (UToronto)	Ramina Alwarda, Darby Bates, Kevin Joshy, Kimberly Strong

Fig. 2 shows the final campaign schedule. The results presented in this report are those taken during the period labelled "intensive phase" that comprises 18 fays from June, 2nd to June 19th.

There were 2 mayor deadlines for the data submission: one roughly 3 months after the campaign, on Sept. 30th 2024. While initially planned as the final deadline, a second deadline was set roughly 8 month later, May, 16th 2025. The reason for this generous time line when compared to the previous campaign is two fold:

- Many groups fairly new to the DOAS technique participated in this campaign, hence particular weight was put on giving the groups the possibility to improve their data processing after feedback was received.
- The weather conditions were very unfavourable in terms of temperature and cloud coverage and hence posed challenges not encountered during the previous edition of the CINDI campaign.

Regarding data availability, many of the CINDI-3 campaign data sets can be found and accessed on the CINDI-3 Research Drive in the relevant folders. The data used for the analysis in this report will be eventually hosted by ESA Validation Data Center (EVDC) and made available through links on this webpage: `https://frm4doas.aeronomie.be/index.php/cindi-3` after a first draft of a publication describing these results is submitted.

The reminder of this report is organized as follows: Sect. 2 gives an overview of the products, the measurement protocol and the data availability. Sect. 3 describes the analysis and support given during the campaign as well as after the campaign in terms of a web-tool (for elevation scan measurements) and accompanying summary PDF-documents (for horizon scans and high Solar zenith angle (SZA) zenith measurements). Sect. 4 presents the results from the daily horizon scans performed. Sect. 5 presents the post-campaign results from the inter-comparison of zenith only measurements during medium and high SZA while Sect. 6 presents the post-campaign results from the elevation scan analysis during medium and low SZA. Beside the individual processing, a number of instruments also delivered the level 1 (i.e. corrected for dark current, offset, as well as possibly straylight) spectra (L1) files to the Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations (FRM4DOAS) centralized processor. Results of these and comparisons with own processing are presented in Sect. 7. Finally, Sect. 8 comprises a summary of the main findings, conclusions, and recommendations for future campaigns.

Table 2: participating instruments

#	Institute	Instrument type	Spectral range [resolution]/ nm	FOV/°	fit software	detector T/°C
01	ABOM	Airyx SkySpec-1D	296 – 460 [0.6] 441 – 584 [0.6]	0.3[1]	QDOAS v3.6.5	20
02	UMelb	Airyx SkySpec-1D	296 – 410 [0.5] 408 – 555 [0.6]	0.3[1]	QDOAS v3.6.5	20
03	KNMI	Airyx SkySpec-1D	333 – 491 [0.65]	0.5[1]	KMDOAS v2.2	20
06	UToronto	Airyx SkySpec-1D	300 – 540 [0.6]	0.3[1]	QDOAS v3.6.5	20
07	SUWON	Airyx SkySpec-1D	300 – 460 [< 0.7]	0.3[1]	QDOAS v3.7.2	20
14	LMU	Airyx SkySpec-1D	300 – 460 [0.6]	0.3[1]	QDOAS v3.6.5	20
21	BIRA-IASB	Airyx SkySpec-2D	297 – 462 [0.55] 408 – 553 [0.58]	0.3[1]	QDOAS v3.7.4	10
24	CNR-ISAC	Airyx SkySpec-2D	300 – 405 [0.45] 405 – 565 [0.6]	0.3[1]	QDOAS v3.4.6	20
28	IUPHD	Airyx SkySpec-2D	300 – 550 [0.6]	0.3[1]	heiDOAS v1.2	20
32	AIRYX/FING	Airyx SkySpec-2D	300 – 550 [0.3]	0.3[1]	QDOAS v3.5.0	20
33	RAL	Airyx SkySpec-2D	310 – 410 [0.45] 405 – 550 [0.6]	0.3[1]	QDOAS 3.7.2	20
39	USTC	Airyx SkySpec-2D	297 – 408.5 [0.43] 419.1 – 565.6 [0.6]	0.3[1]	QDOAS v3.2.0	20
04	KNMI	Mini-DOAS Hoffmann	297 – 515 [0.74]	0.9	KMDOAS v2.2	5
05	MPIC	Tube-DOAS	300 – 460 [0.65]	0.6	QDOAS v3.5.0	20
15	AIOFM	MAX-DOAS-2D	290 – 450 [0.4]	0.5	QDOAS v3.6.5	25
16	AIOFM	MAX-DOAS-2D	220 – 430 [0.35]	0.5	QDOAS v3.2	-40
17	AIOFM	MAX-DOAS-1D	290 – 420 [0.55]	0.5	QDOAS v3.2	25
18	AUTH	PHAETON	280 – 539 [0.5–0.6]	1	QDOAS v3.6.0	10
19	AUTH	DELTA	305 – 523 [0.85]	0.8	QDOAS v3.6.0	-50
20	BIRA-IASB	MAX-DOAS-2D	304 – 383 [0.35] 402 – 543 [0.5]	< 1	QDOAS v3.6.10	-50
25	IUPB	MAX-DOAS-2D	315 – 376 [0.21] 411 – 545 [0.5]	1	nlin v7.135	-45
29	IUPHD	PMAX-DOAS	285 – 565 [1.3]	0.26	heiDOAS v1.2	-28
30	INTA	RASAS-III	422 – 539 [0.6]	0.5	LANA v9.2.6	-30
40	PKNU-NIER	AQ-Profiler	280 – 795 [0.5]	2.2	QDOAS v3.2	20
22	BIRA-IASB	SEMPAS	285 – 505 [1]	0.15[2]	QDOAS v3.6.10	-50
23	BIRA-IASB	NO2 camera	430 – 490 [0.6 – 0.8]	20/ 0.04	NO2CAM-spec. v2.3.1	≥(RT-20)
27	IUPB	IMPACT	402 – 505 [~1]	1.5 [3]	nlin v7.135	-30
08	LATMOS	SAOZ	270 – 640 [1.37]	10	SAOZ V1.47- 1d37765	RT
09	LATMOS	mini-SAOZ	270 – 820 [0.71]	8	SAOZ V1.47- 1d37765	RT
10	LATMOS	mini-SAOZ	270 – 820 [0.75]	8	SAOZ V1.47- 1d37765	RT
34	KNMI	Pandora-1S	280 – 530 [0.6]	1.5	BlickP v1.8.62	25 – 26
35	Luftblick	Pandora-1S	280 – 530 [0.6]	1.5	BlickP v1.8.62	25 – 26
36	Luftblick	Pandora-1S	280 – 530 [0.6]	1.5	BlickP v1.8.62	25 – 26

[1] All SkySpec instruments have an asymmetrical field of view, roughly 0.3° vertically and 0.9° horizontally
[2] Sempas has an asymmetrical field of view, the resolution in the horizontal dimension is 7.65°.
[3] IMPACT has 1.5° for low elevation angles, 7.5° for 15° elevation, 12° for 30° elevation and 35° for zenith.

2 Semi-blind intercomparison description and Data overview

This section gives a general overview of the measurement protocol and data collected during the semi-blind inter-comparison by MAX-DOAS-type instruments CINDI-3 (Subsect. 2.1) and outlines the recommended filtering approach for the post-campaign data delivery in Subsect. 2.2.

2.1 Measurement protocol and data overview

Experience from previous inter-comparison campaigns, particularly CINDI and CINDI-2, has shown that the level of agreement between MAX-DOAS observations can be significantly affected by imperfect co-location and a lack of synchronization in the viewing directions of the various instruments. This issue is especially critical when comparing tropospheric trace gases such as NO_2, which can exhibit strong spatial gradients due to localized pollution sources.

To address this, all spectrometers participating in the semi-blind inter-comparison exercise during CINDI-3 were installed in close proximity on the Remote Sensing Site (RSS) platform at the CESAR station in Cabauw. In addition, a strict protocol for the timing of spectral acquisitions was implemented to ensure synchronization across instruments.

The measurement acquisition protocol is described in detail in Deliverable 7.1: Campaign Planning Document (D7.1), which includes a comprehensive table outlining the required acquisition sequences. The most important elements of the protocol are summarized below.

During the 18 days of the semi-blind inter-comparison campaign, measurements were scheduled to start and end each day at twilight with 1-minute zenith observations. Between 04:10 Coordinated Universal Time (UTC) and 18:10 UTC, instruments followed an hourly measurement schedule, which varied depending on the instrument type (1D or 2D).

These hourly measurement schedules are illustrated in Fig. 2. The left panel shows the schedule for 2D-MAX-DOAS instruments, the middle panel corresponds to 1D-MAX-DOAS instruments, and the right panel highlights deviations from the standard hourly schedule during the noon hour. This midday adjustment allows time for the acquisition of daily reference measurements (10 measurements) as well as upward and downward horizon scans.

Outside the defined hourly schedule, before the first (pre-04:10) and after the last (post-19:09) scheduled measurements, all instruments performed continuous zenith observations with 1 min integration time.

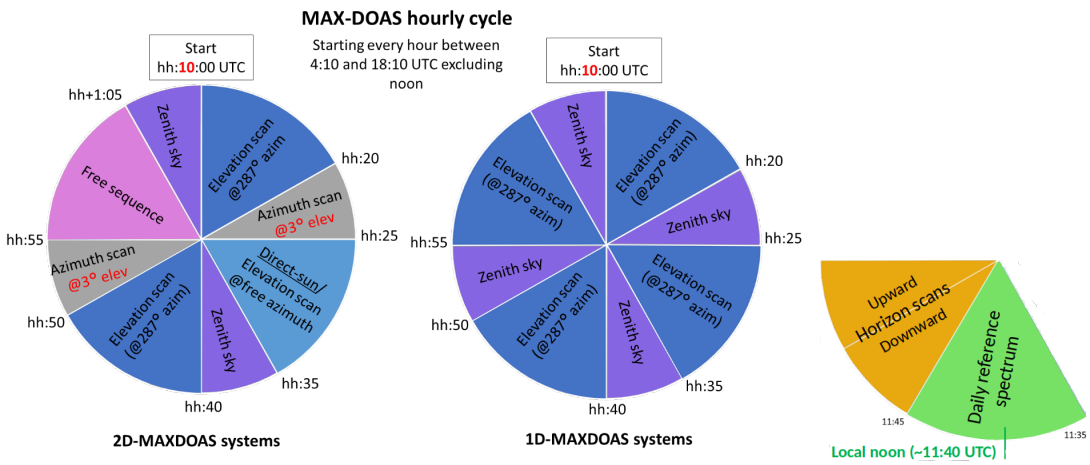

Figure 2: MAX-DOAS measurement protocols. Left panel: hourly measurement protocol for 2D MAX-DOAS systems; Middle panel: hourly measurment schedule for 1D MAX-DOAS systems; Right panel: Changes in the measurement protocol during the local noon hour from 11:35 to 11:55, where the reference time slot and the horizon scans replace parts of the otherwise scheduled measurements.

It was decided to include 10 products in the semi-blind inter-comparison. The products together with their abbreviation as well as the fitting range is summarized in Table 3. For details on the fitting parameters, we refer to D7.1.

Each of the products was analyzed with three different references:

- the mean of the 10 daily references on each day (DAILYREF): most groups used the mean of all 10 zenith measurements, however some groups applied a filtering on those measurements and only used the mean of a subset. In cases of rapid concentration changes, this might lead can lead to slightly different columns in the reference and hence to additional differences in the product.

Data product	Abbreviation	Fitting range/ nm
NO_2 (VIS range)	NO2VIS	425.0–490.0
NO_2 (alternative VIS range)	NO2VIS-SMALL	411.0–445.0
NO_2 (UV range)	NO2UV	338.0–370.0
O_4 (VIS range)	O4VIS	425.0–490.0
O_4 (UV range)	O4UV	338.0–370.0
HCHO	HCHO	336.5–359.0
HCHO (extended range)	HCHO-WIDE	324.5–359.0
HONO	HONO	335.0–373.0
O_3 (Chappuis bands)	O3VIS	450.0–540.0
O_3 (Huggins bands)	O3UV	320.0–340.0

Table 3: Overview of data products, abbreviations, and spectral fitting ranges.

- the zenith measurement after each elevation scan (SEQREF)
- the mean of the first 9 references measurements from June, 6th 2024 (FIXREF)

In the following, we refer to product-reference type combinations such as NO2UV_FIXREF. This example refers hence to the DSCD fits for NO_2 in the UV-range fitting window with a fixed reference from June, 6th. This report will hence cover 30 different product-reference type combinations.

Not all instruments listed in Table 2 covered the spectral ranges of all products. Therefore, not all products are submitted for all instruments. Additionally, not all instrument data providers provided the measurable products analyzed with all possible references. An overview of which product-reference type combination is delivered by which instrument, is shown in Fig. 3.

Note that instruments #8, #9 and #10 are zenith looking instruments only, hence naturally, they only deliver 90° Viewing Elevation Angle (VEA) measurements for which no SEQREF measurements are possible. Further, instrument #22 joined the campaign after June, 6th and hence has no measurements on the day selected for FIXREF.

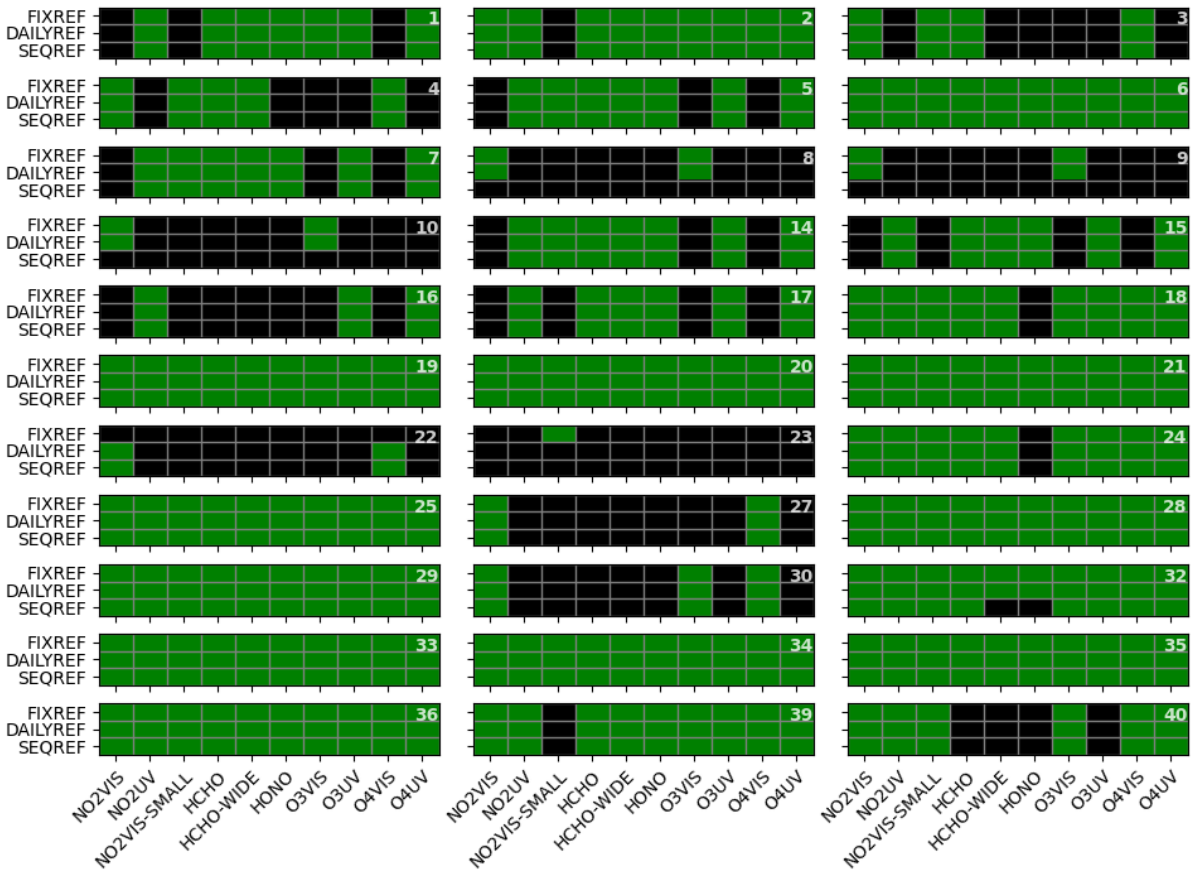

Figure 3: Overview of submissions of product (x-axis)-reference type (y-axis) measurements for the different instruments as indicated in the upper right panel of each image. Green indicates that at least 1 day of data is submitted, black indicates that no data is available for that instrument of that product-reference type combination.

Some instruments arrived slightly later while others had technical issues and have therefore holes

in the dataset. Other instruments were removed from the site before the official end of the final period. The latter is due to the fact that the campaign plan changed slightly during the campaign since the weather conditions were not favourable. An overview of the total number of instruments submitting data for each product-reference type combination on each day is presented in Fig. 4. While the official intensive phase was extended to June, 19th 2024, for several products the baseline comparison is not calculated for that day since we require a certain percentage, see Sect. 5 and Sect. 6, of valid measurements for each time in order to calculate the baseline.

Figure 4: Overview of the number of instruments delivering a certain product-reference type combination each day.

2.2 Recommendation on filtering

During the CINDI-3 work shop (WS) in Heidelberg (17th March – 19th March 2025) two important decisions regarding the data products were taken:

1. The reference measurement used for the DSCD fits for the sequential reference (SEQREF) analysis should always be the zenith measurement at the end of each scan.

2. It was suggested that groups perform a filtering based on fitting Root mean square (RMS) to avoid the submision of clearly corrupted data.

The suggested approach for 2. was first outlined as follows and distributed on 27 March 2025: "During the meeting, it was decided that, if filtering of any data is needed, each participating group should apply this filtering to their own data. Any filtering should be mainly based on RMS. Since the RMS for a given time for a given product, spans often more than 1 order of magnitude between the different instruments and since the instrument owner/operators know their own observations best, it was decided that the PIs are responsible for the filtering of their own data. Since the RMS increases with decreasing light, RMS is usually a function of SZA. Hence, a fixed limit for the RMS filter is not a good idea. PIs should plot the RMS of their products (as a function of SZA) and set appropriate limits. If identified problems (e.g. condensation) were present during a certain time period, this should be clearly stated."

A refinement (originally suggested by D. Karagkiozidis) of the instructions was circulated on 28 April 2025 as follows: "Group the data separately for each product-reference type-elevation angle combination into SZA bins of roughly 10° during the elevation scan measurements time slot. The SZA during the ELEV time slot range roughly from 28° to 85°, hence 6 bins of 9.5° each could be a good choice. For the twilight zenith measurement timeslots (roughly from 84.5° to 94.5° *) you could use 10 bins of 1°. Calculate the 95th percentile of the RMS distribution in each of these SZA angle bins (per product, per reference type and per elevation angle) and use this as the limit on the data in each of those bins."

3 On-site data analysis and post-campaign support

This section comprises two parts, where Subsect. 3.1 describes the daily routine of measurement taking, data submission, data analysis and analysis presentation and Subsect. 3.2 describes an online tool developed for the data submitters after the campaign to identify remaining issues.

3.1 On-site analysis and daily routine

During the official phase of the CINDI-3 inter-comparison (2 June – 19 June), participating groups were required to adhere to the conditions of a semi-blind inter-comparison under the auspices of NDACC. This required that all measurements from a given day be submitted by 11:00 Local time (LT) the following morning. Preliminary results, based on initial data analysis and investigations, were then discussed each afternoon during on-site meetings.

To preserve anonymity, the submitted measurements were displayed without identifying the corresponding instruments or research groups. However, in cases where clear anomalies, errors, or biases were observed in individual datasets, the referee team contacted the relevant group to allow further investigation and, if necessary, correction for the remainder of the campaign.

Following submission, each dataset was checked for format compliance. The format is described in D7.1. If formatting errors were detected, the corresponding group was notified within 20 minutes and given an additional hour to resubmit the data in the correct format. This process ensured that results from the previous day's measurements could be included in the daily meetings, held at 15:00 LT. All meeting presentations are archived and available on the KNMI research drive `https://knmi.data.surf.nl/apps/files/files/6385419?dir=/CINDI-3/Meetings`.

Typical analysis plots shown in these meetings comprised

- comparison plots for the diurnal variation of the slant columns for each submitted data product for the main azimuth direction and all viewing elevation angles. These plots came in two flavours:
  - Direct DSCD time series of 1–5 days, e.g. Fig. 5 for NO2VIS_DAILYREF from 11 June. Each panel shows a time series for a separate elevation angle as indicated in the lower right corner in each panel. The solid black line is the median of all instrument measurements. The red lines indicate the number of instruments at each given time. The regular dips correspond to those elevation scans that were only performed by 1-D instruments, compare Fig. 2.
  - Relative (and absolute) difference plots of individual DSCD compared to the median of the data from all instruments. Fig. 6 shows such a plot for the data shown in Fig. 5. Note that the relative difference at 30° arises due to small absolute DSCDs.
- horizon scan plots. Horizon scans were used during the CINDI-3 campaign to assess the ability of the instruments to measure at the correct viewing elevation and, in particular, to investigate any deviation in viewing elevation angle with respect to the expected horizon elevation. A set of horizon scan measurements was performed every day from 11:45 to 11:55 UTC in upward and downward directions (c.f.Fig. 2). An example for horizon scans covering the warm-up phase and the first days of the intensive phase, shown during the first weekly meeting on 4 June is shown in Fig. 8.
- zenith only DSCD comparisons at high SZA for selected products. An example for 15 June 2024, as presented during the daily meeting on 16 June 2024 is shown in Fig. 7.

In summary, the primary objective of the data comparison during the active campaign phase was to monitor instrument performance and to identify potential issues in data processing. Particular emphasis was placed on detecting instrumental problems that could be addressed onsite. These included incorrect elevation viewing angles, obstructions in the line of sight of individual instruments, measurement disturbances caused by participant movement on the instrument platforms and scaffolding, and potential interferences from laser emission lines (from the various LIDAR systems operating at the campaign site) under specific cloud conditions.

As previously mentioned, the referee team provided support to individual participants by highlighting possible issues observed in the submitted data products. In addition to identifying instrumental problems, the team also noted and reported issues arising from limited experience with data analysis procedures.

While BrO and glyoxal measurements were included in the on-site campaign activities, they were not further analysed in the post-campaign main comparison but instead are treated in dedicated working groups. This allowed for a more detailed investigation of the corresponding data processing approaches and enabled open discussion among participants, something not feasible within the constraints of a semi-blind inter-comparison.

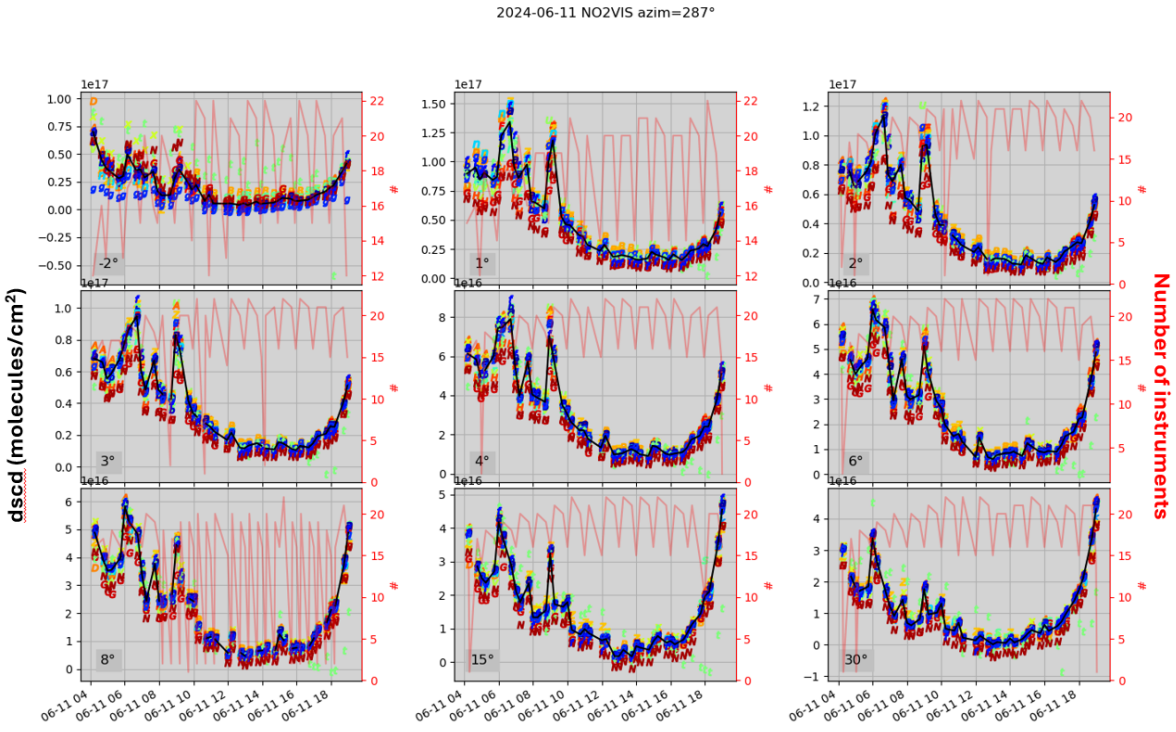

Figure 5: NO2VIS_DAILYREF results from 11 June 2024 as presented during the daily meeting on 12 June 2024. Each instrument is displayed with a certain color and letter that were changed multiple times during the campaign. The black line shows the median of the data from the different instruments while the red line indicates the number of instruments as indicated in the r.h.s axis.

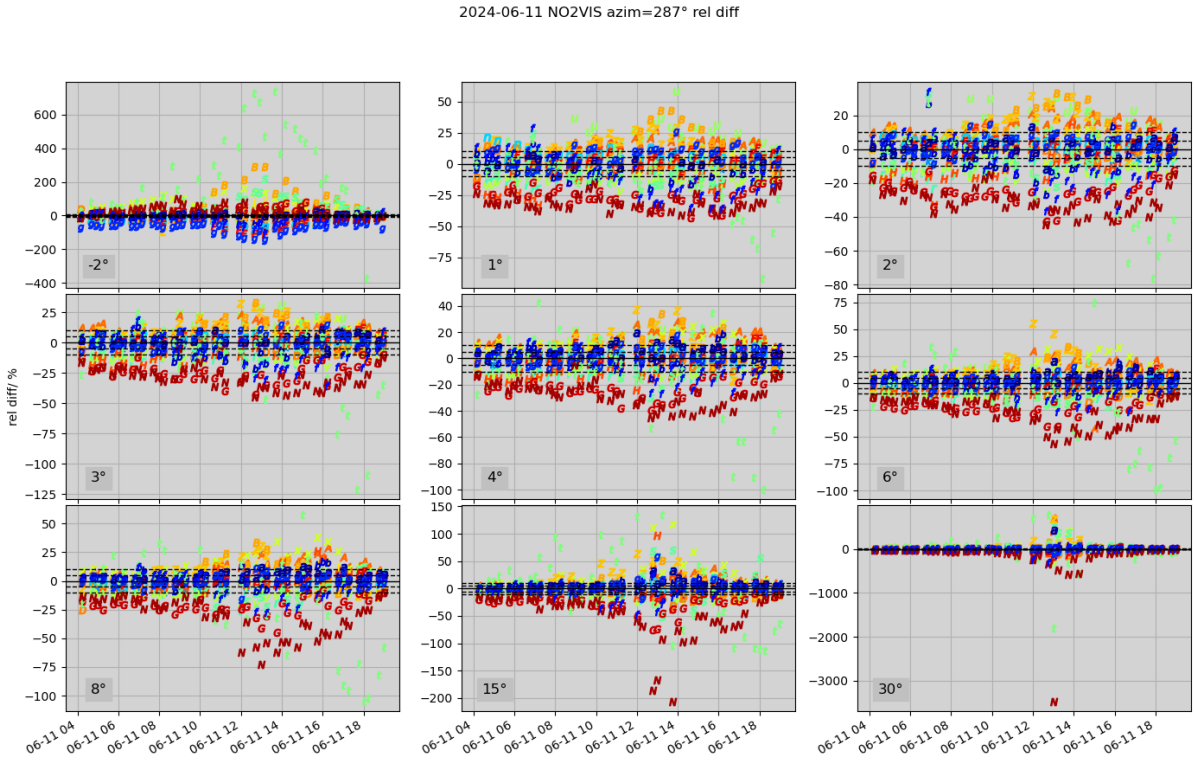

Figure 6: NO2VIS_DAILYREF DSCD differences between individual instruments and the median of the instruments from 11 June 2024 as presented during the daily meeting on 12 June 2024. Each panel shows a different elevation angle as indicated in the lower left corner. 5% and 10% deviations are indicated with dashed lines.

3.2 Post-campaign support

Post campaign support comprised mainly 3 types of feedback:

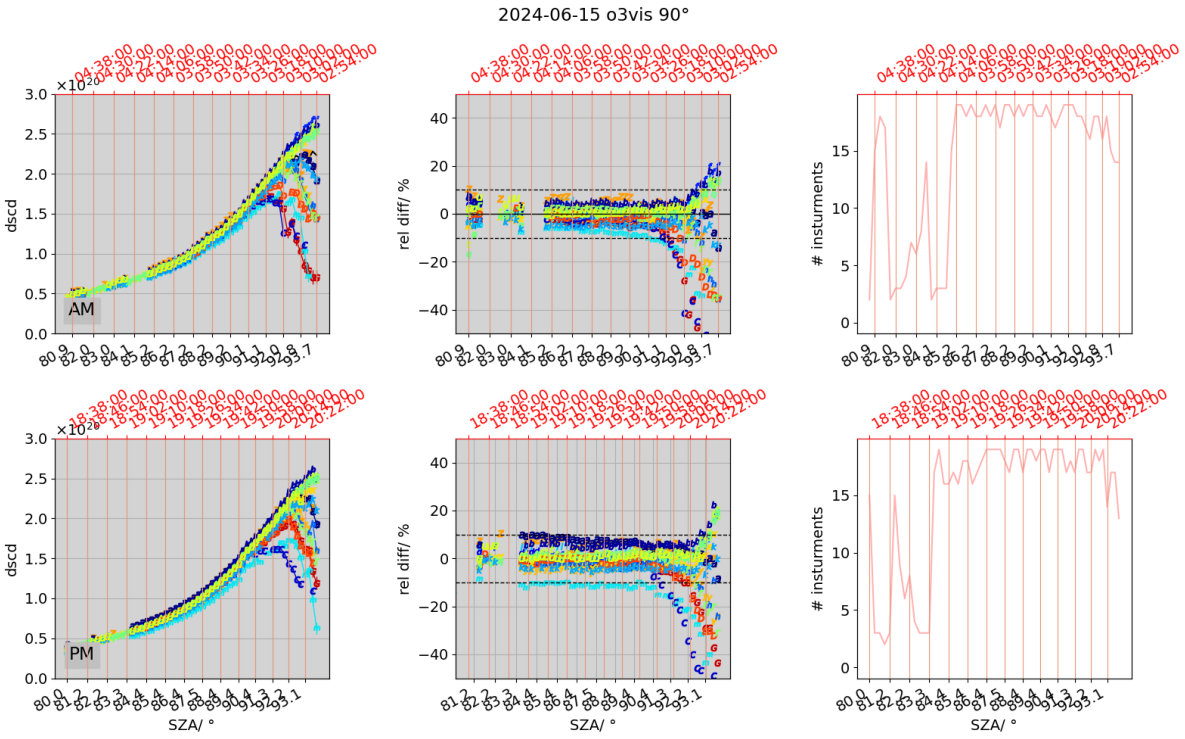

Figure 7: O3VIS_DAILYREF DSCD at AM (upper panels) and PM twilight from 15 June 2024. Right panels show the DSCD of the individual instruments, middle panels show the deviation from the median and right panels show the number of instruments at each time.

- giving continuous information about which is the current data version that is used for each instrument for each day for each product-reference type (in form of a plain text file). Quick checks on whether new data-versions introduced a worseing [3]
- interactive feedback on the DSCDs for selected products for the elevation scans, see Subsect. 3.2.1
- instrument specific feedback on the DSCDs for selected products for zenith measurements only, see Subsect. 3.2.2

3.2.1 Online tool for elevation scan assesssment

A web-based tool for data analysis and visualization of the CINDI-3 campaign was developed by M. M. Friedrich (BIRA) and was made available online (28 Oct. 2024) to all participants for reviewing their data prior to final submission. The tool remains currently accessible at: https://cindi3.pythonanywhere.com/ (user: cindi3, password: oh_cindi3_was_so_rainy).

The main purpose of the provision of the online tool was to give participants the possibility to detect where remaining processing issues might be present, since the aim of the inter-comparison is not to show the differences in ability to [TODO what do I actually want to write here] but to assess the quality of the instruments.

Due to time limitations imposed by the free server hosting, long-term access to the tool is not guaranteed. However, the final datasets and tools for interactive access will be made available without password protection on the CINDI-3 website (https://frm4doas.aeronomie.be/index.php/cindi-3). This release will follow the submission of the first draft of the inter-comparison report.

In the online tool, only measurements acquired as part of elevation scan sequences, and only those recorded by both 1D and 2D instruments, have been included. As a result, there are 29 measurements per day for each elevation angle. Considering the data from 2 to 19 June, the theoretical maximum number of measurements per elevation angle is 522.

The tool includes eight data products: NO2VIS, NO2VIS-SMALL, NO2UV, O4VIS, O4UV, HCHO, HCHO-WIDE, and HONO (ozone data are not included) for the three reference types DAILYREF, FIXREF, and SEQREF. The main user interface is shown in Fig. 9.

Three types of plots can be visualized:

[3] Specifically, a number of groups introduced issues after changing the sequential reference from their initial choice to the zenith measurement after.

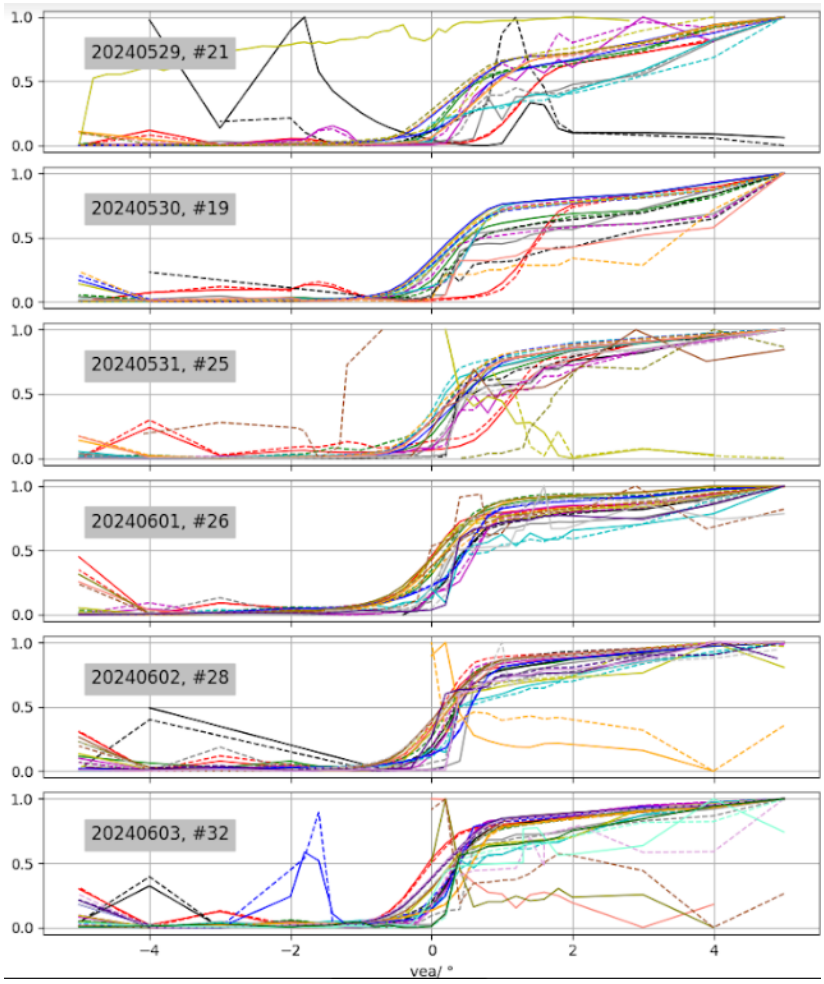

Figure 8: Horizon scan plot as shown in the weekly meeting on 4 June. Each panel shows all instruments horizon plots for a different day as indicated in each panel. The number following the date indicates how many horizon scans are shown. Since this shows both upward (solid lines) as well as downward (dashed lines) scans, the number is approximately double number of instruments submitting horizon scans.

- Direct time series (see Fig. 10)
- Scatter plots comparing individual instruments against the median (see Fig. 11), with interactive functionality to display the time and date of each data point by hovering
- Meta-plots, showing data quality indicators such as the slope from a Theil–Sen analysis or the RMS of differences relative to the median for individual elevation angles (see Fig. 12).

3.2.2 Collection of PDF files for zenith only measurements

To complement the feedback on the elevation scans via the online tool, pdf documents were provided for each instrument. For up to 8 product-reference type combinations (O3VIS, NO2VIS, NO2VIS-SMALL, O3UV with reference types DAILYREF and FIXREF), documents were distributed, each comprising 38 pages showing for each day once the DSCD over the whole SZA range and once showing a zoom. Each document started with a general description of the data displayed. An example of the first two pages is shown in Fig. 13. These documents were distributed on 30 March 2025.

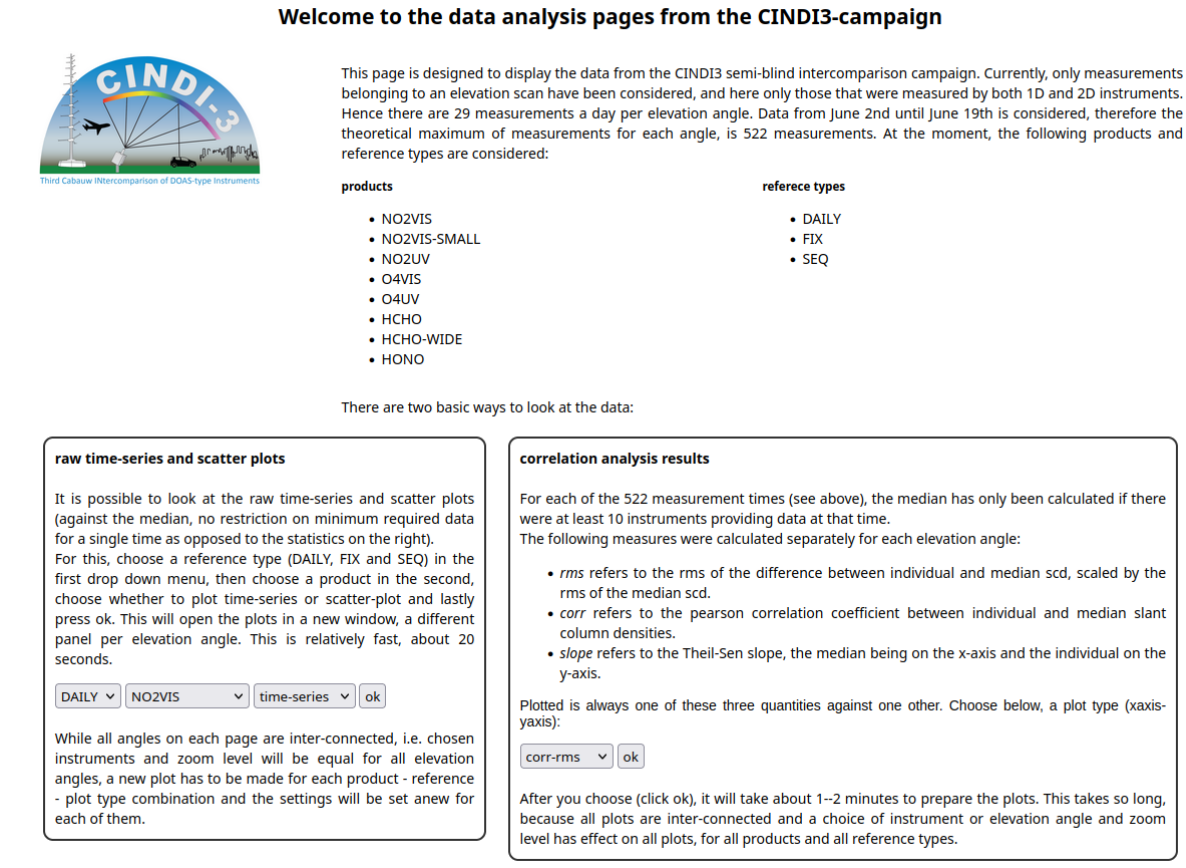

Figure 9: Starting interface of the web analysis tool.

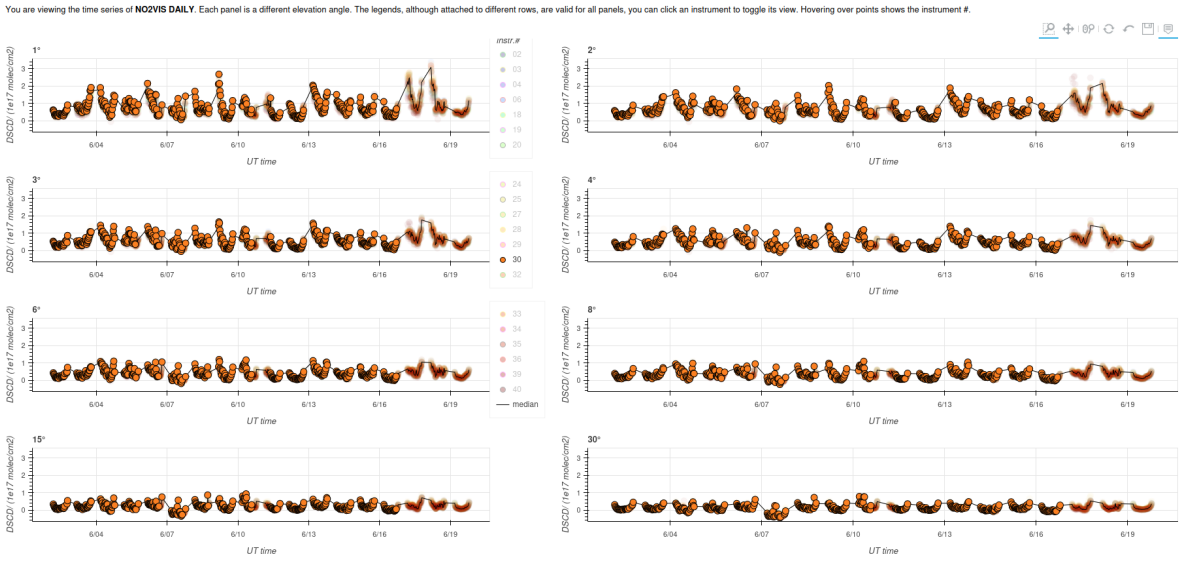

Figure 10: Example of time series for selected instruments for NO2VIS_DAILY.

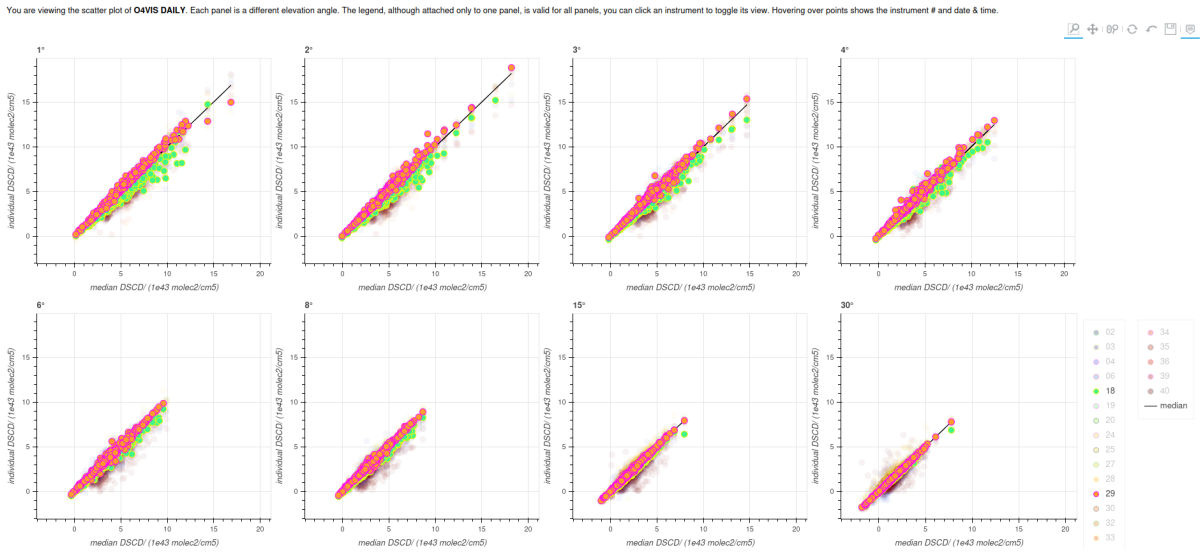

Figure 11: Example of scatter plot for selected instruments for O4VIS_DAILY. Note that hovering over the individual data points showed the exact time and date of the measurement in question.

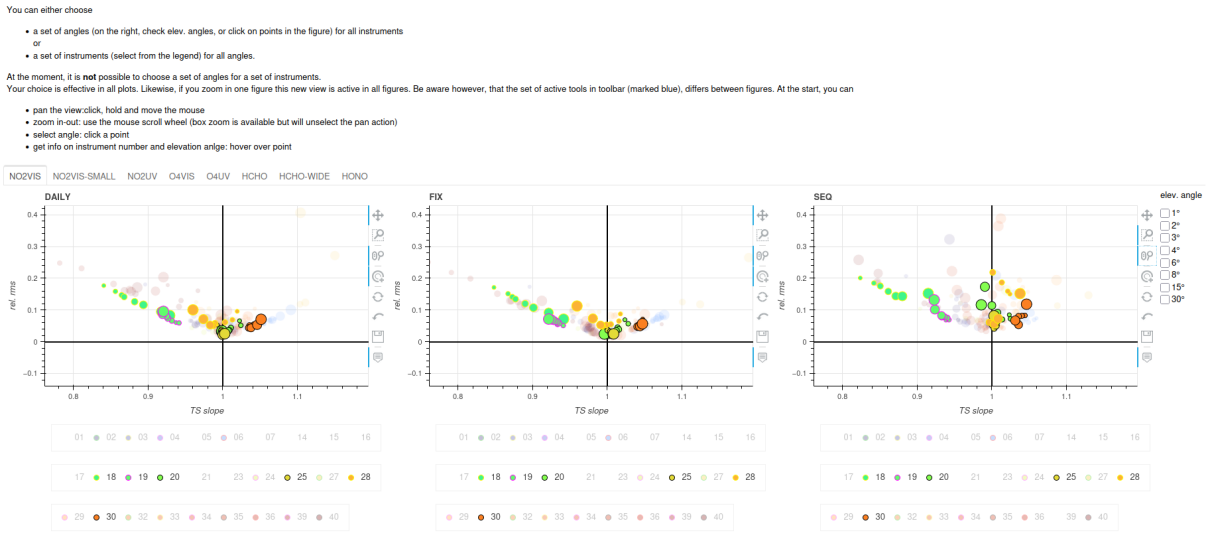

Figure 12: Example of meta analysis plots for selected instruments for NO2VIS. Note that after instrument and angle selection, the data product can be switched directly for easy inter-product comparisons.

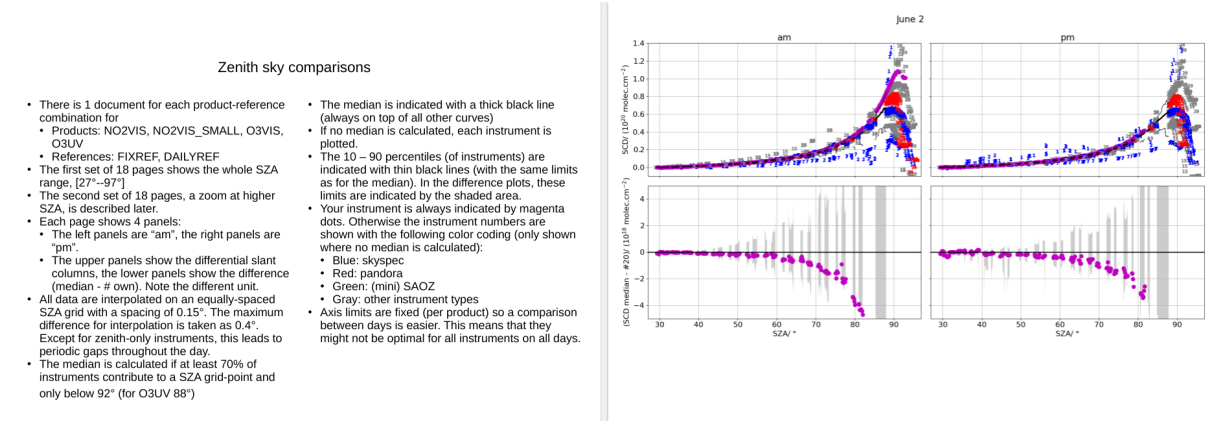

Figure 13: Example of first 2 pages for feedback provided for zenith only measurements for O3VIS_FIXREF to instrument #20.

4 Horizon scans

4.1 General considerations

Horizon scans, measurements of sky radiance intensity across the vertical horizon, were systematically performed each day during the noon hourly cycle (see Fig. 2) during the semi-blind inter-comparison period with steps of 1° between -5° and -2° and between 2° and 5° and with steps of 0.2° between -2° and 2°. Although absolute calibration is limited by common obstructions near the horizon, these scans provide a straightforward and effective method for assessing the elevation pointing stability of MAX-DOAS instruments. The data were fitted using a combination of an error function and a first-degree polynomial (see Eq. 1), capturing the integral shape of the instrument's effective FOV and accounting for linear background trends.

$$I = \frac{a}{2} \cdot \mathrm{erf}\left(\frac{x - x_0}{w}\right) + s(x - x_0) + y_{\mathrm{off}} + \frac{a}{2} \tag{1}$$

This fitting approach was previously applied during the CINDI-1 campaign (Kreher et al., 2020). The width parameter w of the error function relates to the standard deviation σ of the corresponding Gaussian (whose integral is the error function) as:

$$w = \sqrt{2}\sigma \tag{2}$$

Consequently, the Full Width at Half Maximum (FWHM) of the instrument's effective FOV is given by:

$$\mathrm{FWHM} = 2\sqrt{2\ln 2}\sigma = 2\sqrt{\ln 2}w \approx 1.665w \tag{3}$$

There is certain ambiguity between w and the combination of slope s (necessary to account for background trends), amplitude a and offset y_{off}. This is demonstrated in Fig. 14.

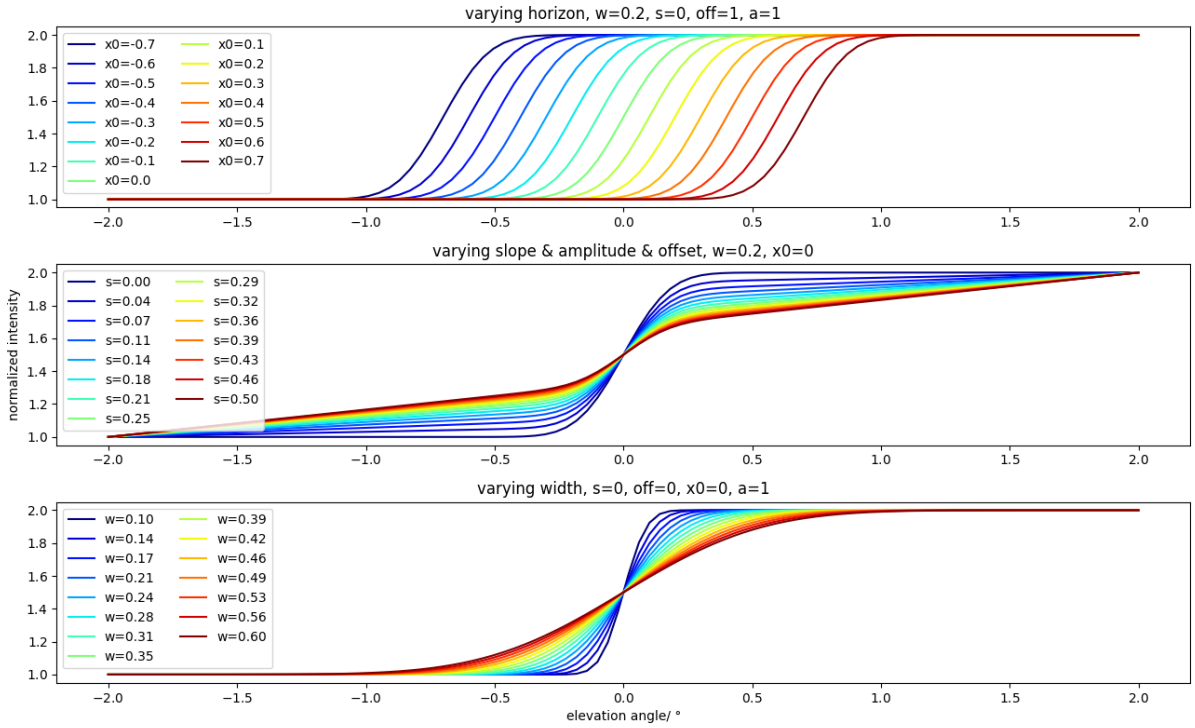

Figure 14: Graphs of Eq. 1, keeping different fitting parameters constant and varying others, as indicated in the panel headers and panel legends. Note the similarity between curves in the second and third panel, indicating some ambiguity between the different fit parameters.

Using unfiltered horizon scans is therefor not a practical approach to determine the FOV. To demonstrate this ambiguity for imperfect measurements, we show an example in Fig. 15. Here, we show downward scans on two consecutive days for instrument #24 on June 5th and June 6th. The actual measurements are indicated with red circles (June 5th) and magenta squares (June 6th), while the corresponding fits are indicated with a solid black (June 5th) and blue dashed (June 6th) line. It is clearly visible that, although the actual measurements are very similar, the fit for the width and slope are very different. This is mostly due to the outlier at 0.2° on June 6th.

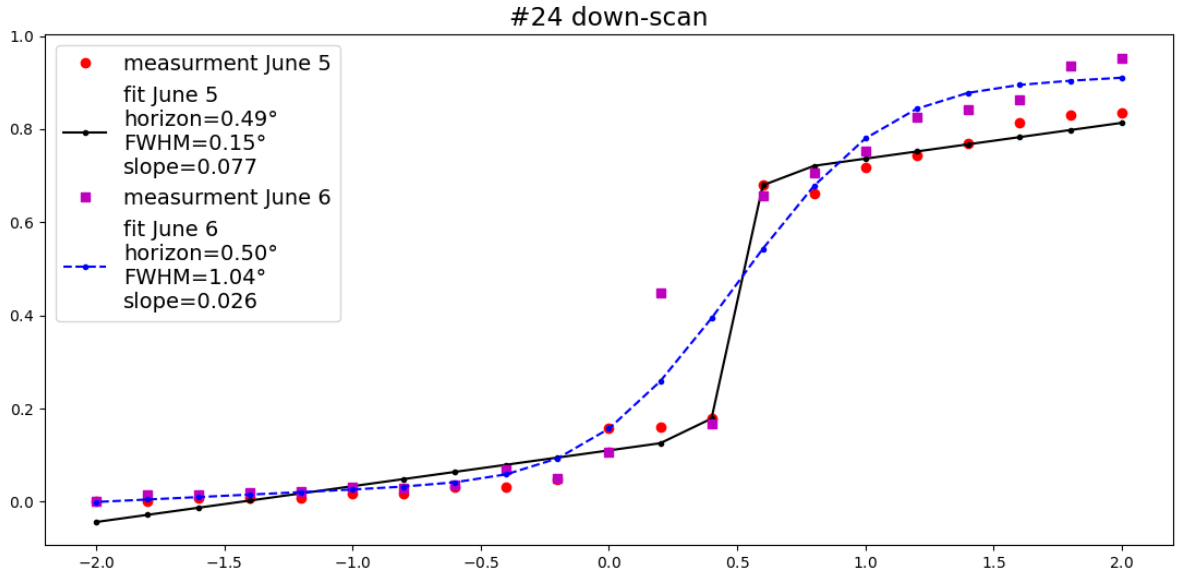

Figure 15: Example showing ambiguity of slope and width of error function for imperfect measurements. Measured (red circles and magenta squares) and fitted (solid black line and dashed blue line) upward horizon of instrument #24 in the visible on June 5th and June 6th as indicated in the legend.

The time series for the horizon (left panel) and FWHM (right panel, left axis) and slope (right panel, right axis) for downward scans of instrument #24 are shown in Fig. 16. Several instruments had elevation scan sequences that included more frequent outliers, missing measurements and repeated angles.

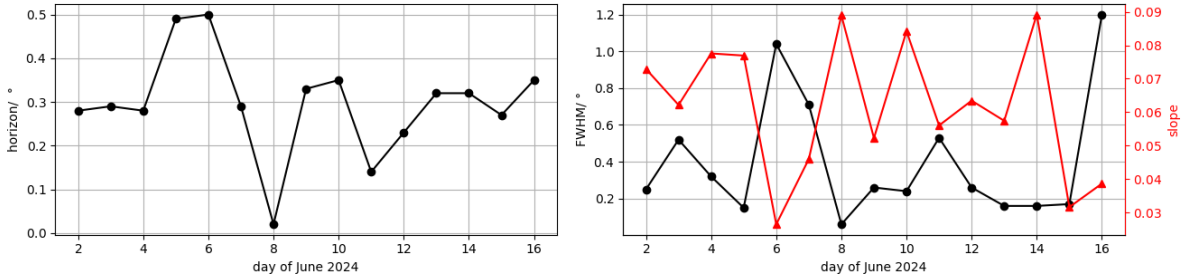

Figure 16: Time series of fitted horizon and FWHM for the upward scans in the visible channel of instrument #24

4.2 Overview of Horizon scans for all instruments

While in general, elevation angles in the range [-2, 2]° were used for the fits [4], instrument #28 FOV was blocked at -1.5°, likely by the scaffolding railing. Therefore the fitting range was adjusted to [-1.2, 2]°.

For the purpose of horizon scans, data files from two NO_2 products with daily noon reference were used: For the visible channels, NO2VIS_DAILYREF was used, and for the ultraviolet (UV)-channel, NO2UV_DAILYREF was used.

After the CINDI-3 WS in Heidelberg (17th March – 19th March 2025), it was decided to recommend a filtering of the data based on fitting RMS to allow obvious outliers to be filtered out. The recommended approach is described in Subsect. 2.2. Unfortunately, several groups also applied filtering on the relative intensity (INORM) data during the horizon scan time-slot (11:45 – 11:55 UTC). For a number of instruments, this resulted in much fewer measurements being present in the horizon scans. This heavily impacted the fit quality.

It was hence decided to replace the INORM data during the horizon scan measurement time slot in the final files for the following instruments by the data delivered for use at the CINDI-3 WS: #2,

[4] larger ranges were tested but more often than not resulted in worse fitting results, also due to the uneven spacing in elevation angle steps beyond the above indicated range

#6, #32, #34, #35, and #36 for the visible channel and #2, #5, #6, #32 and #33 for the UV-channel. In the final dataset, fitting results (such as DSCD or RMS) for the measurements during the horizon scan are set to Fill-value. Note that the fitting results for measurements taken during the horizon scan time slot are not used for any purpose.

In Fig. 17 and Fig. 18, we show time series of the fitted horizon from the relative intensity from the visible channel (using the INORM delivered in NO2VIS_DAILYREF files) and the UV channel (using the INORM delivered in NO2UV_DAILYREF files), respectively.

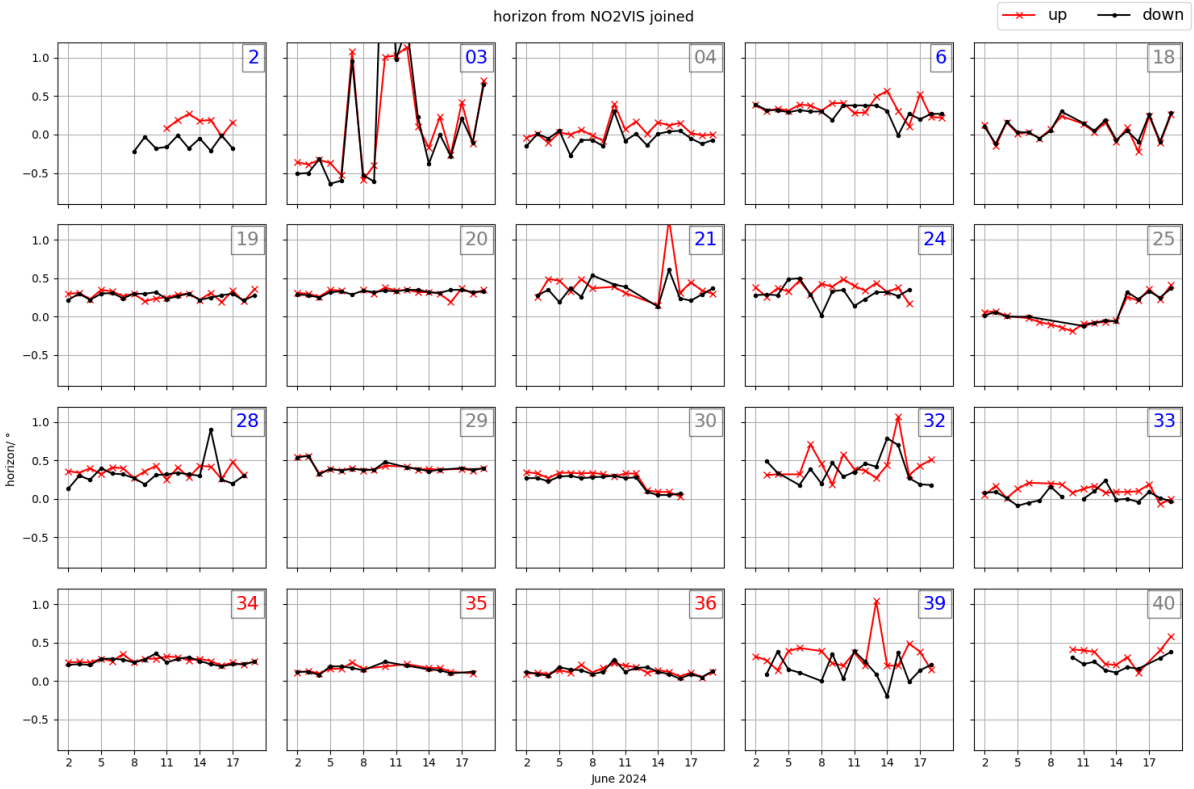

Figure 17: Fitted horizons as a function of time for different instruments using the relative intensities submitted in the NO2VIS_DAILYREF files. The instrument numbers are color coded based on instrument type, see Table 2

We summarize these results in terms of means and standard deviation (both up- and down- scans, as well as in both channels) in Fig. 19.

From Fig. 17 and the blue and cyan curves in Fig. 19 it can be seen that in the visible channel, a range of instruments (#19, #20, #34, #35, #36) were extremely stable during the measurement period (considering up- and down scans separately), varying in a horizon range, Δhorizon less than 0.3°. Some other instruments were adjusted once during the measurement period but otherwise are also very stable (#25, #29, #30) which is only visible in Fig. 17 since the detail of adjustment is hidden in the more information-dense plot of Fig. 19.

As alternative way to compare the different reaction of upward and downward scans in the two channels, we show the standard deviation of the horizon in visible (left panel) and the UV (right panel) channel in Fig. 20. It can be seen that the instruments generally have a similar spread of fitted horizons for the down-ward and up-ward scans, they are close to the gray 1:1 line. Note (right panel) that instruments #16 and #17 did not have sufficient down-scan horizon measurements and are hence located on the gray dashed line, their down-scan horizon was manually set to 1.

The large variation in horizon seen for instrument #3 was real (c.f. Fig. 17, Fig. 19 or Fig. 20). The horizon scan information was used by the data providers to interpolate the delivered DSCDs for VEA 2°, 3°, 4°, 6°, and 8°.

4.3 Differences between upward and downward scans

As already visible in Fig. 17 andFig. 18, most instruments show similar fitted horizons, independent of scanning in the down-ward or up-ward direction. However, some instruments show a systematic difference, such as #5. The instrument owners of that instrument are aware of the backlash issues and consider it not important since in operational set-up, down-scans are never performed. A systematic investigation of the differences between up-ward and down-ward differences is presented in Fig. 21.

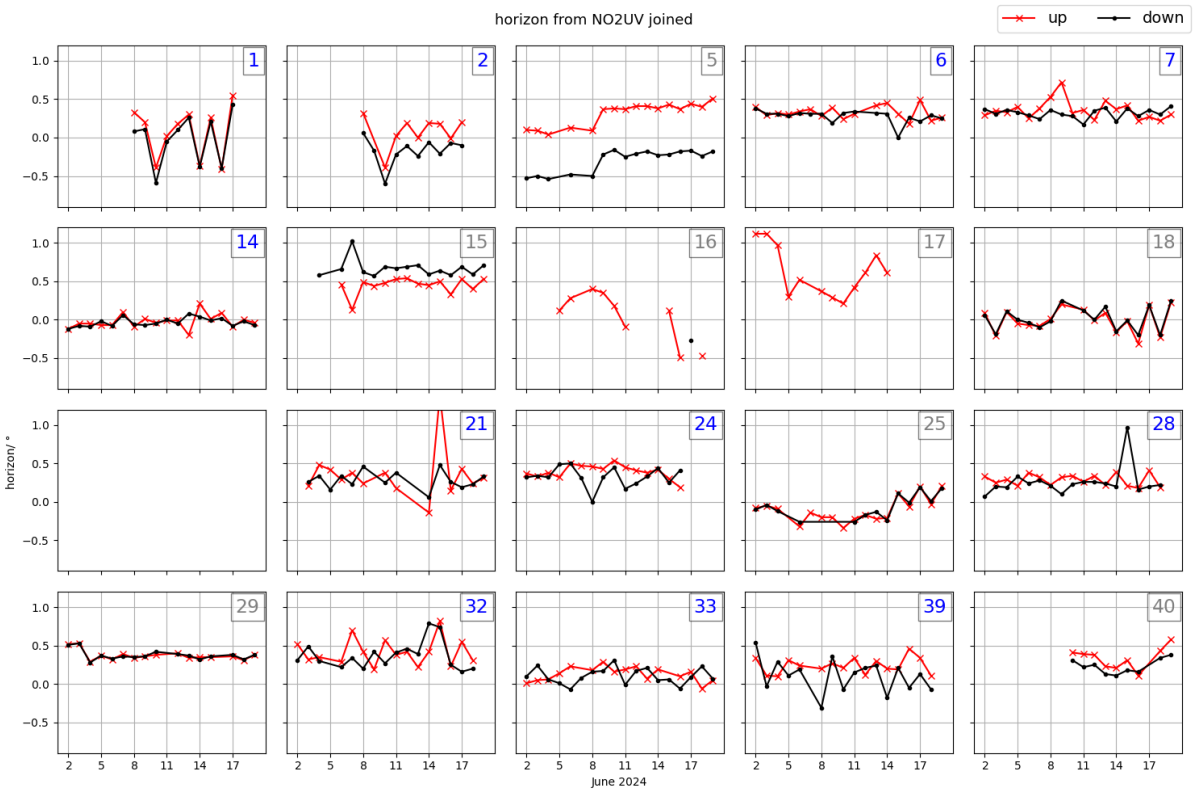

Figure 18: Fitted horizons as a function of time for different instruments using the relative intensities submitted in the NO2UV_DAILYREF files.

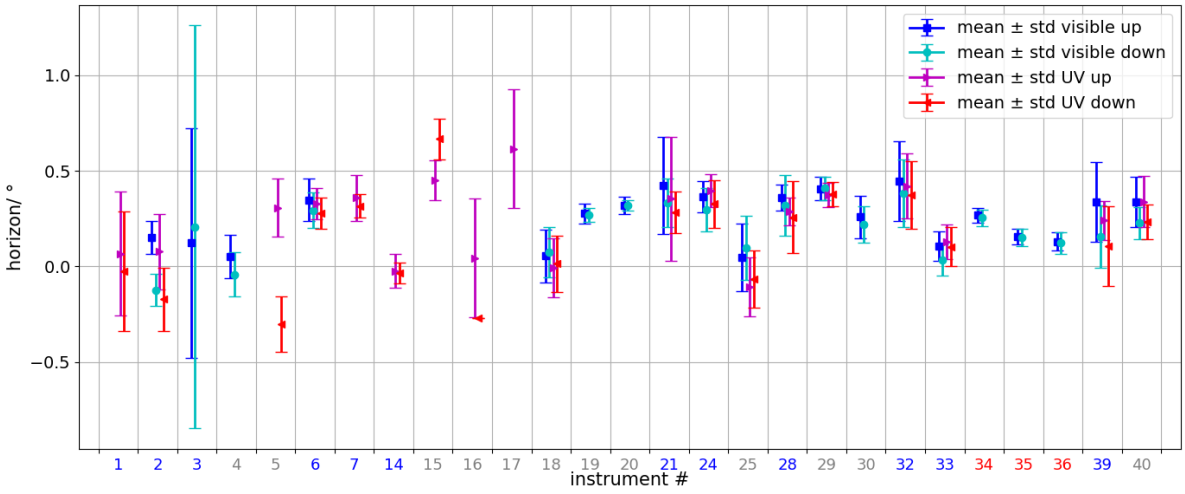

Figure 19: Mean and standard deviation σ for all instruments for up- and down- scans in both channels, as indicated in the legend.

It can be seen that for most instruments, the mean difference is very close to 0, indicating no such instrumental backlash issues. Apart from #5, smaller issues are seen for instruments #1, #2, #4, and #15. It is further visible that the range of differences between up- and down-ward scans are larger but not-systematic for instruments #3, #21, #28, #32 and #39.

4.4 Differenecs between the two channels

It is also of interest to investigate the differences of the fitted horizons between the two chanels. Systematic differences can arise from not perfectly parallel aligned entrance optics for the two channels. In the light of the aforementioned backlash issues for some instruments, and to simplify the analysis, we restrict this analysis to up-ward scans. Fig. 22 shows the difference between fitted horizons for the up-ward scans of the visible and UV channel for those instruments that provide at least 1 day of coincident horizon measurements in the visible and UV channel. It can be seen that there are 3

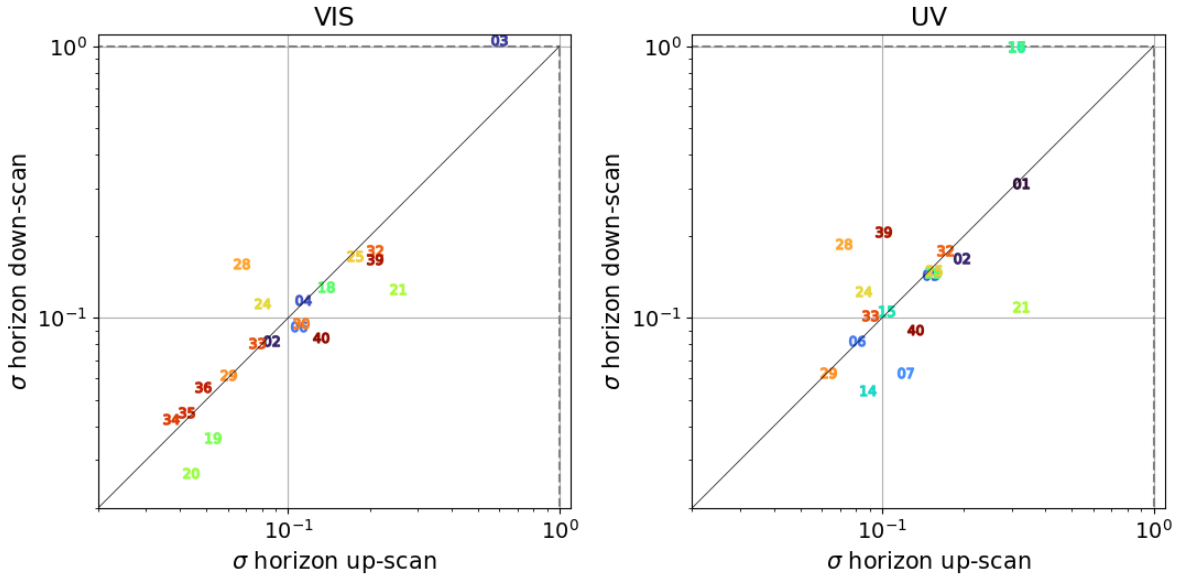

Figure 20: Standard deviations σ of fitted horizons in visible (left) and UV channel right for up-scan (x-axis) against down-scans (y-axis). Those instruments that either have none or only 1 up- or down-scan horizon, are assigned a value of 1 and appear hence on the gray dotted lines. This is only the case for #16 and #17 in UV. #3 has in fact a σ horizon larger than 1 for the visible down-scans.

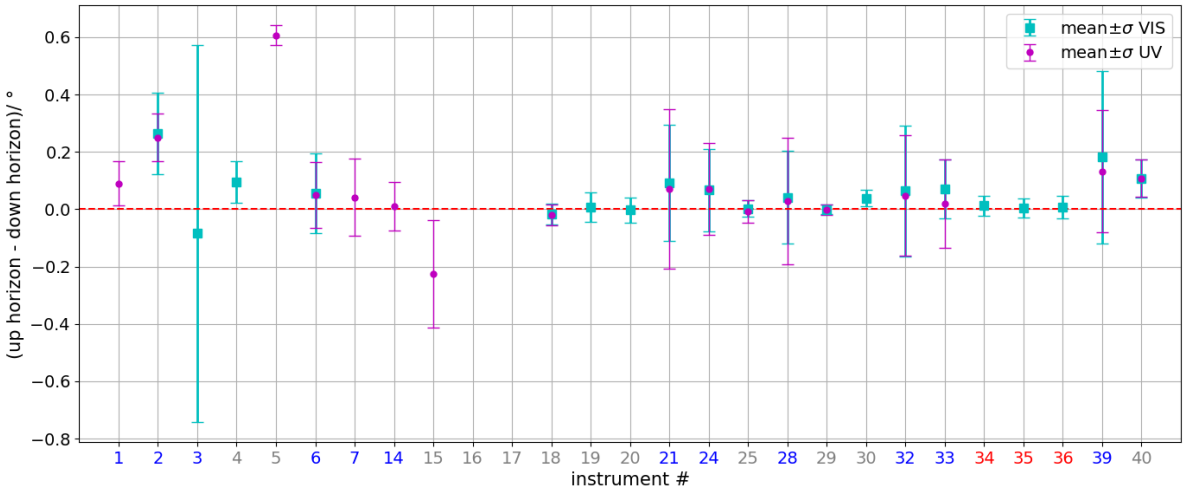

Figure 21: Mean differences (± standard deviation) between fitted horizon of up-scans and down-scans in the visible (cyan squares) and UV (magenta circles) for all instruments that provide at least 1 day of coinciding up- and down-scans.

instruments (#18, #25, #29) for which the difference in the mean is larger than twice the standard deviation. Together with the observation that that the difference between upward and downward scans in the two channels is both very similar and very small (compare Fig. 21), this further suggests that the differences between the two channels are systematic and hence indicate a misalignment of the optics of the two channels.

It is further visible that instruments #2, #21, #32, and especially #39 have large random differences in the fitted horizon between the two channels. This points rather to problems in the fit (possibly due to too few measurements or incorrect elevation angles[5]) than to issues with misalignment of the optics of the two channels.

[5]The Skyspec instruments have in common that the instrument software prevents physical elevation angle increment changes if the difference between the original and new angle is smaller than the measurement error from the inclinometers. This causes sometimes a mismatch between the actual angle of the motor and the angle saved in the measured data

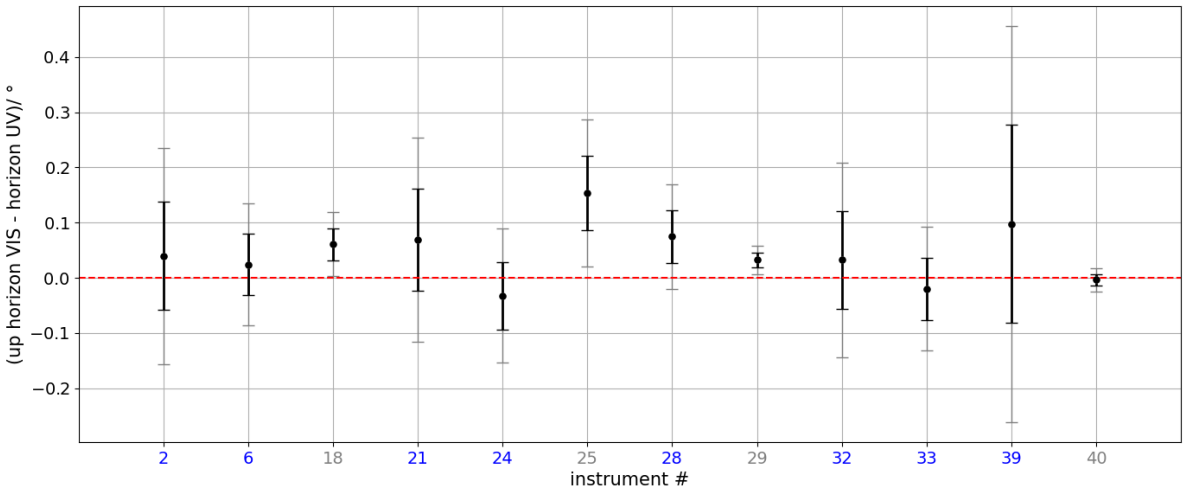

Figure 22: Mean differences ($\pm\ \sigma$ black, $\pm\ 2\sigma$ gray) between fitted horizon of up-scans in the visible channel and up-scans in the UV for all instruments that provide at least 1 day of coinciding up-horizon measurements in the visible and UV.

4.5 Final considerations

While an incremental step of 0.2° in VEA is good for many instruments, SkySpec instruments have the limitation of not adjusting to a new VEA when the difference between old and new VEA is smaller than the accuracy from the inclinometer reading. Hence it might be advisable to leave the choice of VEA increments during the horizon scan more flexible and up to the PI.

It is advisable to clearly state to data submitters that any filtering for bad DSCD fitting results based on e.g. RMS should not be applied to the INORM, especially not for the INORM during horizon scans.

While it is interesting to see the backlash issues of some instruments, since most measurement protocols during operational use of the instruments foresee only 1 scanning direction for elevation scans, it might be better usage of the time to either take two consecutive zenith scans in the same scanning direction or allow more time for a single scan.

5 Analysis of zenith pointing measurements from 50° to 92° SZA

The DSCD data from the inter-comparison campaign were evaluated separately for zenith-only during times of median to high SZA and elevation scan measurements during the day. The former comparison is described in this section for a subset of products, NO2VIS, NO2UV, O3VIS and O4VIS and a subset of reference types, DAILYREF and FIXREF. We first outline the construction of the baseline datasets in Subsect. 5.1 before presenting the results for the different products in Subsect. 5.2.

5.1 Description and motivation of comparison approach

The MAX-DOAS measurements were scheduled daily starting at 04:10 UTC and ending at 19:09 UTC with hourly cycles as displayed in Fig. 2. This translates roughly to 85° SZA in the morning (85.2° at the start of the intensive phase of the campaign and 84.7° at the end of the intensive phase) and 85° SZA in the evening (85.4° at the start, 84.1° at the end). Before and after these MAX-DOAS measurement schedule, the instruments were operated in zenith pointing mode and most instruments covered SZA at least up to 92°. As for the measurements during the hourly cycle (c.f. Fig. 2), the total measurement integration time was set to 60 s.

While we match each measurement for the MAX-DOAS elevation scans without performing any time interpolation, the comparison of the zenith measurements follows a different approach. The interpolation was performed on an equidistant SZA grid with a spacing of 0.15° between 50° and 92°, only interpolating points if the nearest point is closer than 0.15° [6]. This ensures a maximum of measurements to be compared while avoiding insensible interpolations. The choice is motivated by considering the SZA difference during the course of a day considering time differences of 1 min, see upper panel of Fig. 23 and the time difference corresponding to an SZA difference of 0.15° during the day (see same figure, lower panel).

As is evident from Table 2, there are only 8 instruments that have detector cooling to below -10° C and additionally 4 instruments with a temperature stabilization to 10° C or 5° C, while the majority of instruments has either a detector temperature stabilization at 20° C (13 instruments) or no direct temperature stabilization of the detector (8 instruments).

Under low light conditions, a low detector temperature is critical since the signal is weak (which also means that the shot noise is low since it scales with the square root of the signal) and the detector noise becomes a dominant source of error since it scales with the detector temperature. A cooled detector drastically reduces thermal noise and hence improves the signal-to-noise ratio (SNR).

In Fig. 24[7], we show the frequency of the logarithm of the relative intensity, INORM, (upper left panel) and the median of the log of the relative INORM (scaled by the 90th percentile since the raw relative intensities vary over several orders of magnitude between the instruments) as a function of SZA (lower left panel). As expected, the relative intensity decreases with SZA. Considering the median RMS as a function log of the scaled INORM (lower right panel), we see that the increase in RMS with decreasing INORM for instruments with a cooled detector (blue) is much flatter. This in turn translates into a flatter increase of RMS with increasing SZA (upper right panel). From the lower left panel of Fig. 24 it is evident that the median (in each SZA bin) of the logarithm of the scaled relative intensity decreases with increasing SZA almost equally for most instruments. #20 seems to behave very differently spanning a much larger range in INORM. However, it was clarified with the instrument providers that this arises due to a bug in the data format conversion and is not real. As other instruments in the SZA range [80, 92], #20 varies roughly 3 orders of magnitude and not 5.

If we use the median of all measurements as the baseline comparison, then those instruments with a cooled detector, and hence more trustworthy measurements under low light conditions, would be penalized. The fitting RMS is inversely related to the SNR and hence offers itself naturally as a weight to calculate the mean, instead of using the median. Hence, we use a weighted mean as the baseline comparison, where the weight is calculated from the inverse-RMS of each measurement.

Only measurements at times at which at least 10 instruments data provided interpolated measurement values are included in the mean data-set. This also affects the total number of measurements used to calculate the valid percentage. The total number equals the total number of measurements for which a mean was calculated, i.e. times for which at least 10 instruments provided measurements. This explains the different total number (last row) in the statistic figures Fig. 34 to Fig. 41 when considering different product-reference type combinations.

[6] to avoid edge effects, the interpolation was performed on [49, 93] and a mask was applied to cut out the region [50,92].

[7] After final submission, it was discovered that the INORM provided by instrument #20 had a bug in its calculation: The total intensity was normalized twice by the integration time. This explains the large range of INORM for that instrument in these figures.

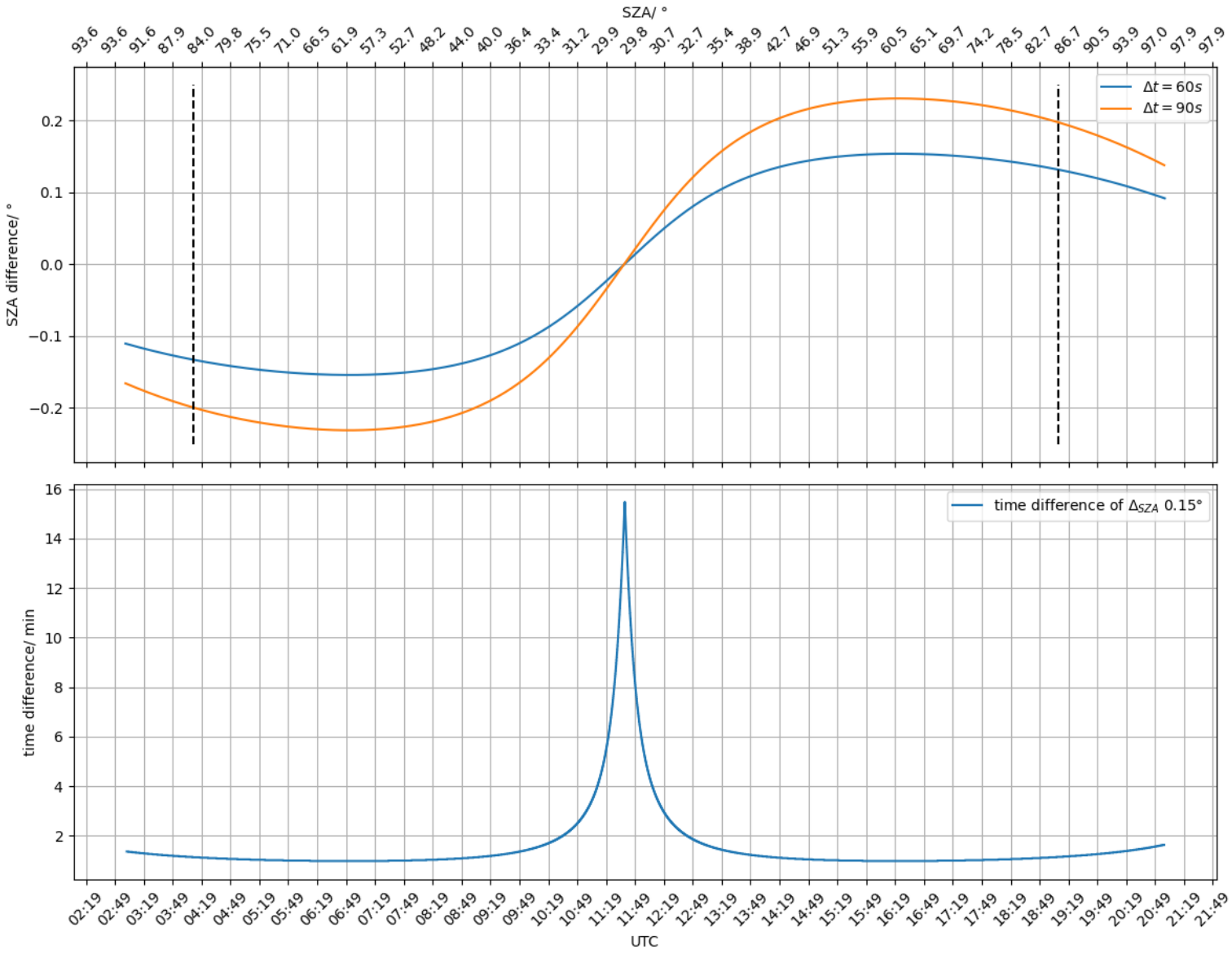

Figure 23: Upper panel: Differences in SZA between 60 s and 90 s during the course of a measurement day. Indicated with vertical dashed lines are the start and stop of the MAX-DOAS hourly cycles. Lower panel: Time difference corresponding to a difference of SZA of 0.15° over the course of a day.

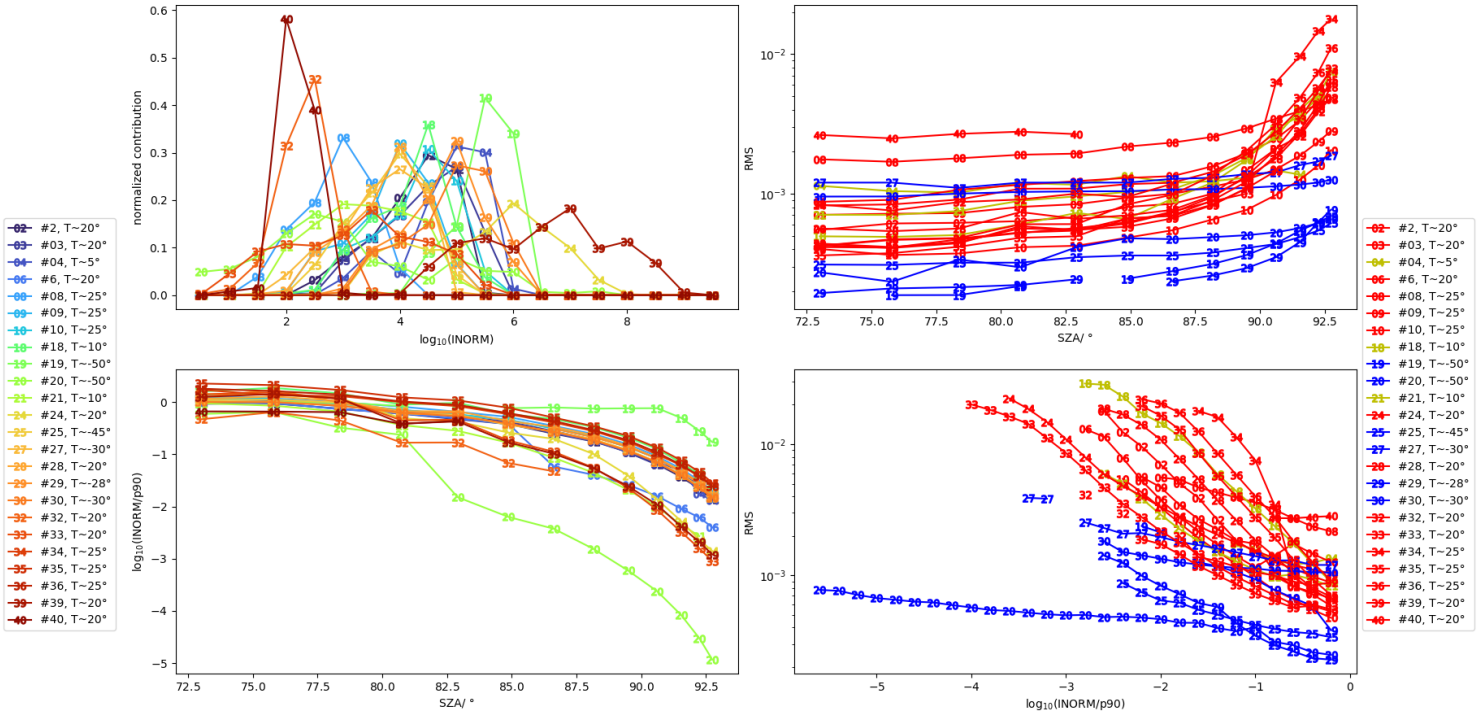

Figure 24: Distribution of the logarithm of relative intensity for each instrument (upper left), median of logarithm of scaled (by 90th percentile) relative intensity as function of SZA (lower left), median RMS as function of SZA (upper right) and scaled logarithm of relative intensity (lower right) for product-reference type combination NO2VIS_DAILYREF

5.2 Comparison of individual instruments to the baseline weighted mean

Fig. 25 shows an example time-series (left panel) of the mean (black) and instrument 6 (red) for O3VIS_DAILYREF and the scatter plot between the two (right panel, mean on x-axis) with color coded SZA.

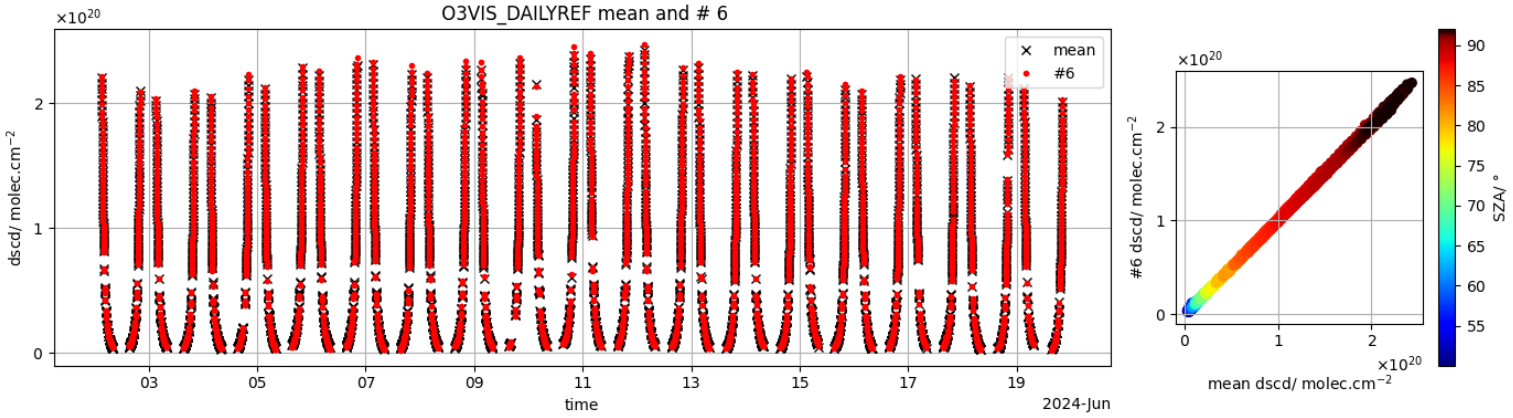

Figure 25: Example of time-series and scatter plot for mean and instrument 6 for O3VIS_DAILYREF.

Fig. 26 to Fig. 33 show the regression scatter plots for all individual instruments (one figure per product-reference type combination, one panel per instrument as indicated in the center top of each panel, individual interpolation on y-axis) over the weighted mean (on x-axis). The RMS of the difference between individual interpolated measurements and the baseline was calculated, as well as the correlation coefficient R and the Theil-Sen regression slope s and intercept I (Theil, 1950; Sen, 1968). The corresponding regression analysis statistics of these indicators are shown in Fig. 34 to Fig. 41.

The colored areas in the latter serve as a visual aid to quickly asses the goodness of a particular statistic of an individual instrument, where cyan indicates and excellent value, green a good value, yellow an acceptable value, orange a marginal value and red a bad value. The same regions were used to assign quality identifiers for each of the instruments, as is described later. The limits were chosen as shown in Table 4.

Category	Correlation r	Slope s	Intercept i	RMS
Cyan (Excellent)	> 0.9997	$\|1 - s\| \leq 0.02$	$\leq 1\%$ of median mean	$\leq 3\%$ of median mean
Green (Good)	$0.997 < r \leq 0.9997$	$0.02 < \|1 - s\| \leq 0.04$	$1\% < i \leq 2.5\%$	$3\% < RMS \leq 7.5\%$
Yellow (Acceptable)	$0.97 < r \leq 0.997$	$0.04 < \|1 - s\| \leq 0.07$	$2.5\% < i \leq 5\%$	$7.5\% < RMS \leq 15\%$
Orange (Marginal)	$0.9 < r \leq 0.97$	$0.07 < \|1 - s\| \leq 0.15$	$5\% < i \leq 10\%$	$15\% < RMS \leq 30\%$
Red (Bad)	≤ 0.9	> 0.15	$> 10\%$ of median mean	$> 30\%$ of median mean

Table 4: Quality categories for correlation, slope, intercept, and RMS, based on specified thresholds.

Additionally, for the percentage of measurements, the limits are 95% for excellent, 85% for good, 70% for acceptable, 50% for marginal. Since the percentage of valid data does not enter the overall quality indicator, it is not included in this table.

As expressed in Table 4, both the intercept, i, as well as the RMS are expressed as a percentage of the median (over all considered times, i.e. 2 June – 19 June, each day in the SZA region [50 –92]°) of the reciprocal RMS weighted mean (see Subsect. 5.1). These median means are given in Table 5.

Comparing these (together with the percentage limits presented in Table 4) with the limits from CINDI-2 (c.f. Kreher et al., 2020), shows that the limits used for CINDI-2 on NO2VIS and NO2UV correspond roughly to the yellow and orange region. Limits on the slope in Kreher et al., 2020 were ± 5% and ±6%, respecively (here, these are located in the center of the yellow region), limits on RMS were 8×10^{15} and 1×10^{16} molec.cm-2 respectively (here, these are located in the yellow region) and limits on the intercept were 1.5×10^{15} and 2×10^{15} molec.cm-2, respectively (here this is located in the yellow and orange region). Limits on O4VIS (slope ±5%, RMS 3×10^{42} and intercept 7×10^{41} molec2.cm-5, see Table 4 in Kreher et al., 2020) roughly corresponds to the yellow region. The limits for O3VIS (slope ±4%, RMS 1×10^{18} and intercept 2×10^{17} molec.cm-2, see Table 4 in Kreher et al., 2020) corresponds roughly to the cyan region.

The choice of using one set of limits for all products was taken such that the level of agreement can also easily be compared between the different products. Note also, that the correlation coefficient was introduced as a fourth quality identifier. The weather conditions during CINDI-3 compared to

product_reference type	median mean value
O3VIS_DAILYREF	4.62×10^{19} molec/cm^2
O4VIS_DAILYREF	1.56×10^{43} molec^2/cm^5
NO2VIS_DAILYREF	2.25×10^{16} molec/cm^2
NO2UV_DAILYREF	1.64×10^{16} molec/cm^2
O3VIS_FIXREF	4.56×10^{19} molec/cm^2
O4VIS_FIXREF	2.17×10^{43} molec^2/cm^5
NO2VIS_FIXREF	2.67×10^{16} molec/cm^2
NO2UV_FIXREF	1.94×10^{16} molec/cm^2

Table 5: Median mean values for the different product-reference type combinations between SZA 50° and 92°.

CINDI-2 were worse in terms of cloud cover, temperature and precipitation; this affected the overall performance of the instruments (lower light conditions) but also enhanced the differences between the different instruments due to different light paths in case of broken (and possibly fast moving) cloud cover.

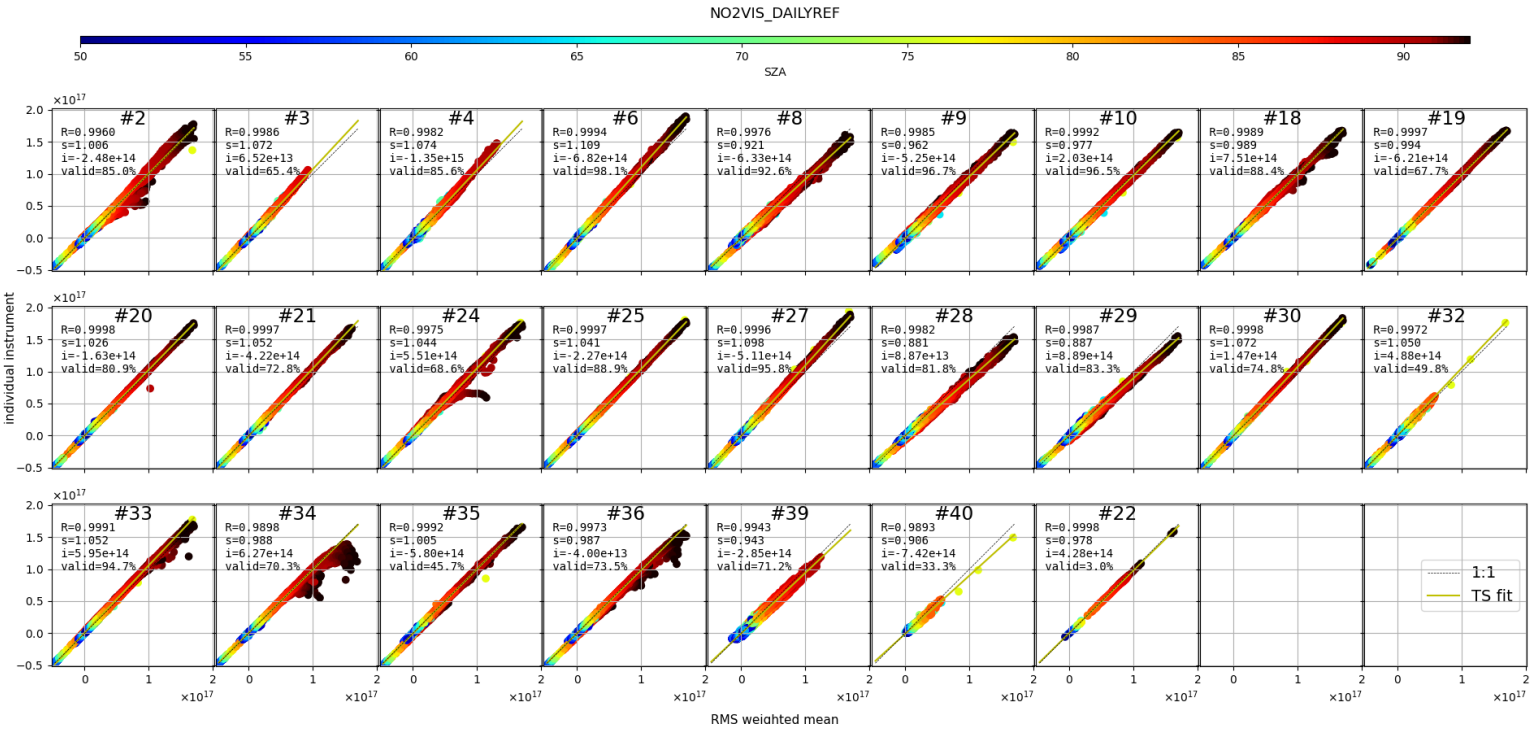

Figure 26: Scatter plots of individual instrument interpolations on a fixed SZA grid over the corresponding weighted mean for NO2VIS_DAILYREF. Details see text.

5.2.1 NO2VIS

The correlation scatter plot for NO2VIS-DAILYREF (c.f. Fig. 26) shows that most instruments correlate very well up to SZA (color coded) up to 92° (maximum shown). The corresponding statistics for this product-reference type combination are shown in Fig. 34. All instruments have an overall correlation coefficient (correlation w.r.t the baseline mean, calculated as outlined above) better than 0.99 (if rounded to 2 decimals) and are hence well in the acceptable region. However, the same is not true for other quality indicators: Instruments #28, #29 and #34 have RMS values judged as bad. Instruments #4, #6, #8, #27, and #30 have two out of four indicators only acceptable.

In order to quantify the goodness of the data of an instrument, we give scores as follows: 0 (in the cyan region), 1 (in the green region), 5 (in the yellow region), 21 in the orange region and 85 in the red region for indicators correlation, slope, intercept and RMS (Note: The percentage of valid data does not enter in this general score).

Hence the best cumulative (over the 4 different indicators) score measurements of an instrument for a particular product-reference type combination can have, is 0 if all indicators are judged excellent. A score below 5 means that no indicator is worse than good, a score below 21 means that no indicator is worse than acceptable and a score below 85 means that no indicator has a bad value. The total

score of the instruments is shown in Fig. 42. Lines are drawn at the positions below which no bad values occur (red), below which no bad and marginal values occur (orange), below which all values are at least good (yellow), below which at least 3 out of 4 indicators have an excellent score and the remaining good (green) and below which all scores are excellent (cyan).

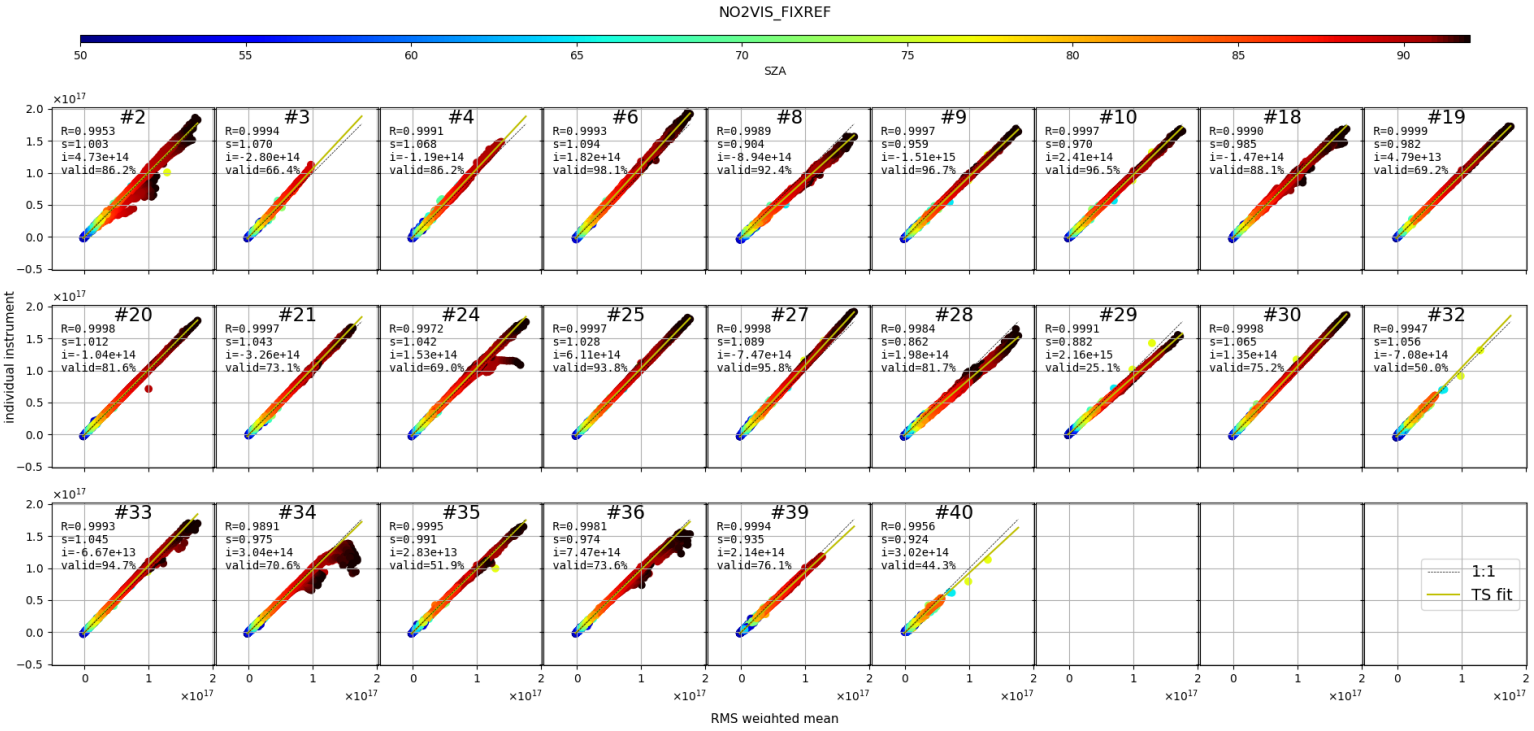

Figure 27: Same as Fig. 26 but for NO2VIS_FIXREF

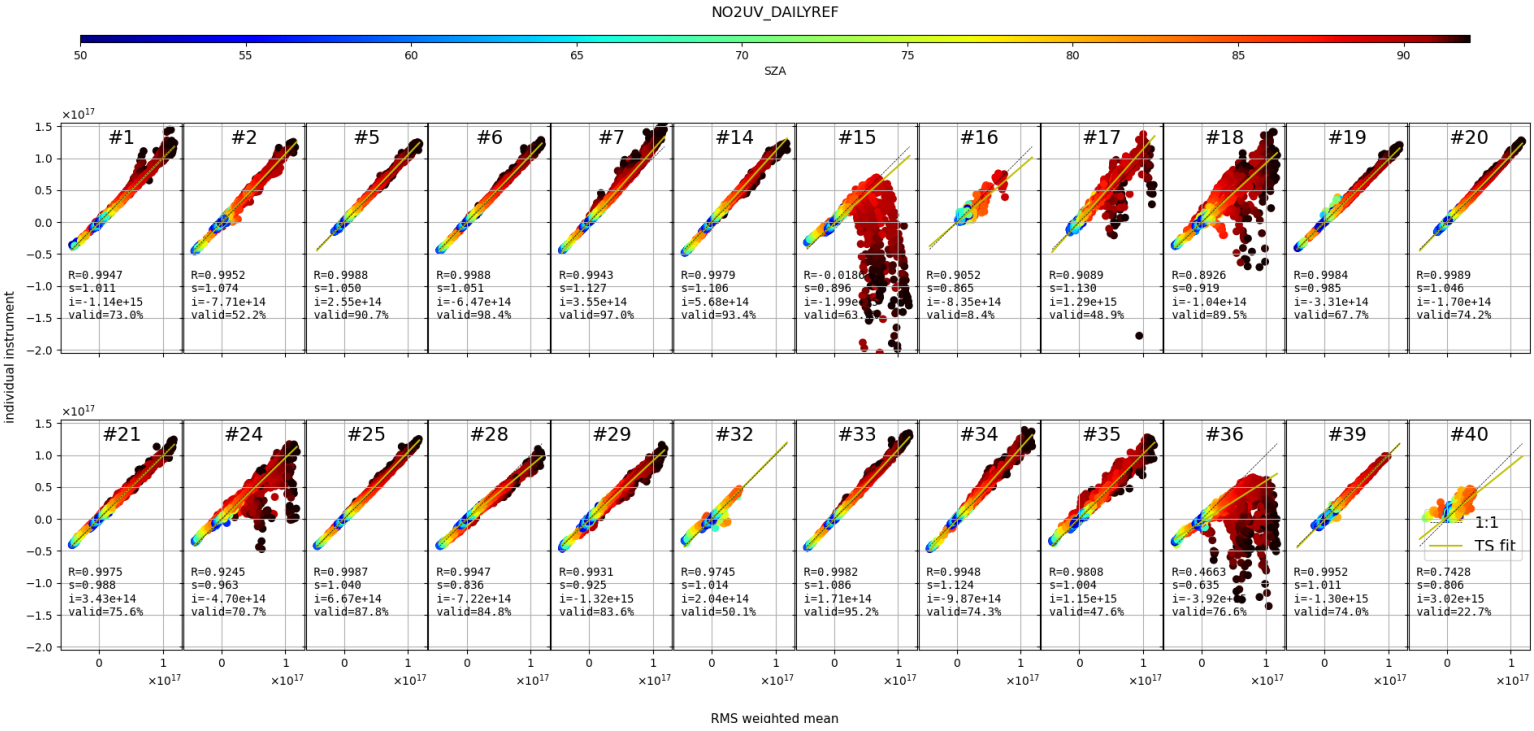

Figure 28: Same as Fig. 26 but for NO2UV_DAILYREF

5.2.2 NO2UV

From Fig. 42 it is evident that only 7 instruments have scores better than marginal for all quality indicators for at least one of the reference types for the NO2UV product (#5, #6, #19, #20, #21, #25, #39), further 6 instruments have no quality identifier judged as bad (#1, #2, #6, #14, #32,

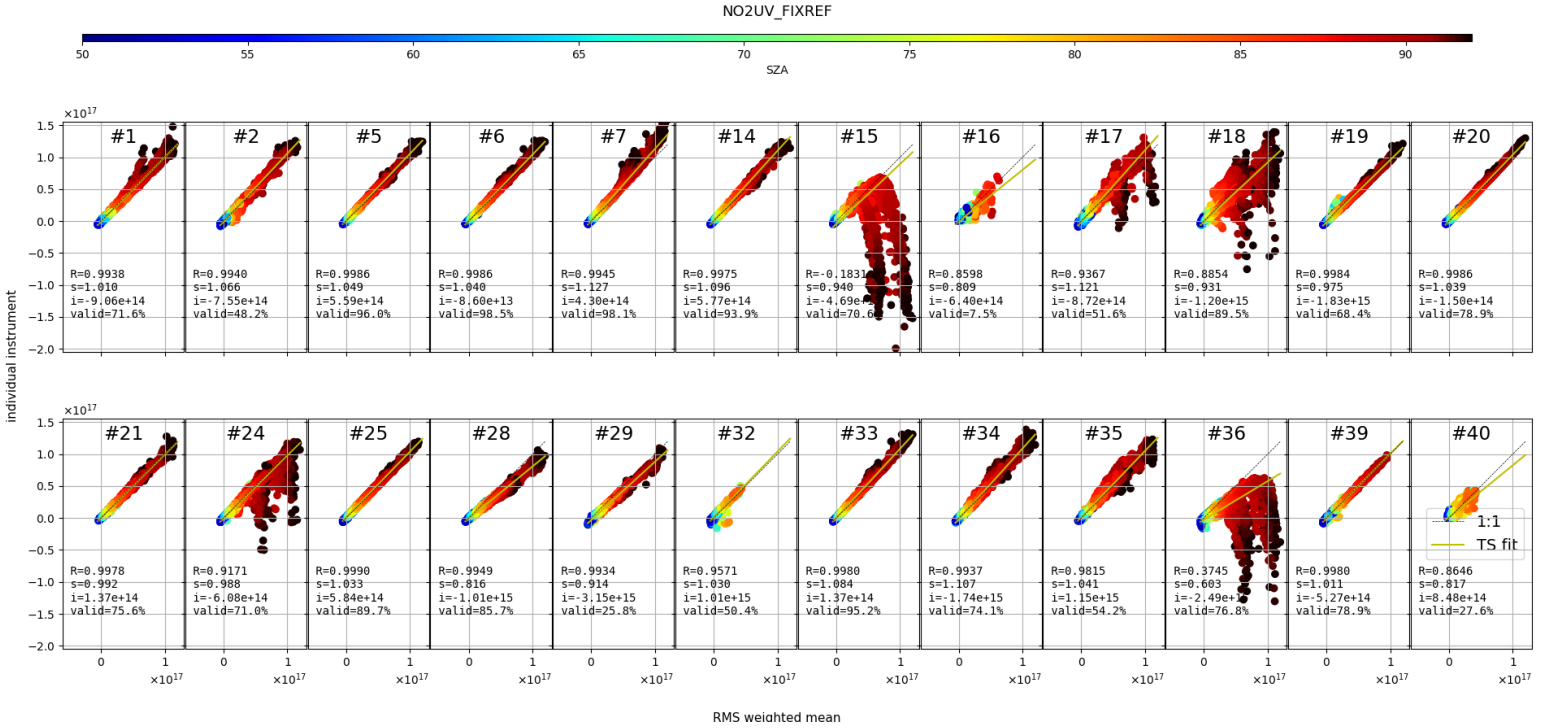

Figure 29: Same as Fig. 26 but for NO2UV_FIXREF

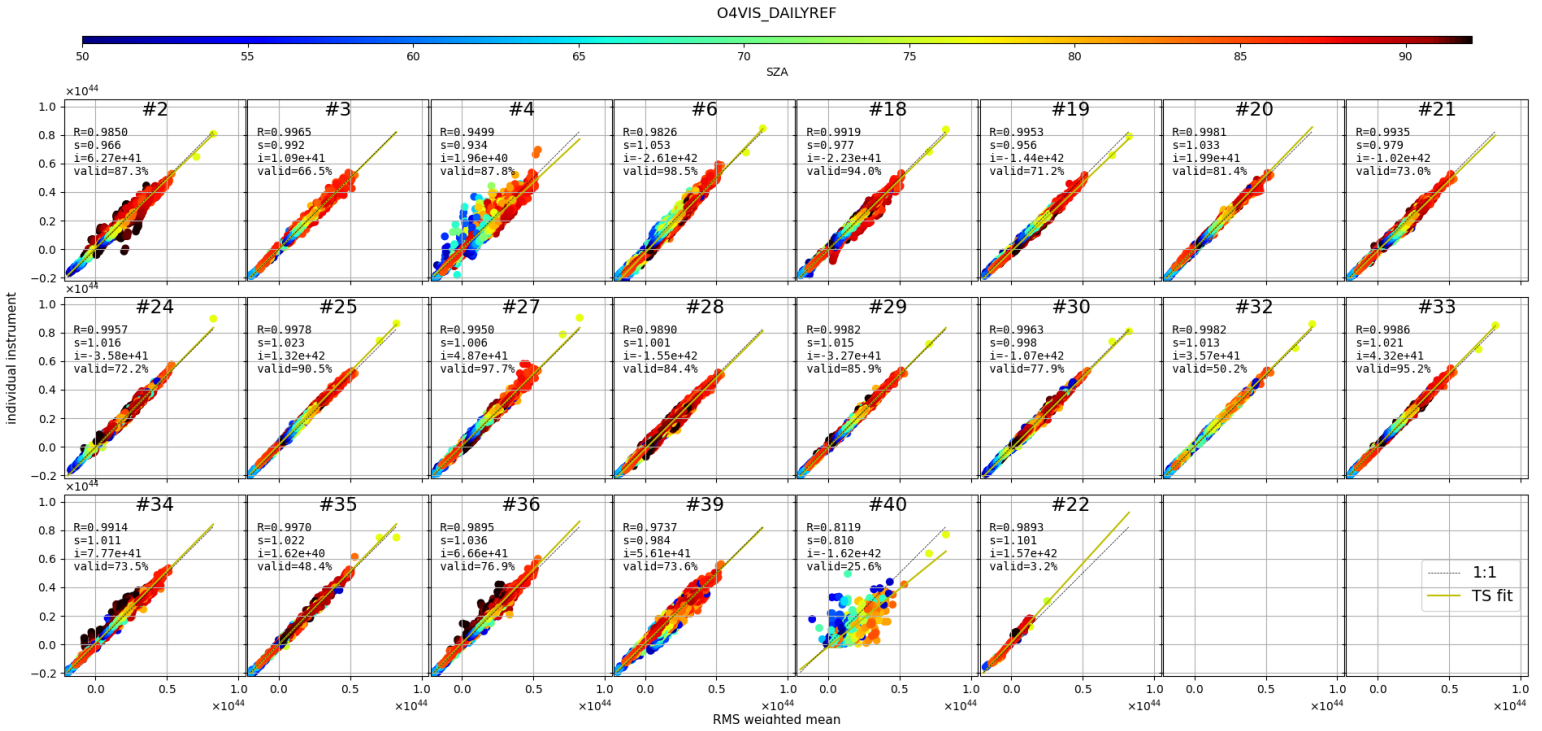

Figure 30: Same as Fig. 26 but for O4VIS_DAILYREF

#33). A bit less than half the instruments have at least one quality identifier for both reference types for the NO2UV product judged as bad (#7, #15, #16, #17, #18, #24, #28, #29, #34, #35, #36 and #40). Fig. 43 shows the same as Fig. 42 but in the SZA range [50°, 90°]. As can be seen, the results do not significantly change if the considered region for NO2UV is limited to slightly lower SZA and in fact sometimes even worsen for other products such as O_3 since DSCD range is reduced. From Fig. 28 and Fig. 29, this was expected, considering the generally larger scatter at all SZA.

5.2.3 O3VIS

For O3VIS, Fig. 42, Fig. 40 and Fig. 41 it can be seen that only one instrument (#34) has one quality idenfier (the RMS) for both reference types judged as bad. Regarding the correlation (both reference

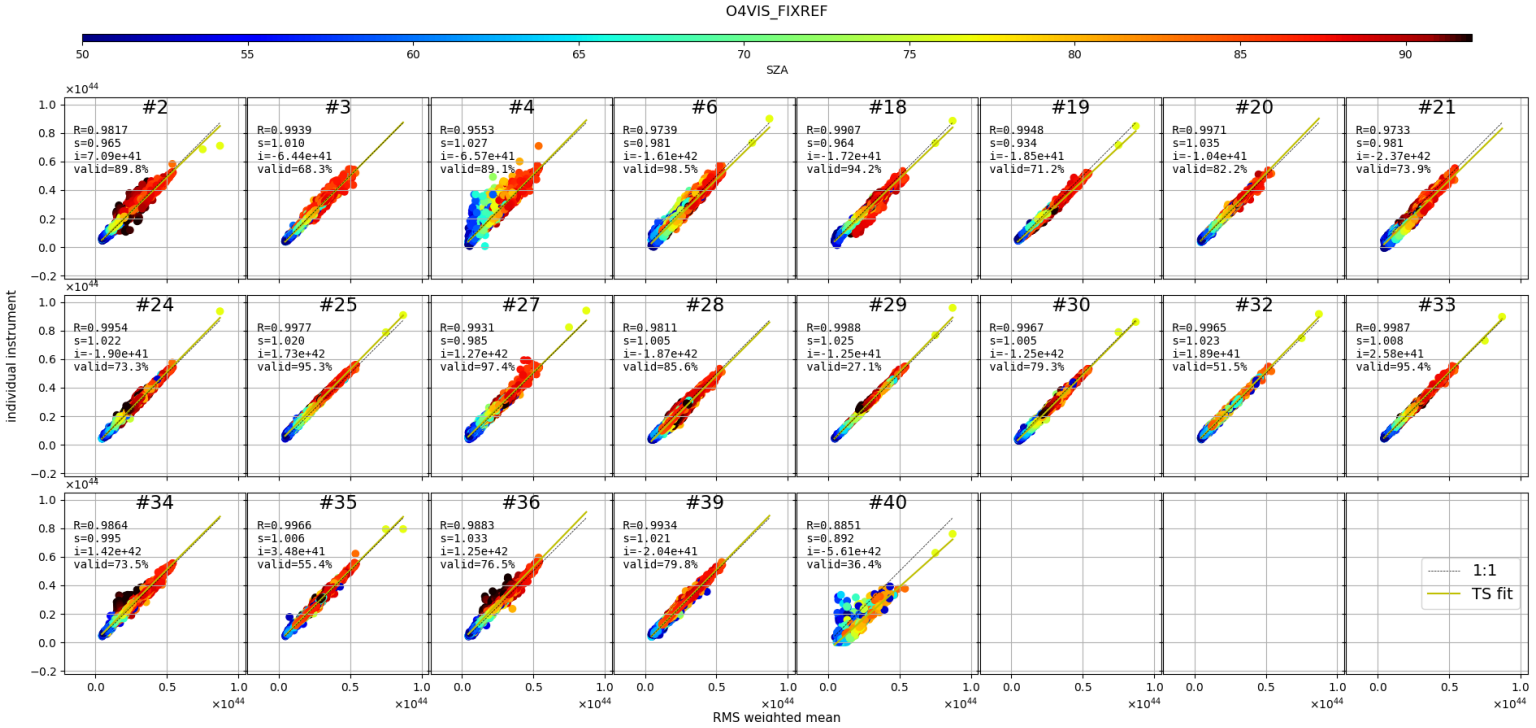

Figure 31: Same as Fig. 26 but for O4VIS_FIXREF

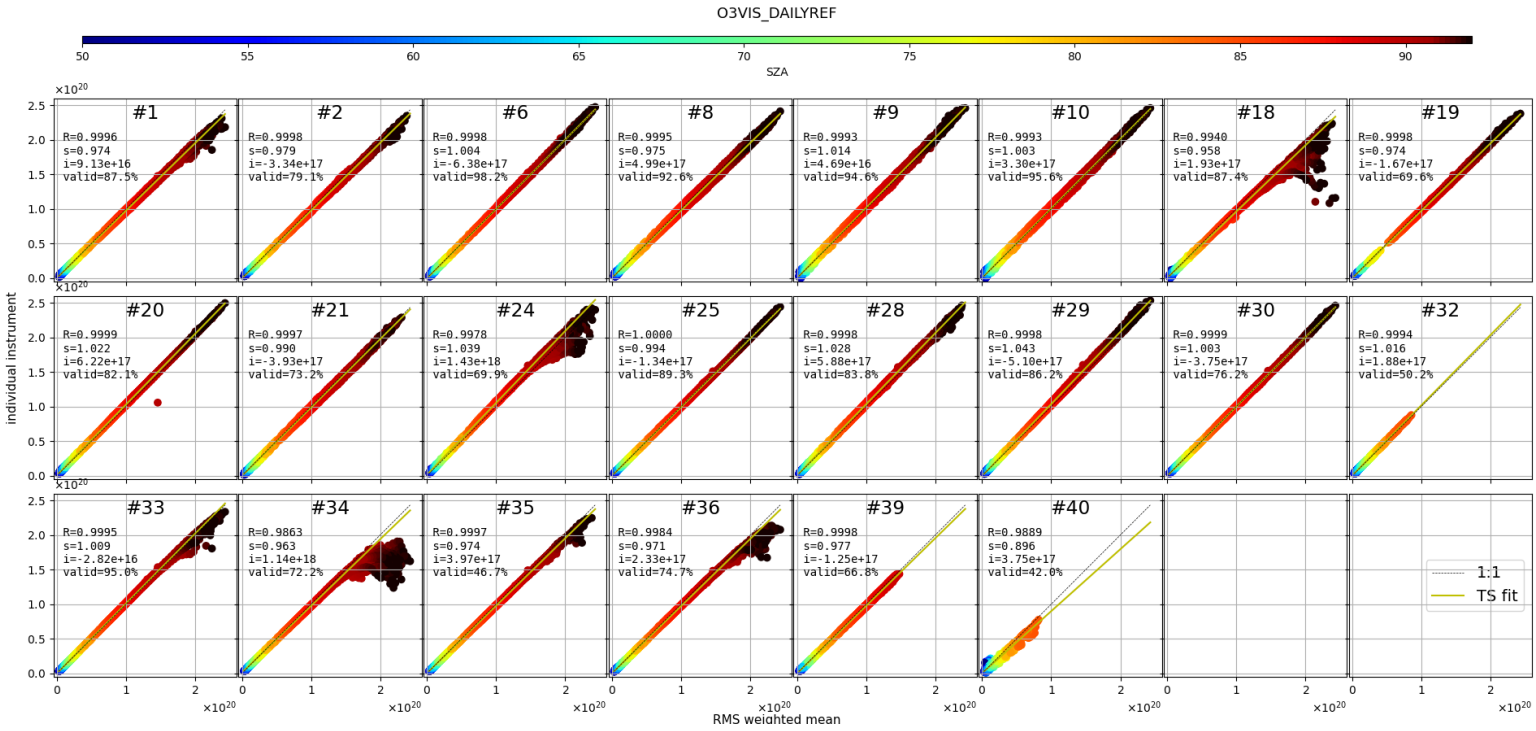

Figure 32: Same as Fig. 26 but for O3VIS_DAILYREF

types), only 2 instruments have a value judged acceptable (#18 and #40), all other instruments have a correlation judged good or excellent. Further it can be seen that two instruments (#9 and #10) have significantly higher intercepts for the FIXREF than for the DAILYREF, but significantly better correlation in FIXYREF than for DAILYREF. This highlights a general and natural trend: While differences in the refrence spectrum for FIXREF will affect mostly the intercept (since it is a difference that is equal for all days), differences in the reference for DAILYREF will mostly affect the correlation (since the difference can vary from one day to the other). An exception from this is #24: For both reference types, the intercept is rather high (but still in the acceptable region) and the correlation at the edge of the good region. This indicates that the offset is unlikely to be related to differences in the reference. For three instruments, #25, #30 and #39, at least one reference type

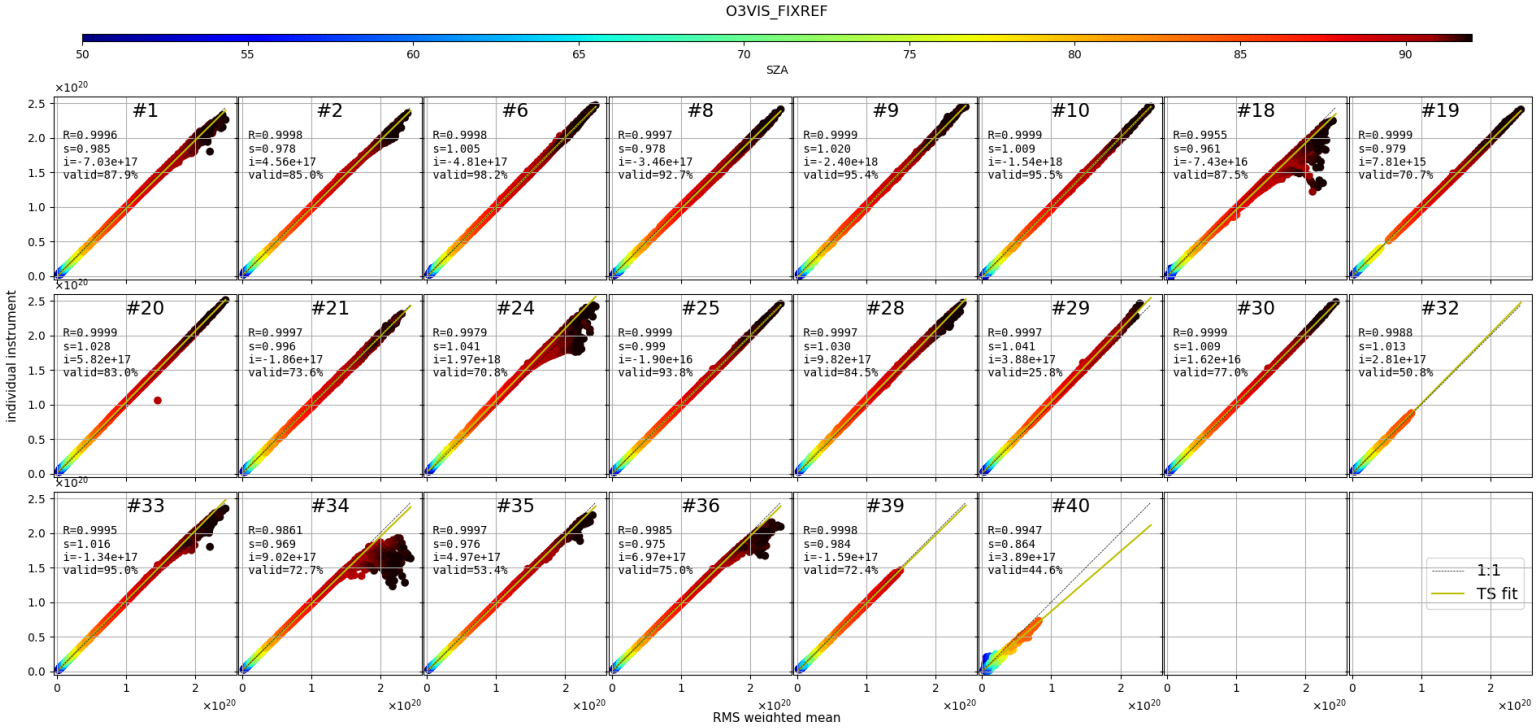

Figure 33: Same as Fig. 26 but for O3VIS_FIXREF

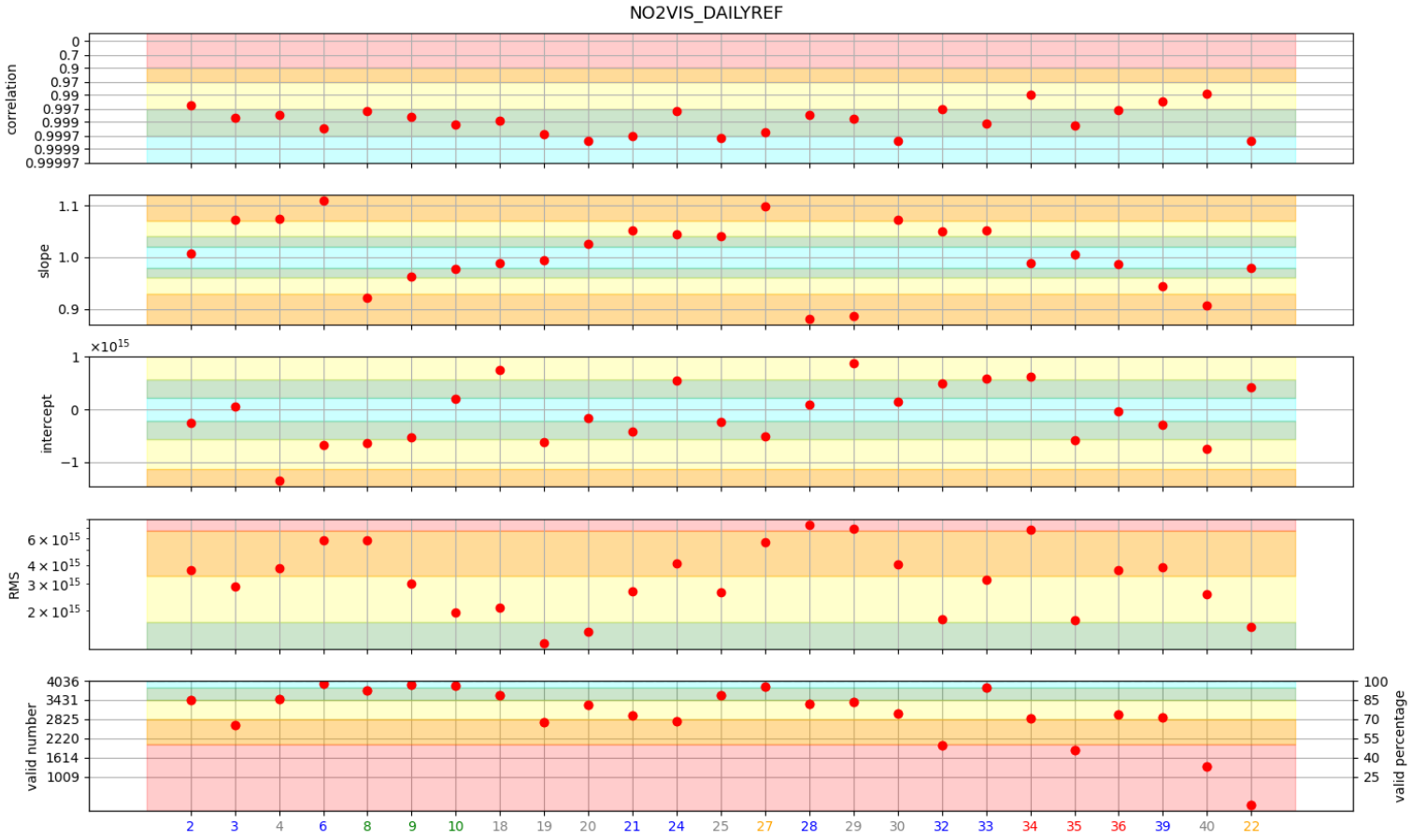

Figure 34: Statistics of the correlation analysis of interpolations at fixed SZA of individual instruments with the weighted mean for NO2VIS_DAILYREF. Details see text.

has all quality identifiers judged as excellent.

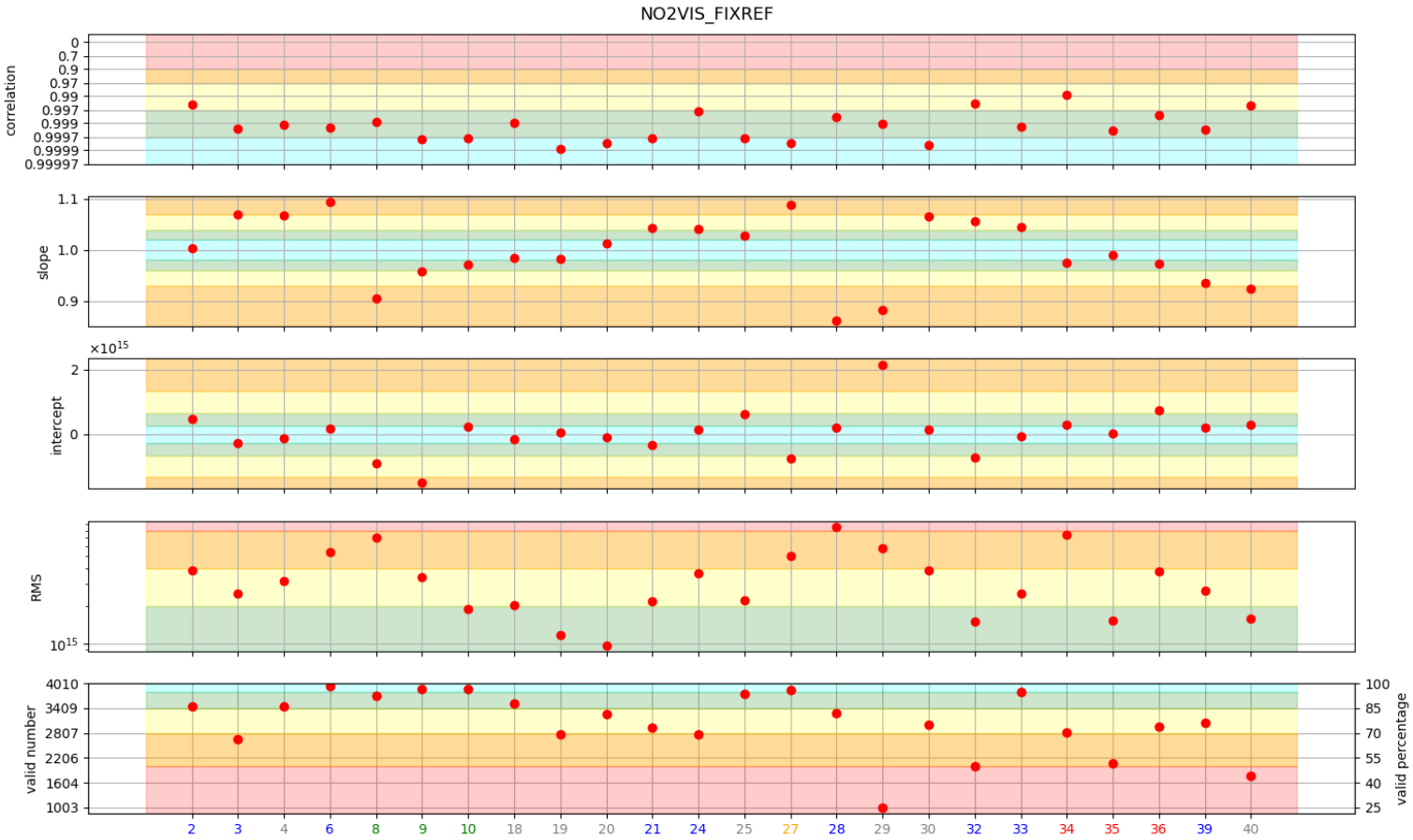

Figure 35: Same as Fig. 34 but for NO2VIS_FIXREF

5.2.4 O4VIS

For O4VIS, the general scatter is higher (no correlation coefficients better than 0.999), but also with only 1 instrument (#40) with a correlation less than 0.9 (but better than 0.7). This is different from NO2UV, where the spread in correlation ranged from 0.999 to 0. Also w.r.t. the RMS, only #40 is classified as bad. Regarding the slope, most instruments (except for #4, #6, #19 and #40) are judged as at least good. Larger differences can be seen in terms of intercept. Overall, only two instruments (#20 and #29) have no quality idenfier worse than good in both reference types. Only one instrumen (#40) has at least 1 quality idenfier bad in both reference types and additional 3 instruments (#6, #21 and #22) have at least one quality identifier judged as bad in one refrence type.

5.3 Final considerations

17 instruments (#1, #2, #3, #4, #5, #8, #9, #10, #14, #19, #20, #25, #27, #30, #32, #33 and #39) have no considered and submitted product-reference type combination judged as bad.

Considering instruments of the same type, the following can be seen: For the Pandora instruments (indicated with red labels, #34, #35 and #36) it can be seen that #35 and #36 have a very similar performance, where only the NO2UV products perform not so good, while more than half the products of #34 are judged as bad. Apart from being operated by two different institutions (#34 by KNMI and #35 and #36 by Luftblick), there is a difference in the additional wavelength calibration: While #34 uses a noon spectrum from 5 June 2024, instruments #35 and #36 use a noon spectrum from 19 June 2024 (c.f. Appendix D).

For Skyspec instruments (indicated with blue labels, #1, #2, #3, #6, #7 #14, #21, #24, #28, #32, #33, #39) it can be seen that the O3VIS product (both reference types) is acceptable or better for all instruments that delivered that product. With the exception of instrument #28 (at least 1 quality identifier bad), for NO2VIS, all instruments are in the range acceptable–marginal. The performance for NO2UV is very diverse (e.g. #21 has all quality identifiers better than marginal while #28 has several quality identifiers as bad).

The 3 SAOZ instruments (indicated with green labels, #8, #9 and #10) are all in the good region for O3VIS_DAILYREF (see note above on FIXREF). For NO2VIS, #8 performs worse than

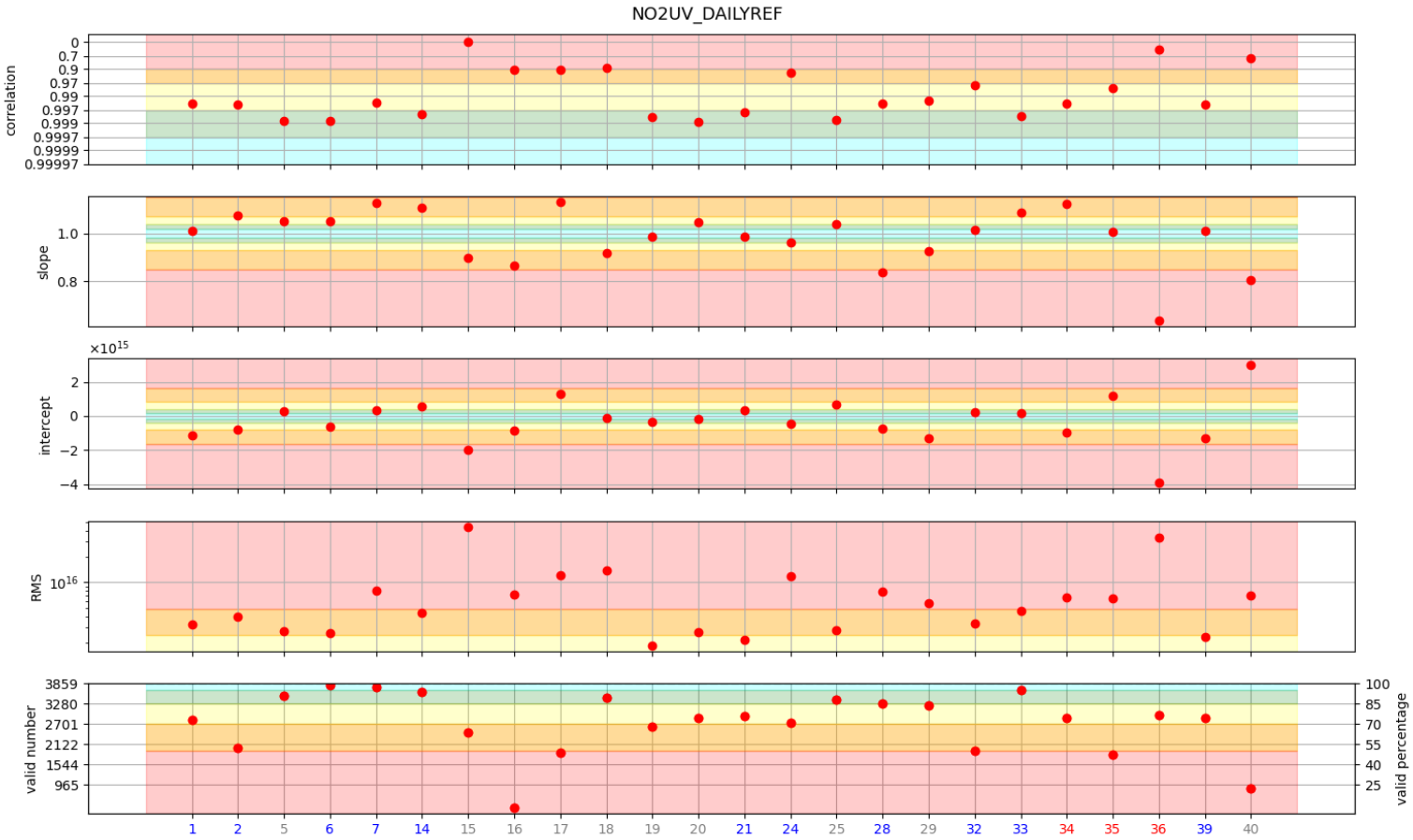

Figure 36: Same as Fig. 34 but for NO2UV_DAILYREF

instruments #9 and #10. As reported in Table 2, the spectral resolution of #8 is worse than the spectral resolution of the other two (roughly double).

Naturally, the diversity for the remaining instruments (indicated with gray labels) that comprises a very diverse group of instruments, is the largest, ranging from 2 instruments that have only excellent quality identifiers for O3VIS (#25 and #30, both have cooled detectors to -45° C and -30° C, respectively) to ones that are in the marginal region for O3VIS (#18 and #40, with temperature stabilization at 10° C and 20° C, respectively). Considering O4VIS, the diversity is even larger, where instruments #20 and #29 (detector cooling to -50° and -28°, respectively) have all quality identifiers at least good, while instrument 40 has the majority of quality identifiers classified as bad.

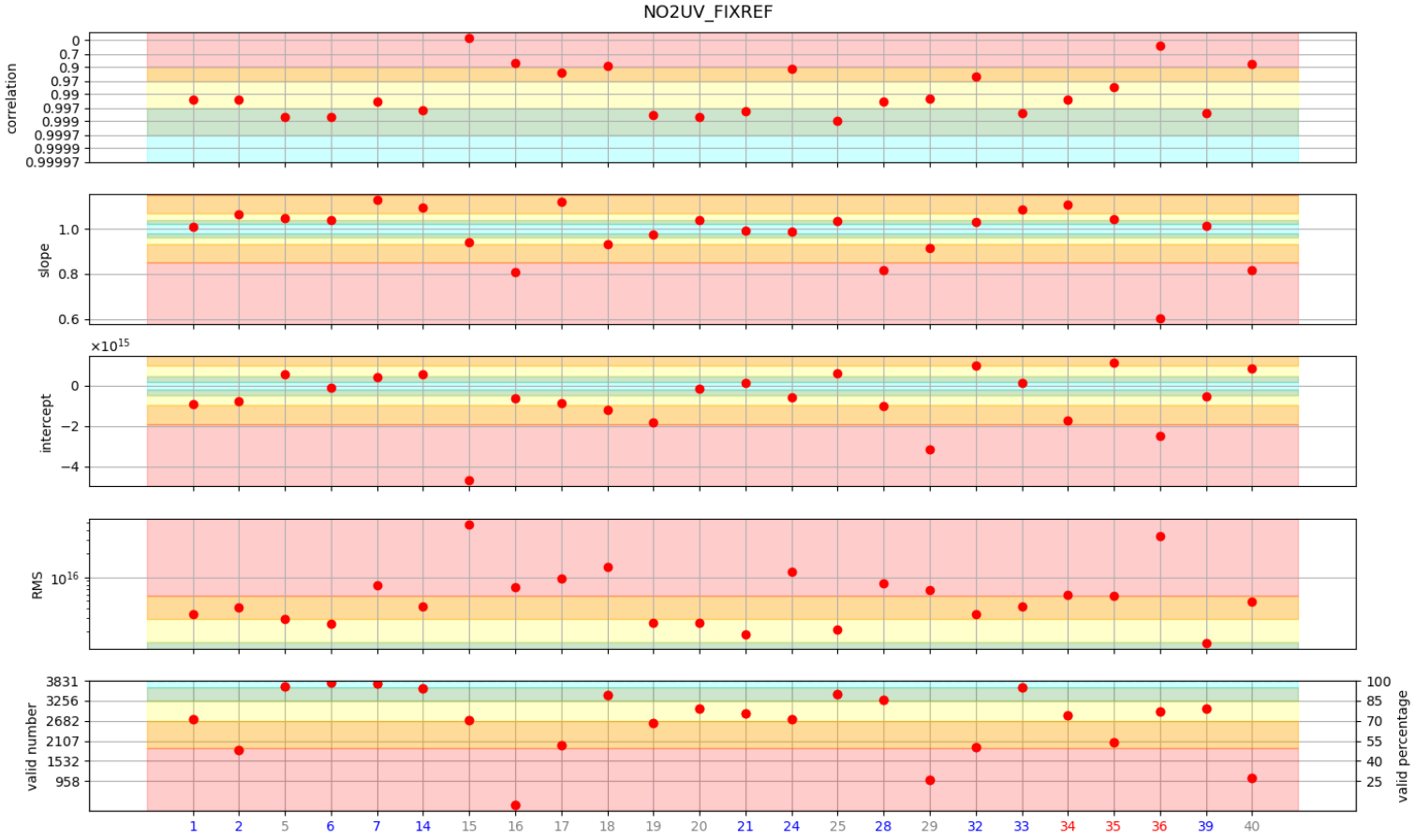

Figure 37: Same as Fig. 34 but for NO2UV_FIXREF

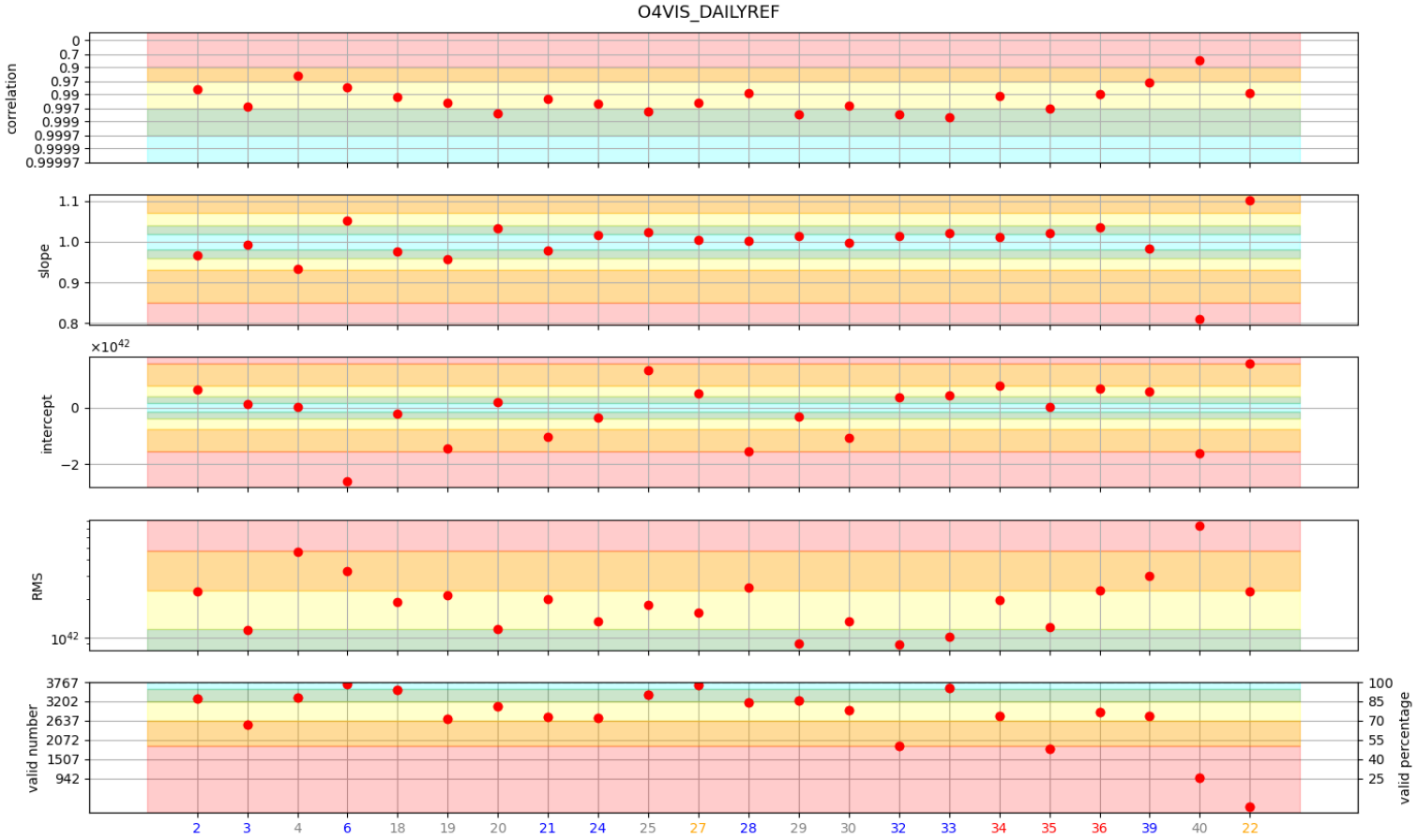

Figure 38: Same as Fig. 34 but for O4VIS_DAILYREF

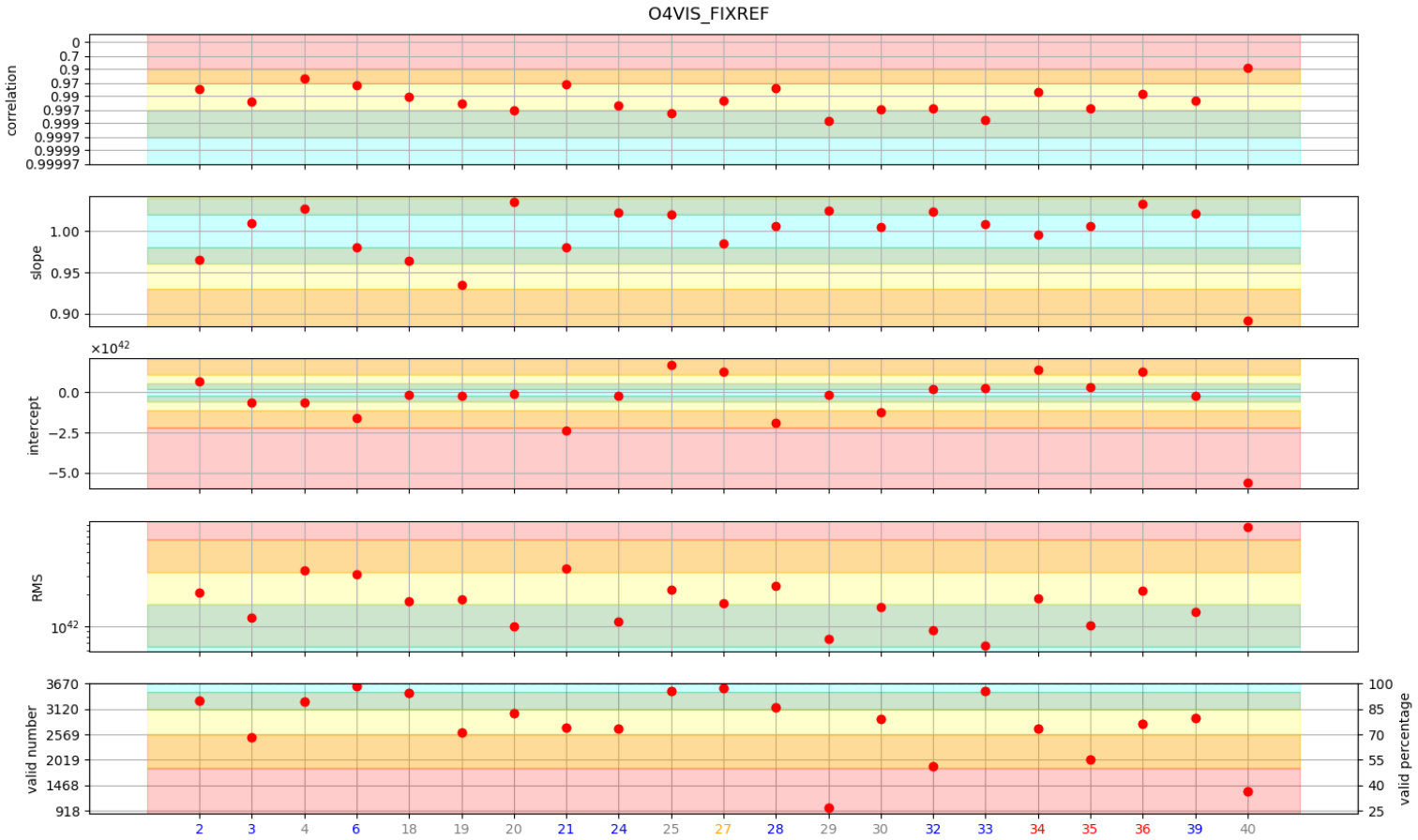

Figure 39: Same as Fig. 34 but for O4VIS_FIXREF

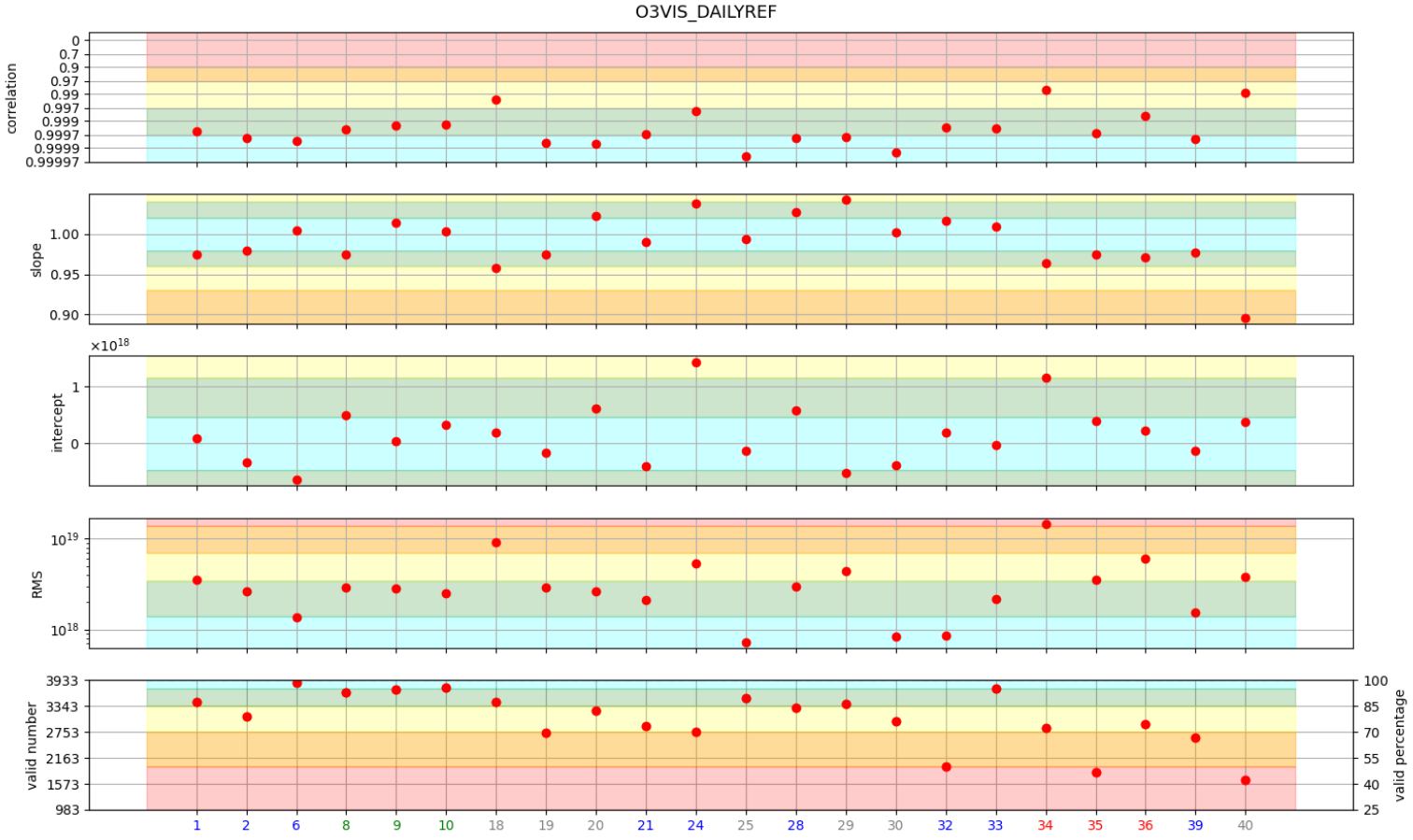

Figure 40: Same as Fig. 34 but for O3VIS_DAILYREF

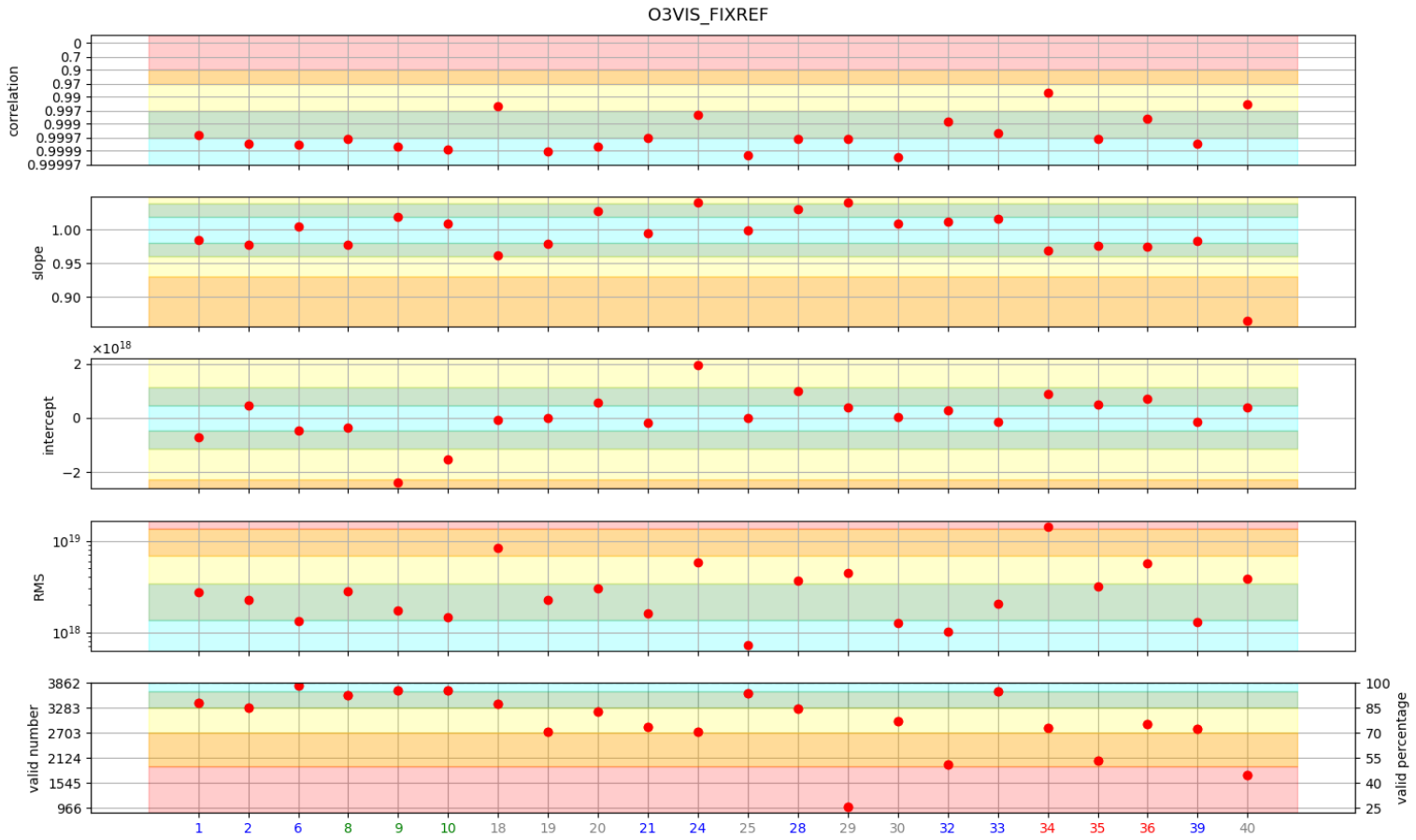

Figure 41: Same as Fig. 34 but for O3VIS_FIXREF

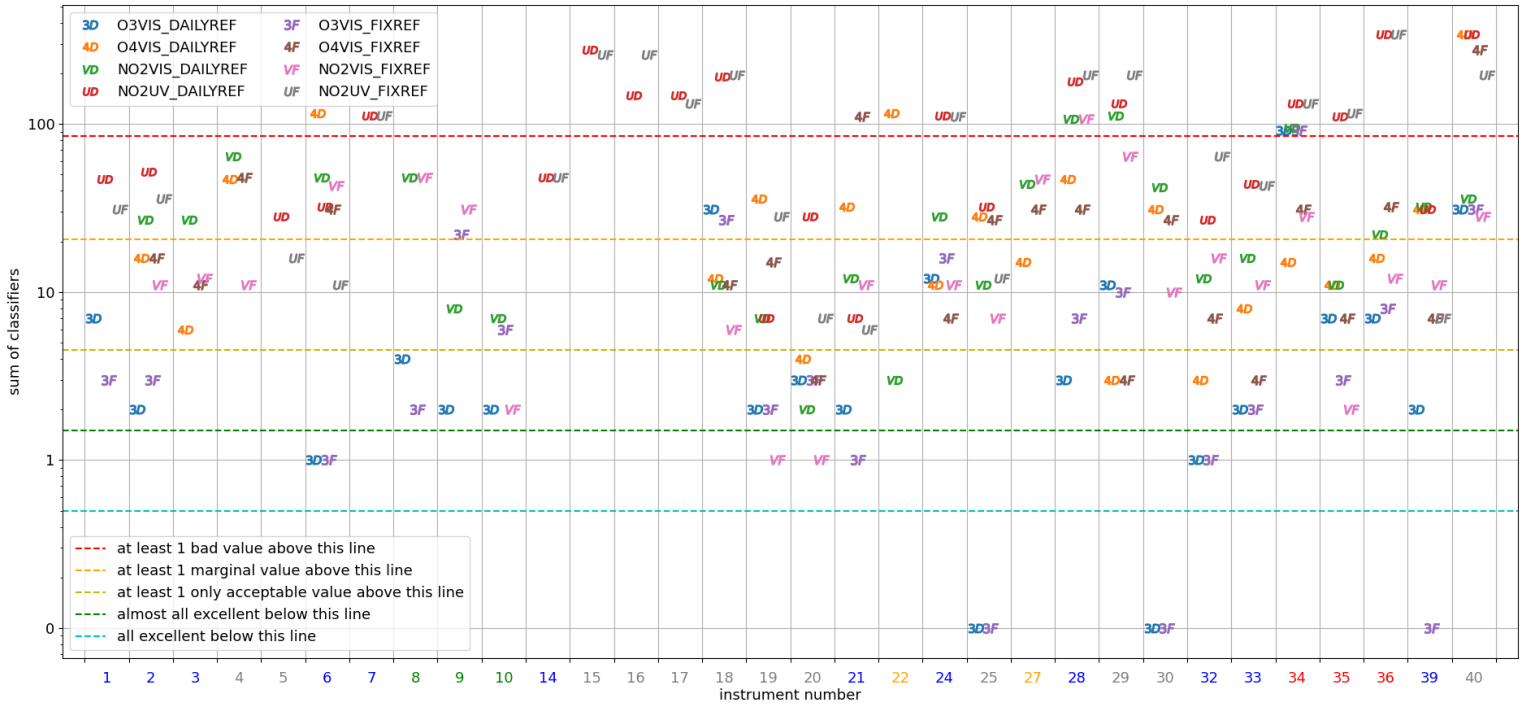

Figure 42: Cummulative scores (y-axis) for the different instruments (x-axis) for the different product-reference type combinations as indicated in the legend in the upper left. The horizontal lines indicate the score zones (see legend lower left). For details, see text.

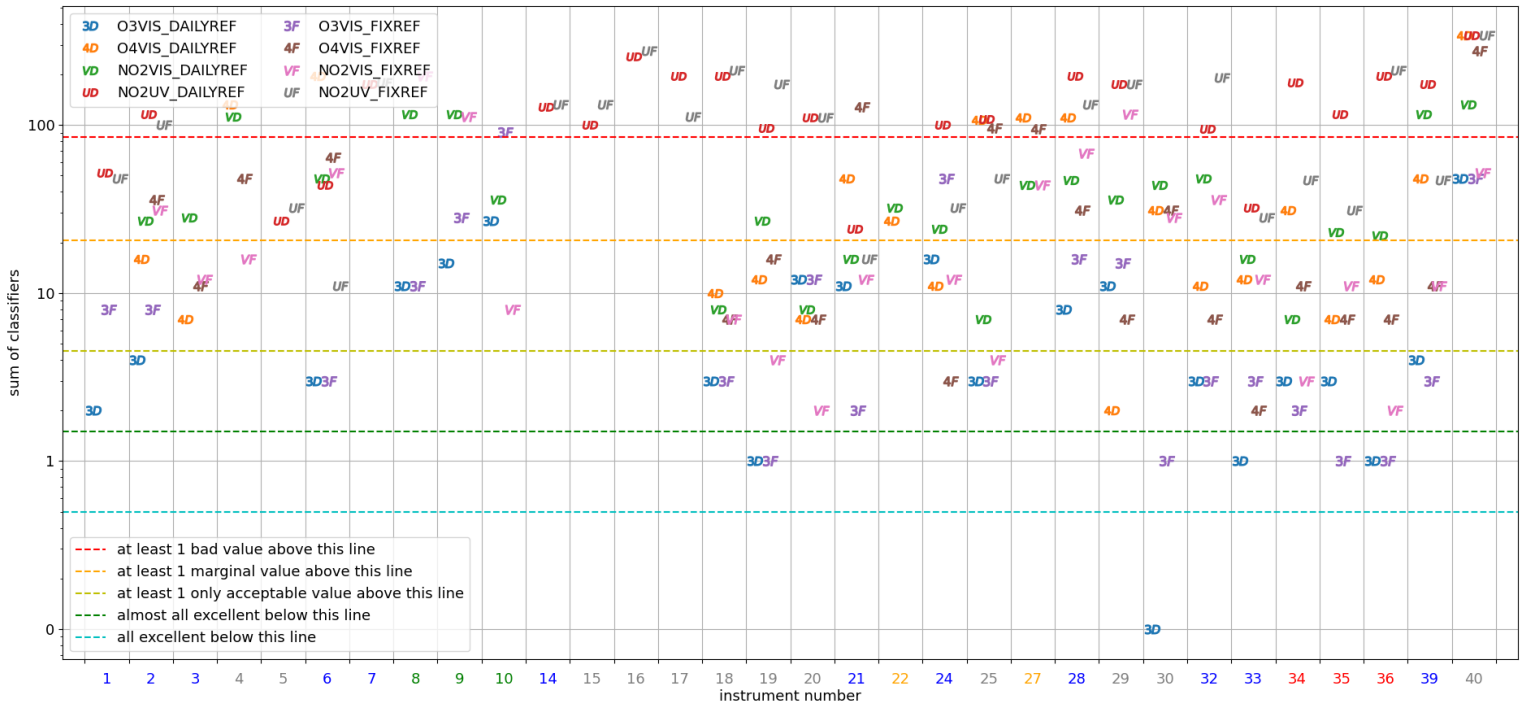

Figure 43: Same as Fig. 42 but for the SZA region [50, 85]

6 Elevation scan analysis

The inter-comparison exercise for the elevation scans included 10 products (c.f. Table 3) with 3 different reference types, FIXREF, DAILYREF, and SEQREF, as outlined in Subsect. 2.1.

Subsect. 6.1 outlines the creation of the reference dataset. Before presenting the results of the inter-comparisons, we show an overview of the time compliance in both measurement channels, UV and visible for the different instruments in Subsect. 6.2. Subsect. 6.3 presents the results from the elevation scan inter-comparison analysis and Subsect. 6.4 contains some concluding remarks.

6.1 Reference dataset construction

As opposed to the reference dataset for the zenith measurements (c.f. Sect. 5) where interpolations on a fixed SZA grid throughout the measurement period were performed, the reference dataset for the VEA scan measurements matches individual measurements. To this extend, zenith measurements were identified in a window ±60 s around the nominal measurement time slot while off-axis measurement were identified in a window ± 2 min around the nominal measurement time slot. The narrower time window of ±60 s for the zenith measurements was used to avoid averaging across multiple measurements.

Since the official measurement period comprised 18 days (2 June – 19 June) with 15 scans per day, each off-axis angle has up to 522 possible measurements, while each zenith angle (recorded twice per scan) has up to 1044 measurements. In Fig. 44 we show a short time series of 90° measurements for 5 instruments, three of which took continuous 90° measurements instead of elevation angle scans, however at a frequency lower than 1/ min (see also information from the questionnaire in Appendix D for those instruments). Note the general good agreement. It is visible that the instruments that measure continuously at 90° capture the fast changing NO_2 above the station. If a window of ±2 min is used for the identification of zenith measurements during elevation scans, instruments that continuously measure at 90° would be penalized. In order to avoid this, a stricter time criterion was used for 90° measurements (±60 s) compared to other elevation angles.

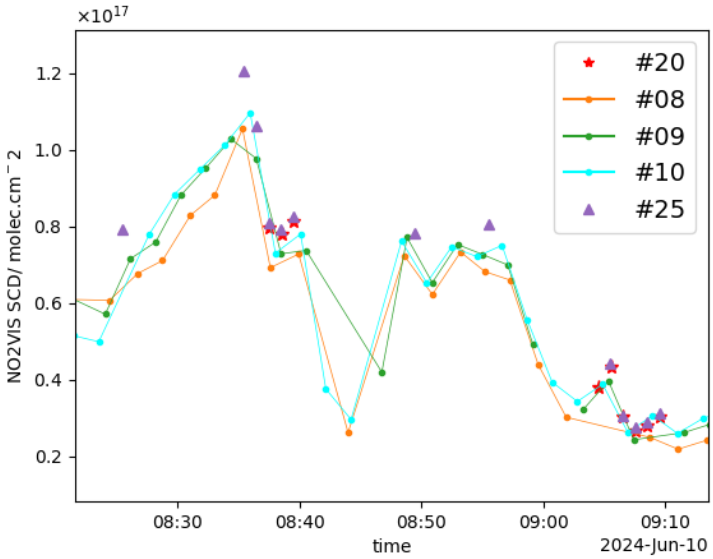

Figure 44: Example 90° NO2VIS_FIXREF time series during morning hours of 10 June for 5 instruments demonstrating the need to limit 90° measurement identification to ±60 s. Note the dip in DSCD during the scan 08:39 and 08:49 UTC

Only instruments that contribute to at least 40% to the measurements (separately considered at each elevation angle) are considered for the mean and median calculation. Further, only times at which at least 70% of these instruments were delivering data were considered in the reference dataset creation. Two baseline datasets are created: one using the median and another one, similar to the investigations for the zenith measurements, using an inverse fitting RMS weighted mean. Unless otherwise stated, the median constitutes baseline dataset.

In Fig. 45, an overview of the coverage (in terms of theoretically reachable 522 measurements per elevation angle per instrument per product per reference type) is given. Since the same requirements for both baseline datasets apply, the coverage is the same and hence only the median is included as an additional "instrument" on the y-axis. Note that instrument #22 joined the campaign late, hence there are much fewer measurements.

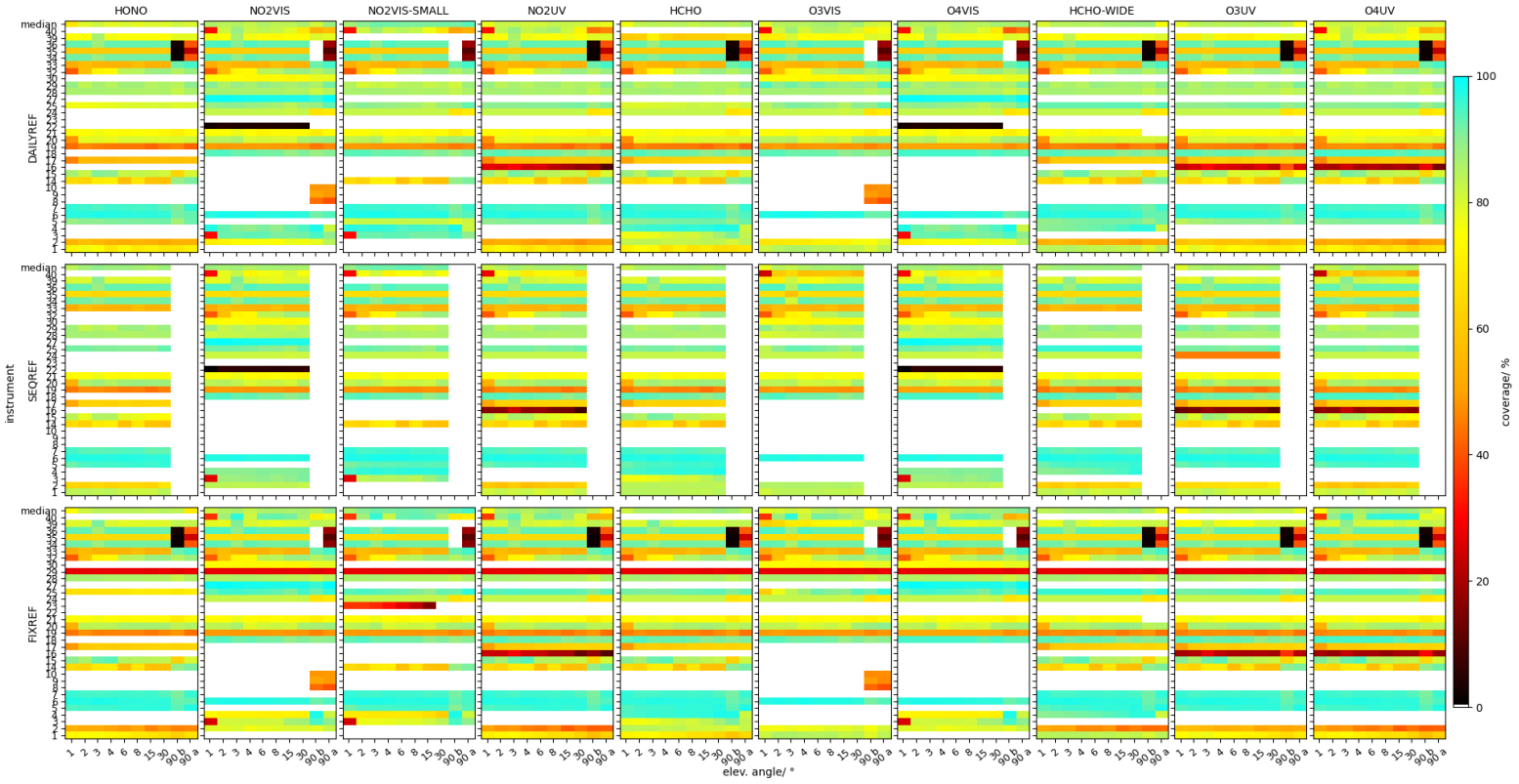

Figure 45: Data coverage for all product-reference type combinations for all instruments and all positive elevation angles. Each product is displayed in a separate column and each reference type is in a separate row, as indicated.

It was decided to exclude the measurement at the negative elevation angle (-2°) from the general inter-comparison. Hence, it is not further considered. However, since some instruments have very different time matches for the zenith measurement before and after an elevation scan, they are considered separately, therefore, the total number of considered elevation angles remains 10: 1°, 2°, 3°, 4°, 6°, 8°, 15°, 30°, 90° before (90 b) and 90° after (90 a).

6.2 Time compliance

Using reference type DAILYREF and considering both NO2VIS and NO2UV, the median time deviation (from the nominal measurement time for each elevation angle) together with the 10th and 90th percentiles are shown in Fig. 46 and Fig. 47 (upper panels), respectively. The lower panels in both figures display the coverage for each elevation angle and instrument. With the exception of this overview, zenith measurements are treated separately as zenith before (90 b) and zenith after (90 a).

As shown in Fig. 46 and Fig. 47, for most instruments, the median of recorded measurement times for all angles are within ±60 s of the nominal time. An exception is Instrument #40, for which the lower elevation angles were often measured about 80 s before the nominal time. This time mismatch could potentially explain the lower agreement of the measurements from this instrument with the baseline that is frequently observed in the following sections.

For the majority of instruments, most measurements (the range between the 10th to 90th percentile) also show a small spread of deviation in time offset from the nominal time. Notable exceptions include the three Pandora instruments (#34, #35, and #36), which exhibited a broader spread of deviations from the nominal measurement times and the three SAOZ instruments that had longer integration times (c.f. Appendix D). This also lead to fewer measurements in the applied time window around the nominal measurement times.

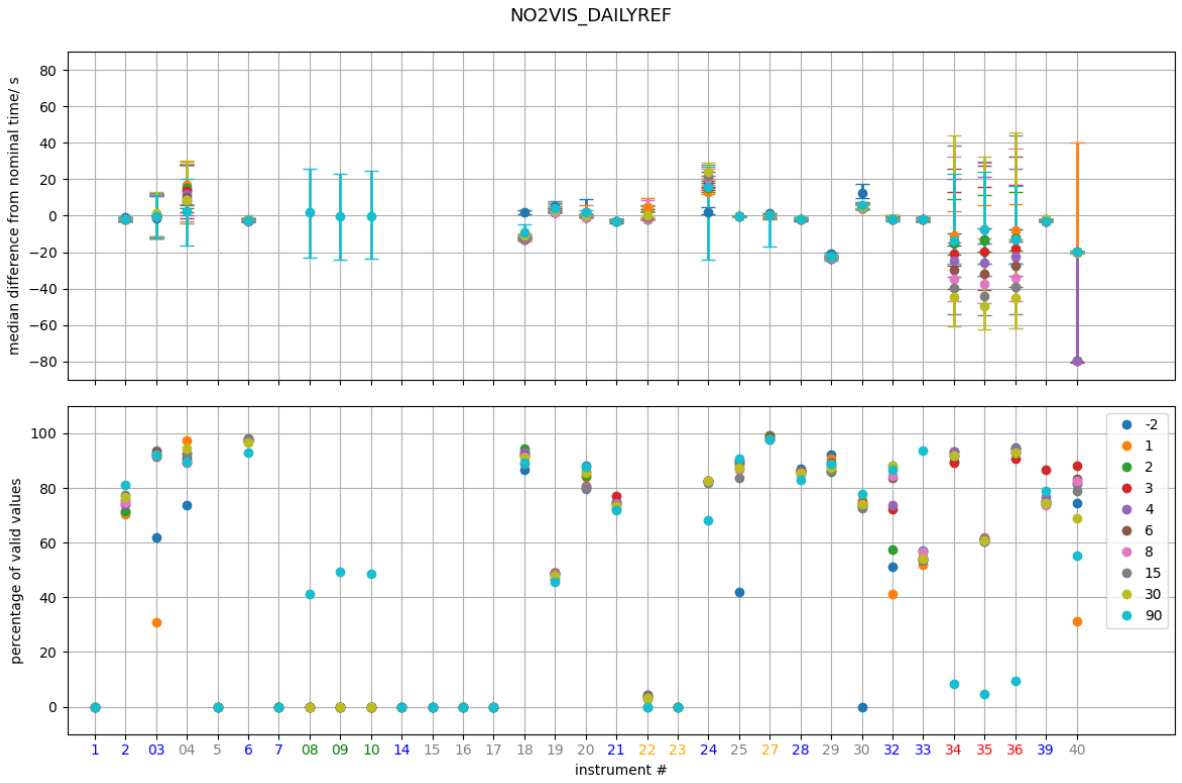

Figure 46: Upper panel: Deviations from nominal measurement times for each elevation angle (as indicated in the legend) for each instrument (on x-axis), reported measurement times as reported in files for product NO2VIS_DAILYREF. Lower panel: Measurement coverage for each elevation angle for each instrument for NO2VIS_DAILYREF.

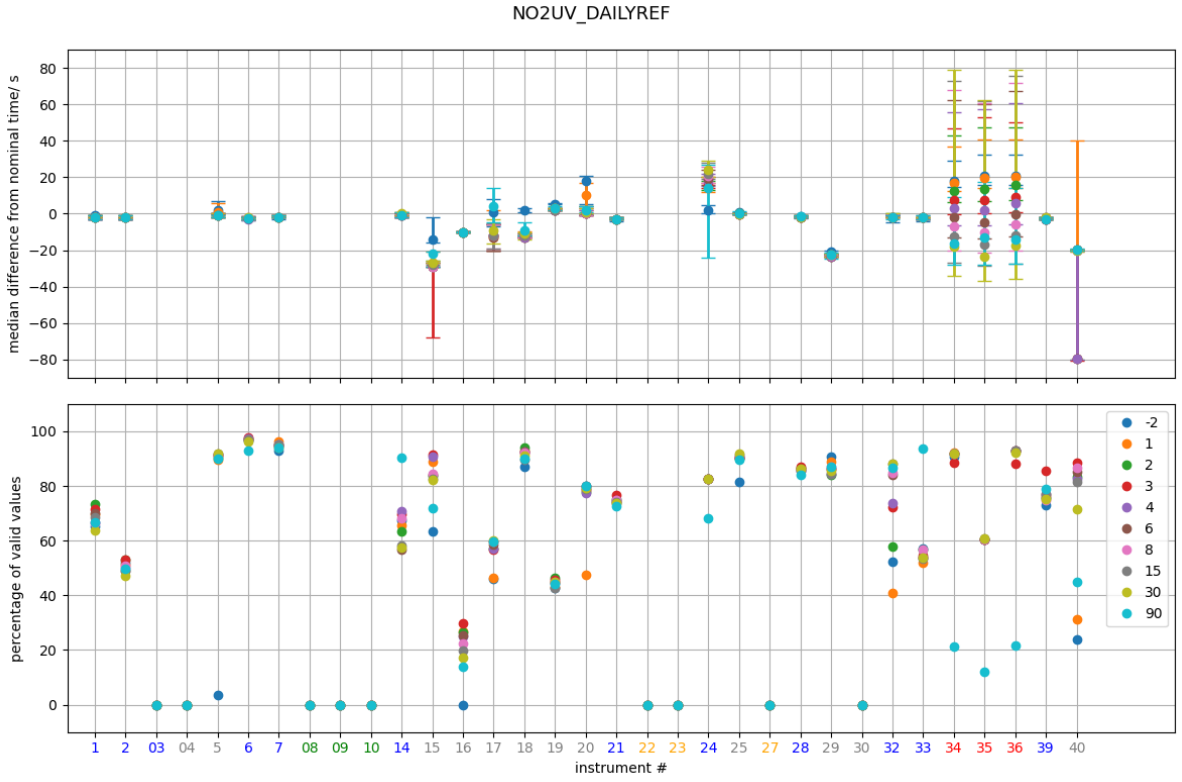

Figure 47: As Fig. 46 but for the UV channel, using measurement times as reported in NO2UV_DAIYREF.

6.3 Inter-comparison of SCD at different elevation angles

In Fig. 48 we show the time series of NO2VIS_FIXREF for instrument #6 as an example for the NO_2 DSCD evolution at three different elevation angles, 15° (red), 30° (black) and 90° (cyan, including both the zenith before and after each scan). This instrument has been chosen over the median for demonstration purpose due to its very high data coverage.

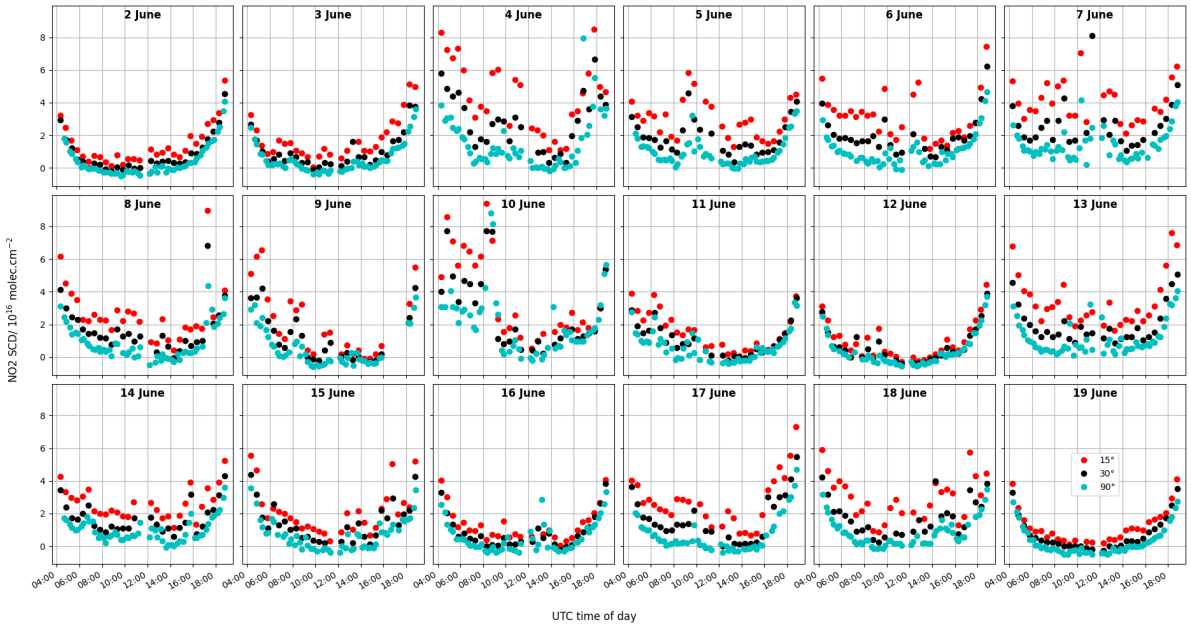

Figure 48: Time series for NO2VIS_FIXREF for instrument #6

To investigate the agreement of DSCD for off-axis measurements, we consider

1. The correlation R with the baseline dataset
2. The RMS of the difference between the individual dataset and the baseline dataset
3. The slope s and
4. The intercept I of the Theil–Sen (c.f.Theil, 1950 and Sen, 1968) fit considering the individual data as the dependent and the baseline dataset as the independent variable.

The RMS and the intercept I are not evaluated directly, but as fractions of the median baseline values. For the 30 product-reference type combinations at the 10 different elevation angles, the medians of the baseline datasets are summarized in Table 6, Table 7 and Table 8.

As an example we show in Fig. 49 regression scatter plots for individual instruments for 1°, 4° and 15° for NO2VIS-SMALL_FIXREF. All regression scatter plots can be found in the Appendix A.

VEA/°	1	2	3	4	6	8	15	30	90 b	90 a
NO2VIS/ (10^{16} molec.cm^{-2})	7.1	6.2	5.4	4.7	4.0	3.4	2.3	1.4	0.7	0.8
NO2VIS-SMALL/ (10^{16} molec.cm^{-2})	7.1	6.2	5.3	4.7	3.9	3.3	2.2	1.4	0.7	0.8
NO2UV/ (10^{16} molec.cm^{-2})	4.9	4.5	4.1	3.7	3.2	2.7	1.8	1.1	0.5	0.5
HCHO/ (10^{15} molec.cm^{-2})	10.6	10.1	9.6	8.9	7.2	6.2	2.5	-1.7	-5.0	-5.3
O3VIS/ (10^{19} molec.cm^{-2})	1.4	1.4	1.4	1.4	1.4	1.4	1.3	1.2	1.0	1.2
O4VIS/ (10^{43} molec^2.cm^{-5})	5.4	5.2	4.9	4.5	4.1	3.7	2.8	1.9	1.1	1.2
HCHO-WIDE/ (10^{15} molec.cm^{-2})	14.2	14.0	13.3	12.8	11.4	10.1	7.1	3.2	0.2	0.2
O3UV/ (10^{19} molec.cm^{-2})	1.0	1.0	1.0	1.1	1.1	1.1	1.1	1.0	0.9	1.0
O4UV/ (10^{43} molec^2.cm^{-5})	3.8	3.8	3.7	3.6	3.4	3.2	2.6	1.8	1.0	1.0
HONO/ (10^{13} molec.cm^{-2})	-4.0	-12.2	-16.2	-18.2	-18.8	-19.5	-17.2	-13.6	-19.0	-16.1

Table 6: Median baseline values for FIXREF.

Similar to the classification for the zenith measurements, see Sect. 5, classification bounds for the 4 quality indicators listed above have been introduced as follows:

1. R: [0 – 0.5] bad, (0.5 – 0.9] marginal, (0.9 – 0.99] good, (0.99, 1.0] excellent.
2. RMS: > 1: bad, [1, 0.6) marginal, [0.6, 0.1) good, [0.1, 0] excellent.

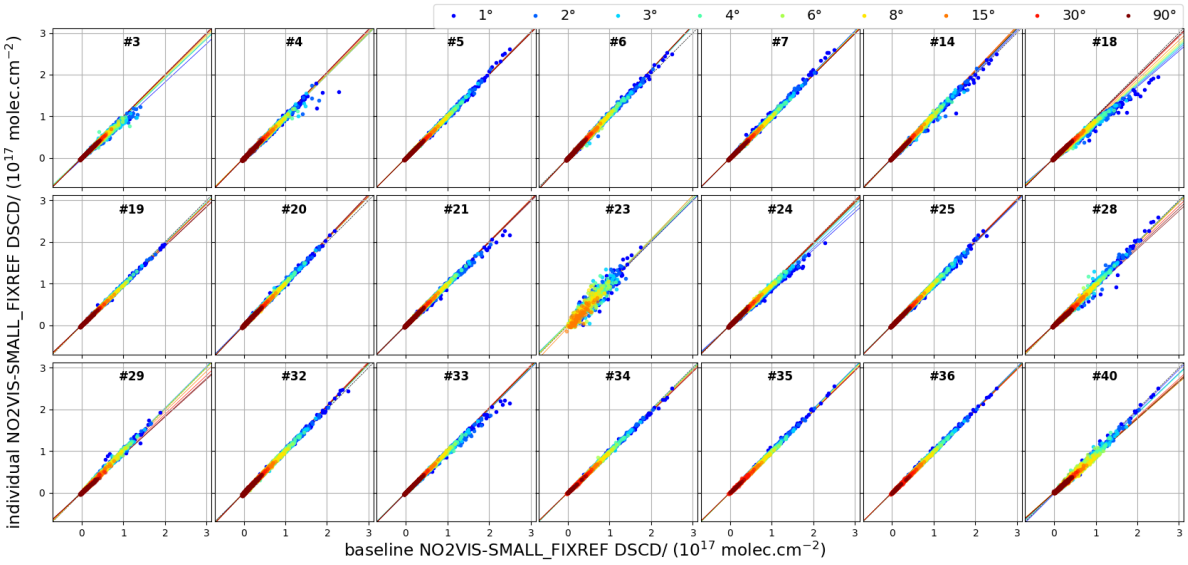

Figure 49: Example of a scatter plot of individual NO2VIS-SMALL_FIXREF DSCDs against the baseline. Each panel is a different instrument as indicated in the panel; Color represents different VEA, as indicated. The Theil-Sen fit for each VEA dataset is also included in each panel with the same color coding. Additionally, the 1:1 line is indicated with back dots.

VEA/°	1	2	3	4	6	8	15	30	90 b	90 a
NO2VIS/ (10^{16} molec.cm^{-2})	6.8	5.7	5.0	4.2	3.5	2.8	1.7	0.9	0.2	0.2
NO2VIS-SMALL/ (10^{16} molec.cm^{-2})	6.6	5.6	4.8	4.2	3.4	2.8	1.7	0.9	0.2	0.3
NO2UV/ (10^{16} molec.cm^{-2})	4.5	4.0	3.6	3.2	2.7	2.3	1.4	0.7	0.2	0.2
HCHO/ (10^{15} molec.cm^{-2})	12.0	12.4	11.7	10.7	9.6	8.1	4.4	0.2	-4.2	-4.4
O3VIS/ (10^{19} molec.cm^{-2})	1.3	1.4	1.4	1.4	1.4	1.4	1.3	1.2	1.0	1.2
O4VIS/ (10^{43} molec^2.cm^{-5})	4.9	4.4	4.2	3.8	3.4	3.0	2.1	1.3	0.5	0.5
HCHO-WIDE/ (10^{15} molec.cm^{-2})	15.0	14.9	14.2	13.6	12.2	11.0	7.5	3.6	0.1	-0.1
O3UV/ (10^{19} molec.cm^{-2})	1.1	1.0	1.0	1.1	1.0	1.0	1.1	1.1	0.9	0.9
O4UV/ (10^{43} molec^2.cm^{-5})	3.2	3.2	3.0	2.9	2.7	2.5	2.0	1.2	0.5	0.5
HONO/ (10^{13} molec.cm^{-2})	10.3	0.6	-3.0	-6.5	-5.1	-8.1	-4.4	-1.3	-7.6	-6.2

Table 7: Median baseline values for DAILYREF.

3. s: Due to the larger deviation from unity, deviation towards smaller values from unity is reciprocal symmetric: $s < 0.5$ or $s > 2$ bad, $0.5 <= s < 1/1.6$ or $2 <= s < 1.6$ marginal, $1/1.6 <= s < 0.8$ or $1.6 <= s < 1/0.8$ acceptable, $0.8 <= s < 0.98$ or $1/0.8 <= s < 1/0.98$ good, $0.98 <= s <= 1/0.98$ excellent.

4. I: Deviations from 0 larger than 0.6 are bad, deviations [0.6, 0.4) are marginal, deviation[0.4, 0.2) are ok, deviations [0.2, 0.02) are good and deviations smaller or equal 0.02 are excellent.

These limits are represented by the color scales used in Fig. 50 to Fig. 72.

VEA/°	1	2	3	4	6	8	15	30
NO2VIS/ (10^{16} molec.cm^{-2})	6.7	5.6	4.8	4.1	3.3	2.7	1.5	0.6
NO2VIS-SMALL/ (10^{16} molec.cm^{-2})	6.4	5.4	4.6	3.9	3.1	2.6	1.5	0.6
NO2UV/ (10^{16} molec.cm^{-2})	4.4	4.1	3.6	3.2	2.7	2.2	1.3	0.6
HCHO/ (10^{15} molec.cm^{-2})	16.5	16.4	16.0	15.2	13.8	12.4	8.6	4.4
O3VIS/ (10^{19} molec.cm^{-2})	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.1
O4VIS/ (10^{43} molec^2.cm^{-5})	4.4	4.1	3.7	3.4	3.0	2.5	1.7	0.8
HCHO-WIDE/ (10^{15} molec.cm^{-2})	15.5	15.3	14.7	13.9	12.7	11.3	8.0	4.0
O3UV/ (10^{19} molec.cm^{-2})	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
O4UV/ (10^{43} molec^2.cm^{-5})	2.8	2.8	2.7	2.6	2.4	2.1	1.6	0.8
HONO/ (10^{13} molec.cm^{-2})	13.3	6.3	1.2	-0.9	-1.2	-2.3	-0.9	2.8

Table 8: Median baseline values for SEQREF.

6.3.1 NO2VIS

In Fig. 50 we show the regression analysis results for all reference types for product NO2VIS. The corresponding scatter plots can be found in Subsect. A.1. As is apparent from the upper left panel of Fig. 50, all but the lower VEAs measurements (1°, 2° and 3°) of instrument #39 (all reference types) and the higher (15° and 30°) VEAs measurements of instrument #27 (SEQREF only) have slope values in the good region and most even in the excellent region.

The deviation for instrument #27 is seen more pronounced (also for other angles) for other products and is likely related to the pixel-angle mapping (#27 is an instrument with imaging capability). According to Table 2, the angular resolution for the higher elevation angles is rather coarse: 7.5° for 15° elevation, 12° for 30° elevation and 35° for 90° elevation. Hence, a larger deviation at larger VEA is to be expected if the NO2 field is inhomogeneous.

A lower slope value (than 1) for #39 means that the DSCDs of #39 are below the baseline. While this could also be explained by a positive offset of VEA (e.g. measurement at nominal 1° VEA is really at ¿1° VEA), the horizon scan analysis (c.f. Fig. 19) seem to indicate that no such offset is present for #39. A similar underestimation was seen in an early product version of instrument #21 before a stray-light correction was applied. Stray-light leads to a an offset in both the spectra intensity I in question and in the reference spectrum intensity I_0 and can hence lead to a multiplicative bias. According to the information from the PI of #39 (see Appendix D) no stray-light correction was applied for #39. Hence, a possible reason for the underestimation might be related to stray-light.

A similar picture is seen for the intercept, see lower left panel in Fig. 50. Also here, instruments #27 (all reference types for 30° VEA) and #39 (only DAILYREF, only 90° VEA) fall outside the good intercept range.

Considering the correlation R, most instruments are in the excellent region for reference types DAILYREF and FIXREF. Exceptions here are instrument #3 for low VEA, #27 for high (but not zenith) VEA (15° and 30°) #28 for 1° and #39 and #40 for most elevation angles. Possible reasons are:

- #3 was very unstable in terms of pointing accuracy and the 1° angle was only submitted on those days where the horizon scan indicated a fairly good pointing. However, given the overall instability (c.f. Fig. 19), it is likely that also on those days, the 1° direction had a larger pointing inaccuracy than other instruments. The 2° angle constitutes most often an interpolation, accounting for the horizon offset from the horizon scans.

- The slightly worse performance of #27 is likely again due to the wider field of view for those elevation angles.

- Given the relatively good pointing accuracy of #28 (c.f. Fig. 19) it is unlikely that the slightly worse correlation for this instrument at 1° VEA can be attributed to the pointing. Considering the scatter plot (c.f. Fig. 96) the worse correlation is rather due to a limited number of 1° measurements that show large deviations from the baseline.

- The systematic underestimation of DSCD of #39 already noted earlier does also affect the RMS. However, it does not explain the worse correlation. The correlation for SEQREF and DAILYREF is of equal quality while the correlation for FIXREF is significantly better, especially for higher VEAs. Given the good time compliance (c.f. Fig. 46) a similar explanation as for #40 (see next point) is ruled out.

- #40: Comparing DAILYREF to FIXREF in Fig. 50 it is evident that the performance in terms of R and slope s is significantly better for FIXREF than for DAILYREF. While instructions were very clear to use the mean of the first 9 measurements of day 6 June as a reference for FIXREF, instructions, specifically w.r.t. filtering were less clear for the DAILYREF. Hence, one explanation might be that the PI of #40 used a different approach for the DAILYREF reference construction. However, considering that also the SEQREF RMS and R are among the worse ones, this explanation is less likely. However, another explanation might be the relatively large consistent offset in zenith measurement time (c.f. Fig. 46) which might lead to slightly different reference each day (and each scan) and therefore affect the correlation in DAILYREF and SEQREF.

In general it can be seen that the RMS in units of the median baseline DSCD for zenith measurements is consistently higher for all instruments for DAILYREF compared to FIXREF. However, this is explained by the smaller range (and median) for DAILYREF for zenith measurements compared to FIXREF (c.f.Table 7 and Table 6).

We can investigate the instrument internal consistence between the different reference types by comparing

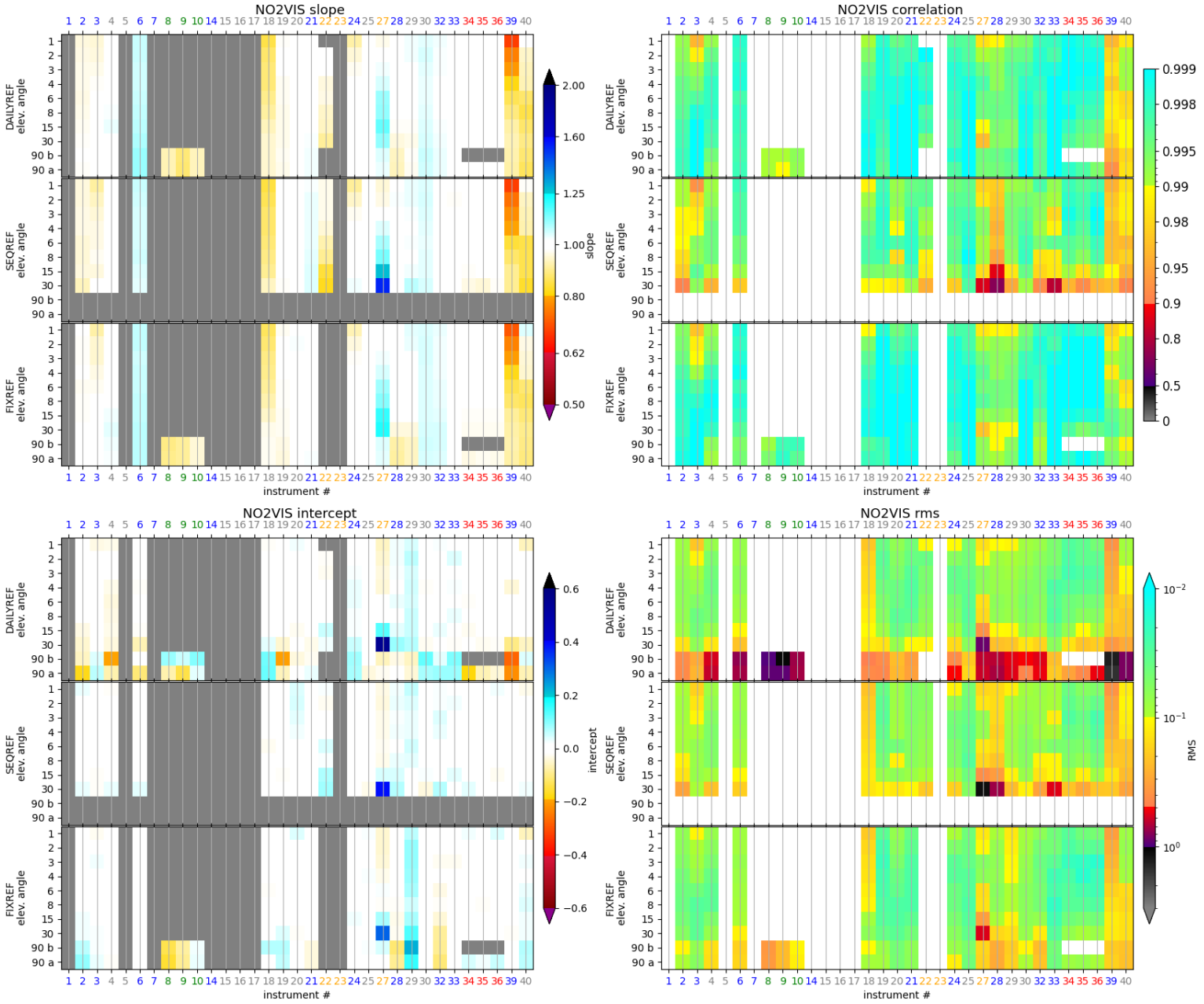

Figure 50: Theil-Sen slope (upper left), intercept (lower left), correlation (upper right) and RMS between the difference of individual instruments and the median of NO2VIS; each panel top-down for DAILYREF, SEQREF and FIXREF for elevation angles as indicated. 90 a denotes the zenith measurement after the scan and 90 b denotes the zenith measurement before the scan. Note that the intercept and the RMS are in units of the median of the baseline. Gray in the left panels (diverging color maps) and white in the right panels, denote missing values.

$$\Delta DSCD(\alpha) = DSCD(\alpha) - DSCD(90_a) \tag{4}$$

both for FIXREF and DAILYREF, to the corresponding DSCD of SEQREF. In Fig. 51 we compare Δ DSCD$_{DAILY}$ to DSCD$_{SEQ}$ and in Fig. 52 we compare Δ DSCD$_{FIX}$ to DSCD$_{SEQ}$. For most instruments (#6, #18, #19, #20, #24, #25, #30, #34, #35, #36 and #40), the difference is non-systematic and small. For instruments #2, #4, #27, #28, #29 and #32, the difference is larger and non-systematic. For half of those (#28, #29 and #30), the difference does not change much considering either DAILYREF or FIXREF (compare Fig. 51 to Fig. 52 and consider the overall small differences in Fig. 53), indicating that the changes in instrument are fairly similar throughout the day compared to throughout the whole campaign.

For three instruments (#3, #21 and #39), the difference is proportional to the DSCD itself. For the first of these, the DSCD using the reference spectrum after the zenith is smaller than the Δ DSCD$_{DAILY}$. This could be caused by saturation of the noon reference spectrum such that the I_0 is underestimated compared to the I at different VEA leading to too large DSCD for DAILYREF

and FIXREF compared to SEQREF. It is less clear what an overestimation of SEQREF (as for #21 and #39) could cause. The PI of #21 notes that there were slight unintended differences in the wavelength calibration applied to the SEQREF analysis compared to the FIXREF and DAILYREF. This could cause small systematic bias.

There are also structural differences between Δ $\mathrm{DSCD_{DAILY}}$ and $\Delta DSCD_{\mathrm{FIX}}$ for instruments #27 and #40, the former showing for some of the 30° VEA having a constant offset, the latter showing for some measurements a linear dependence on Δ $\mathrm{DSCD_{DAILY}}$. The latter might indicate an instrumental change during the measurement period that acts multiplicatively on the DSCD.

Lastly, #33 shows some constant offset, most likely indicating that not all measurements for SEQREF used the 90° reference after the scan.

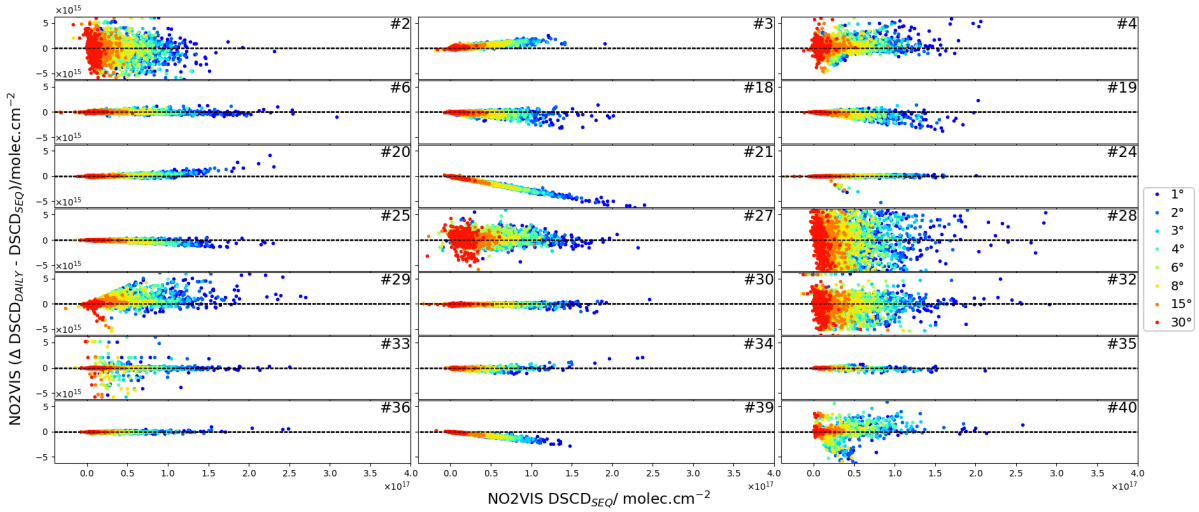

Figure 51: $\Delta$$\mathrm{DSCD_{DAILY}}$ - $\mathrm{DSCD_{SEQ}}$ against $\mathrm{DSCD_{SEQ}}$, color coded is the VEA.

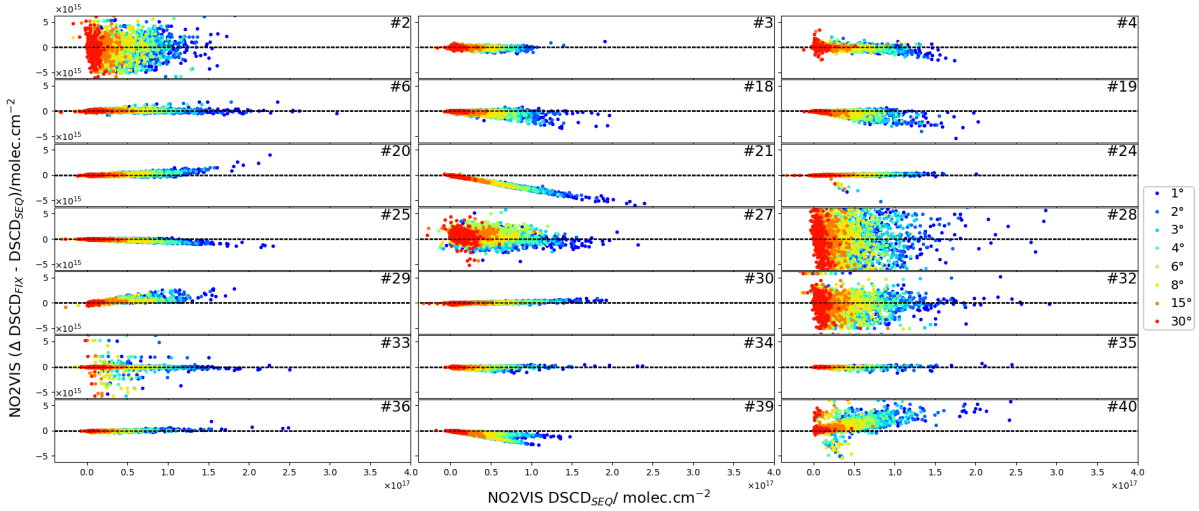

Figure 52: Δ $\mathrm{DSCD_{FIX}}$ - $\mathrm{DSCD_{SEQ}}$ against $\mathrm{DSCD_{SEQ}}$, color coded is the VEA.

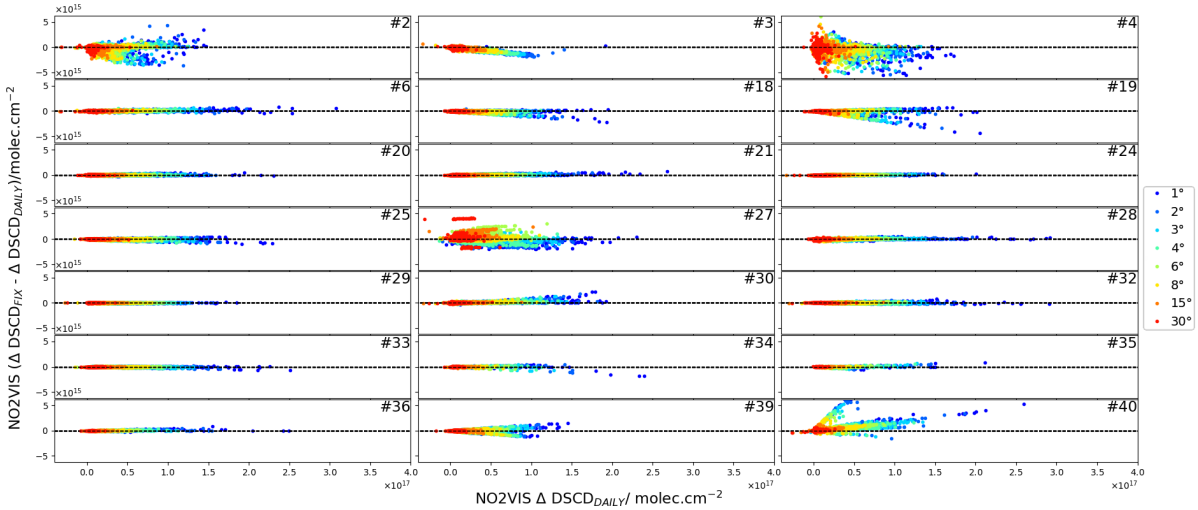

Figure 53: Δ DSCD$_{\text{FIX}}$ - Δ DSCD$_{\text{DAILY}}$ against Δ DSCD$_{\text{DAILY}}$, color coded is the VEA.

6.3.2 NO2VIS-SMALL

The NO2VIS-SMALL was introduced as a product to offer a visible NO_2 product also for those instruments that do not cover the whole wavelength region of the NO2VIS window (c.f. Table 3). Of the 8 instruments that did not cover the NO2VIS wavelength range, 4 provided the NO2VIS-SMALL product: #5, #7, #14 and #23 (FIXREF only, no VEA > 15°). In Fig. 54 we show a comparison of the two products for DAILYREF. It is evident that the correlation between the two baseline datasets (medians) is excellent and is larger than 0.998 for all VEA and larger than 0.999 for VEAs >1°.

Figure 54: Time series (left panels) and regression plots of baseline datasets of NO_2 products NO2VIS and NO2VIS-SMALL for reference type DAILYREF.

In Fig. 55 we show the same type of regression analysis results as for NO2VIS in the last section. It can be seen, that in terms of slope, all instruments are within the good range and most are excellent. For the intercept I, this is also largely true. Except for instrument #3 at small elevation angles (the reason has already been discussed in Subsect. 6.3.2) and instrument #40.

As for NO2VIS, #40 shows better agreement with the baseline for FIXREF than for the other reference types, likely due to the different (shows of as an offset) influence of a slightly different reference due to the general time shift against the nominal measurement time of the zenith measurements. Overall, the performance of the instruments in the NO2VIS-SMALL window is very similar to the performance in the NO2VIS window. Instrument #23 that only delivered NO2VIS-SMALL_FIXREF at the lowest 7 VEAs has a slightly worse agreement with the baseline (visible in R, I and RMS) than most other instruments. However, as #22 and #27, #23 is an imaging instrument and scans through the wavelength space, taking approximately 1 image at a wavelength per second (c.f. Appendix D). With a resolution of 0.3nm it takes hence approximately 5 min to cover the wavelength range (434.5 – 455) nm. One image covers all the different provided low elevation angles. Hence, the time agreement is large compared to the time agreement of all other instruments. This explains the larger spread of the data around the baseline (c.f. Fig. 49). However, no systematic bias and only moderate offsets are detected (s is excellent and I good, see Fig. 50).

Some considerations for #32: While all three reference types for NO2VIS looked good in terms of correlation and slope, NO2VIS-SMALL_SEQREF seems worse. Considering the scatter plot in Fig. 100, it seems as if some measurements use a different reference, they show up on a parallel diagonal to the 1:1 line, suggesting that they were analysed with a spectrum containing a different amount of NO_2. However, in general, there is a larger spread compared to the NO2VIS product. Since this is not seen in either of the other two reference types DAILYREF and FIXREF, it still suggests an issue with the reference selection.

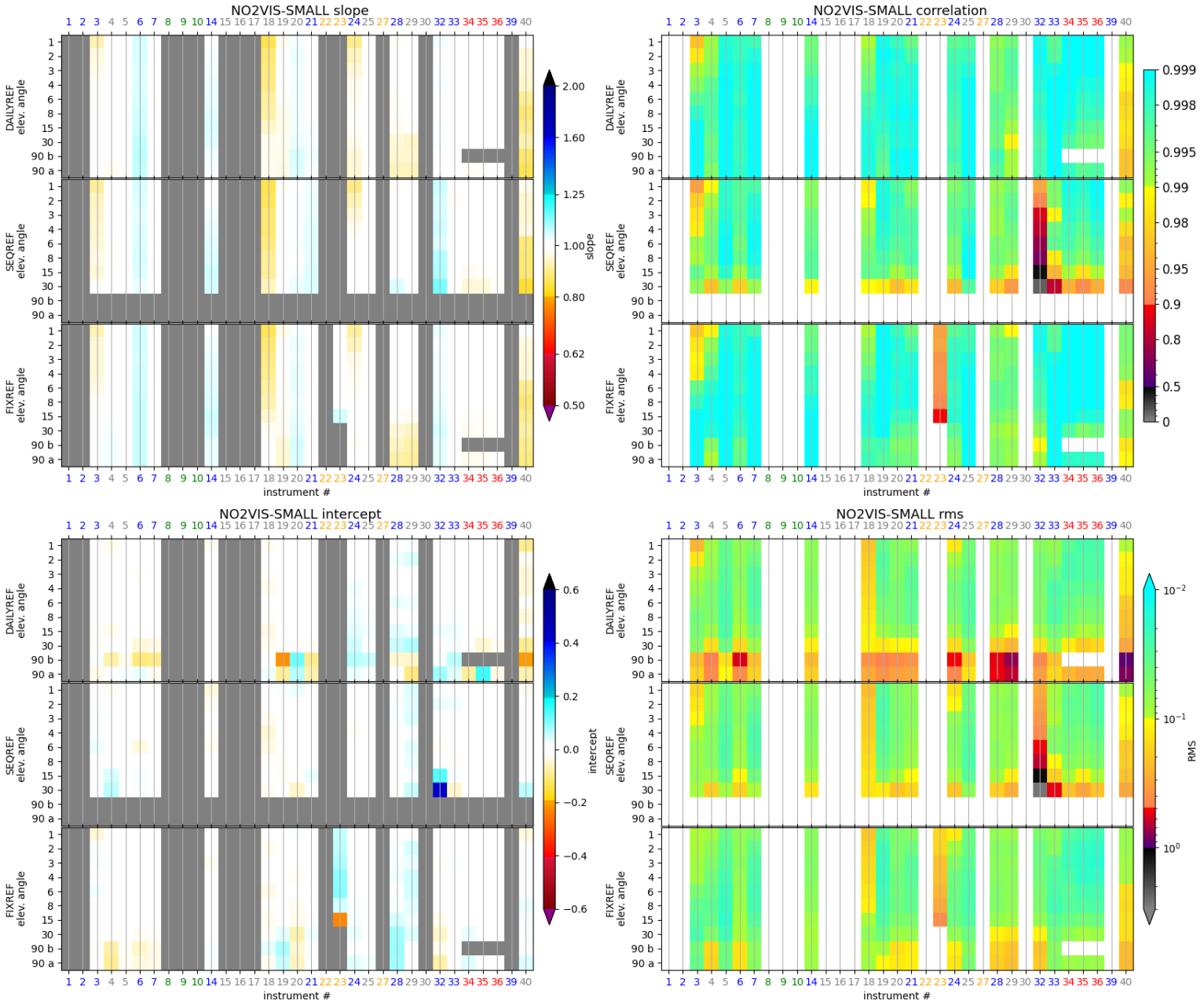

Figure 55: As Fig. 50 but for NO2VIS-SMALL

6.3.3 NO2UV

The air mass factor (AMF) in the UV wavelength region is considerably smaller than in the visible due to the stronger Rayleigh scattering in the UV. This reduced AMF is apparent when comparing the median DSCD values, c.f. Table 7 – Table 8. For lower elevation angles, this means that the area probed is confined closer to the instrument location. Since hence the spatial differences of measured line of sight (los) should be smaller, it could have been expected that the standard deviation between instruments that cover all NO_2 products is also smaller in the UV. However, this is not the case: There are 15 instruments that covered all three NO_2 products for reference type DAILYREF: #6, #18, #19, #20, #21, #24, #25, #28, #29, #32, #33, #34, #35, #36, and #40. In Fig. 56 we show the median (upper panel) and standard deviation at VEA 3° of the three products between those 15 instruments for all times at which at least 10 instruments delivered data. As previously noted, the DSCD of NO_2 is generally lower in the UV due to the smaller AMF. The standard deviation is most often comparable and sometimes, e.g. 7 June, considerably higher. One possibility might be that the weaker absorption strength of NO_2 in the UV and larger interference with other molecules such as HCHO outweigh the better spatial agreement. This can be tested by comparing the RMS in the different retrieval windows.

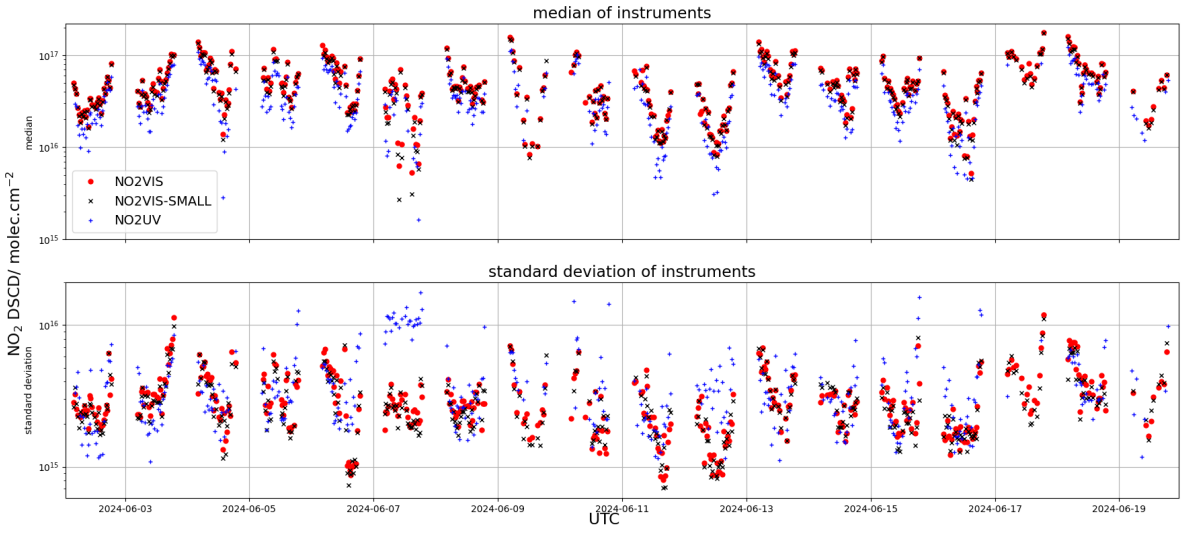

Figure 56: Median (upper panel) and standard deviation (lower panel) time series for NO_2 products with daily reference of 15 instruments that covered NO2UV, NO2VIS and NO2VIS-SMALL for reference type DAILYREF.

In Fig. 57 we show the distribution of the RMS for DAILYREF for all three NO_2 products for each instrument separately. Note that only measurements that were identified as a scan measurement are taken into account. Further, what is shown is the percentage of data, not absolute values in order to increase comparability among the different instruments and products. Of those 15 instruments that deliver all three NO_2 products, 4 instruments (#25, #34, #35 and #36) show a higher RMS in the UV than in the visible. For 7 instruments (#18, #19, #20, #21, #29, #32 and #33), the RMS seem comparable and for three instruments (#6, #24 and #28), the RMS in the UV is even smaller than in the visible. Hence it is more likely that, although the horizontal extend of the probed area is smaller (and hence in case of azimuthal misalignment the los agreement better), the increased scatter leads to light paths that differ (not necessarily in length) to a larger extend. If the NO_2 distribution varies greatly on very small spatial scales, this will increase the scatter in the UV among different instruments (that have unavoidably slightly different viewing directions) more than in the visible.

In Fig. 58 we show the regression analysis results for NO2UV. From the relatively (to the median baseline) larger standard deviation observed for UV as shown in Fig. 56, it is to be expected that also the general correlation is smaller and RMS is larger. This is confirmed: While the correlation and RMS for DAILYREF for NO2VIS-SMALL was for almost all instruments in the excellent region, most instruments are only in the good region for NO2UV.

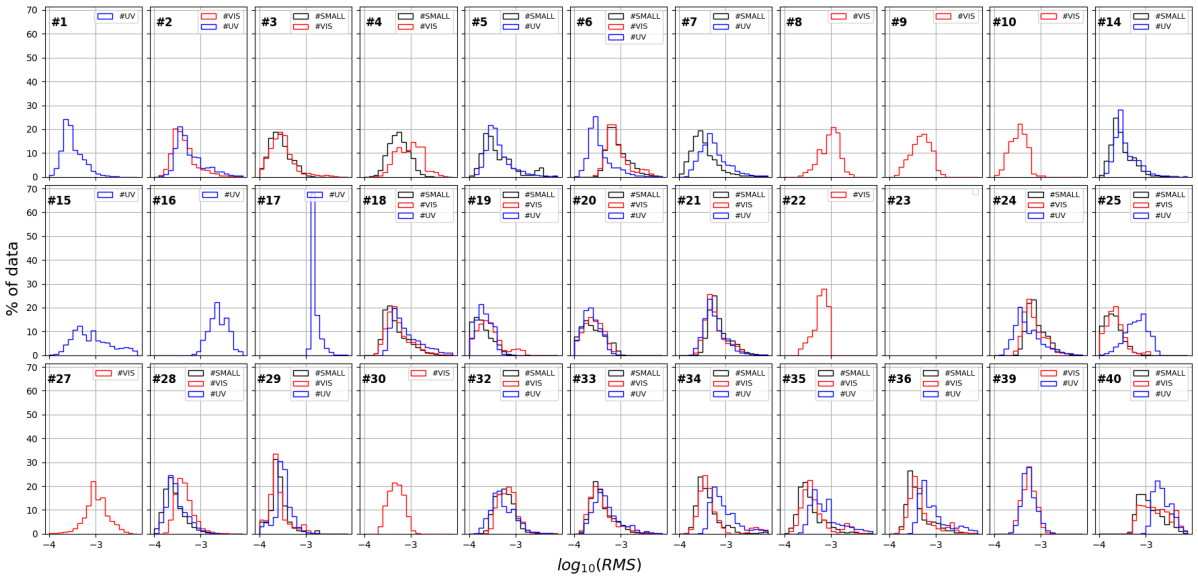

Figure 57: log_{10}(RMS) histograms for all three NO_2 products as indicated. SMALL refers to NO2VIS-SMALL, VIS refers to NO2VIS and UV refers to NO2UV. Each panel shows the distribution of RMS for a different instrument, as indicated. The histogram shows percentage of valid scan data, i.e. only those that match the scan times as outlined in Subsect. 6.2.

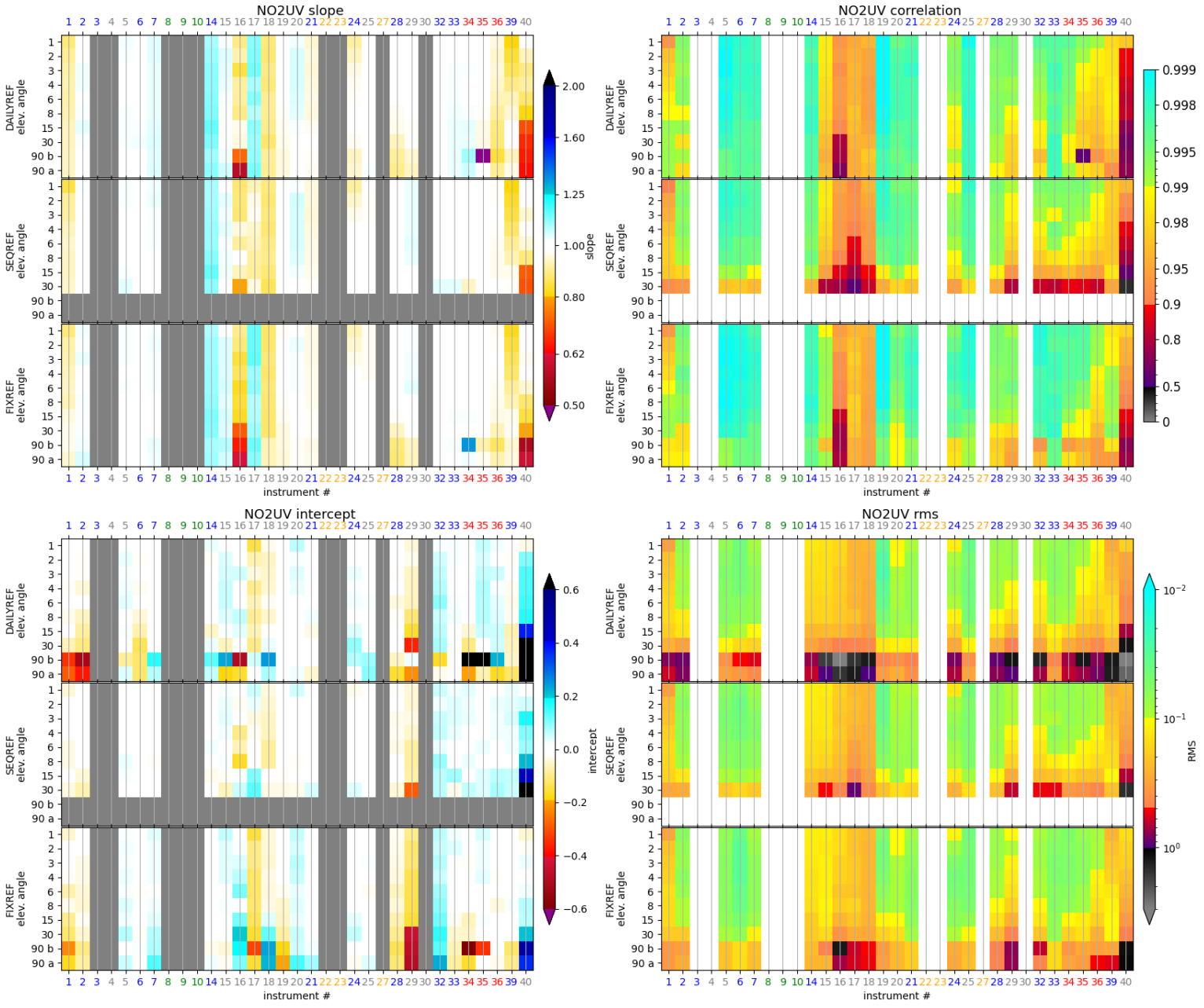

Figure 58: As Fig. 50 but for NO2UV

6.3.4 O4VIS

O4VIS DSCDs are fitted in the same retrieval window as NO2VIS (c.f. Table 3). In Fig. 59 we show the regression analysis result for O4VIS for all three reference types. It can be seen that the slope and intercept are excellent for most instruments and comparable to those of NO2VIS. The correlation of O4VIS is generally slightly worse for O4VIS DSCDs than for NO2VIS DSCDs.

Note that, as for NO2VIS, the performance of #3 (low VEA only), #27 #39, #40 is slightly worse than for most other instruments, likely to the same reasons already discussed. Instruments #4, #6 and #30 however show considerably worse performance for O_4 than for NO_2 in terms of correlation R and for higher VEA also for the RMS. The O_4 absorption features are weaker and broader compared to those of NO_2 in the visible. Hence the polynomial fit baseline correction applied in the DOAS fit has a larger imact on the O_4 retrieval than on NO_2 and hence causes larger uncertainties.

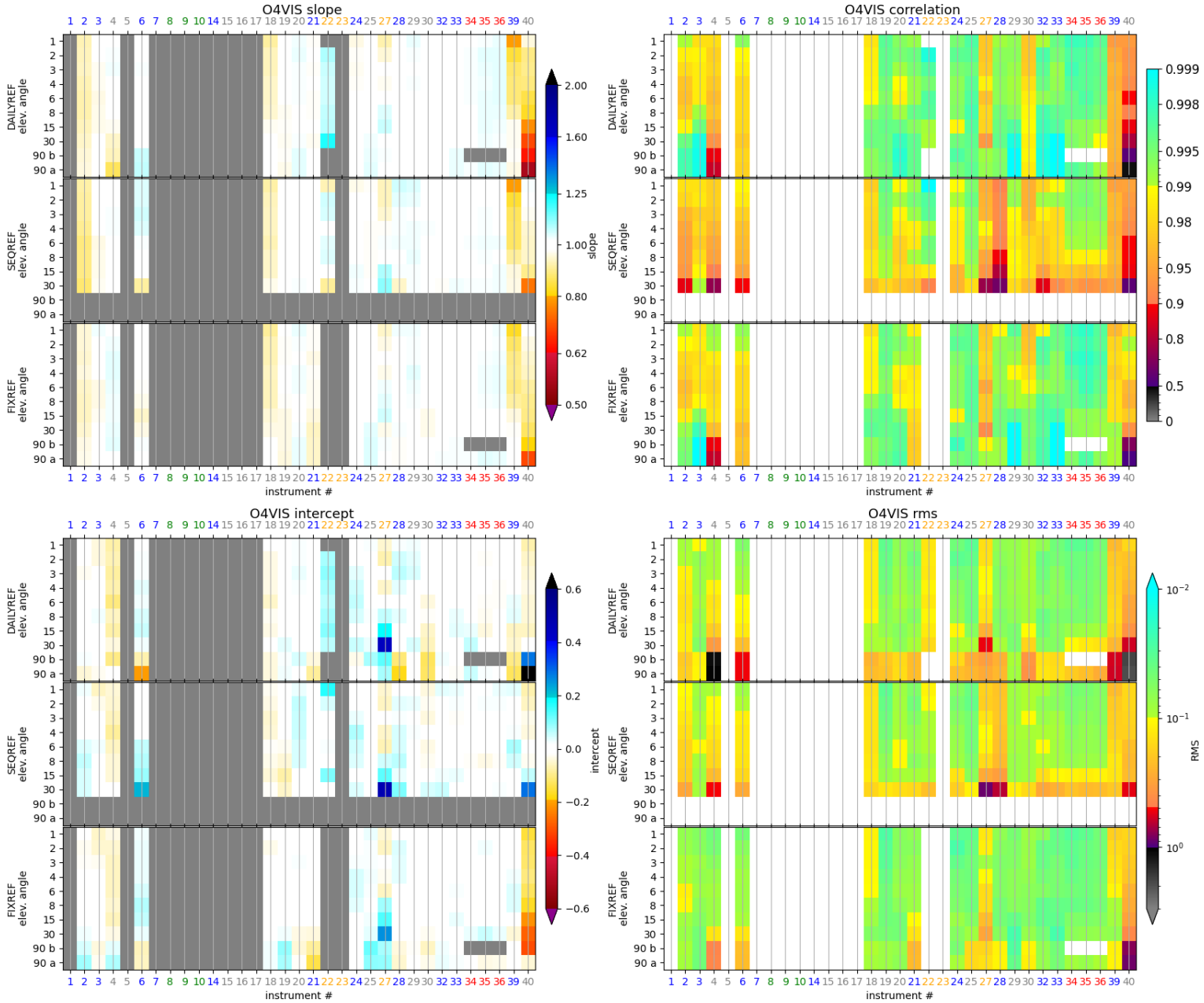

Figure 59: As Fig. 50 but for O4VIS

6.3.5 O4UV

As in the visible, also in the UV, O_4 is fitted in the same wavelength window as NO_2 (c.f. Table 3). We show the regression analysis results for O4UV in Fig. 61. In the visible, (see Subsect. 6.3.4) it was observed that the slope for O_4 DSCDs were comparable to the NO_2 and the correlations were slightly worse.

In the UV wavelength range, with the exception of #18, the slope for O4UV seem in fact better than for NO2UV, compare Fig. 61 and Fig. 58. From Fig. 108 it can be seen that in general, there is a larger spread around the median for #18, but there are also a number of outliers. These outliers are distributed over almost all days, but mostly limited to evening measurements after 17:00 UTC as can be seen in Fig. 60. Similar plots for NO_2 reveal that also for NO_2 in the UV, the performance is worse in the evening. This lower performance is not as visible in the morning. If this lower performance is related to the lack of cooling of the detector, and hence a larger detector noise, then the reason for this not showing up as clearly in the morning might be because of the lower morning temperatures compared to the evening temperatures, especially inside the containers that heated up during the day time.

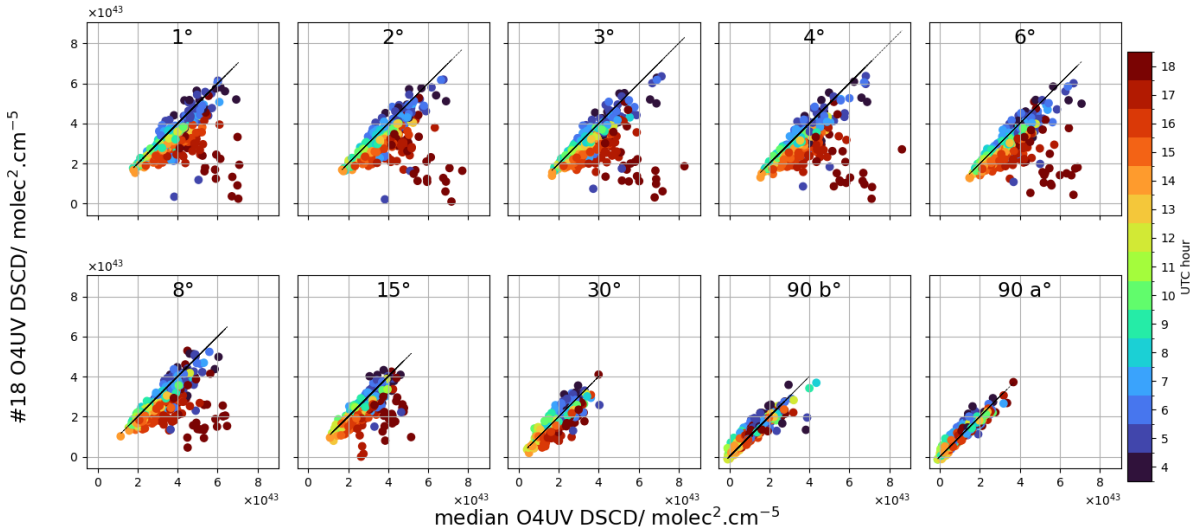

Figure 60: O4UV_FIXREF DSCD #18 against median with color coded UTC.

It is noteworthy to point out #35 DAILYREF at the 90° measurement before the scan. In total, there is a very small number (only 4) of valid values due to the stricter time criterion on the 90° angle as outlined above. This explains the large difference.

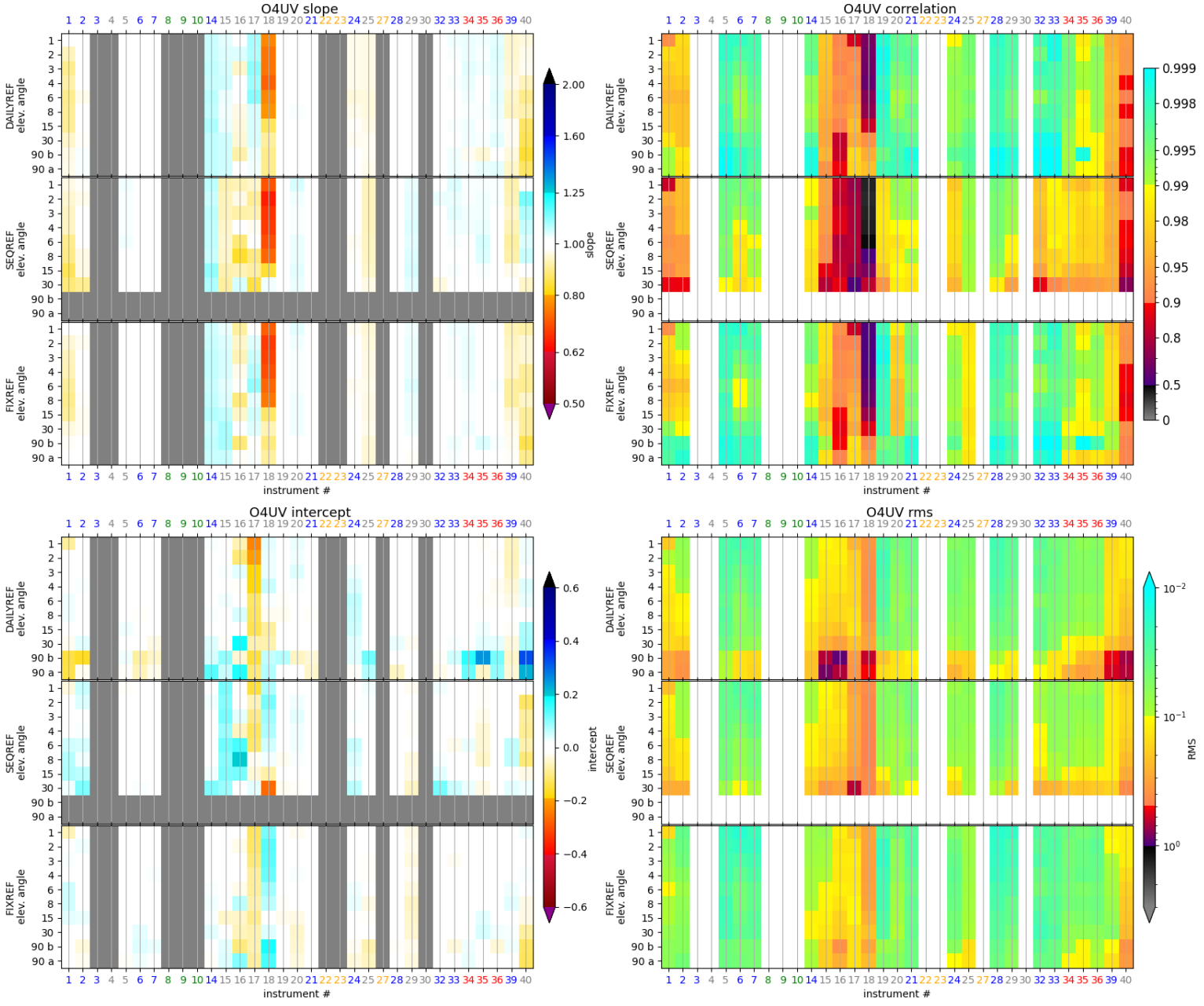

Figure 61: As Fig. 50 but for O4UV

6.3.6 O3VIS

For O_3 in the visible, O3VIS, the slope and correlation of the DSCD regression analysis is exellent for almost all instruments (c.f. Fig. 62) for DAILYREF and FIXREF. The larger deviation towards smaller slopes for instruments #1 and #2 for SEQREF is striking. However, the general correlation for SEQREF is poor, also because the the bulk of the ozone is located in the stratosphere and the large variation in DSCD is related to different AMF in the stratosphere during noon and other times of the day. If instead a zenith reference from the same time is used, the DSCD is hugely reduced and the range much smaller. However, there are still a number of instruments (#19, #25, #30, #34, #35, #36) still show good correlation.

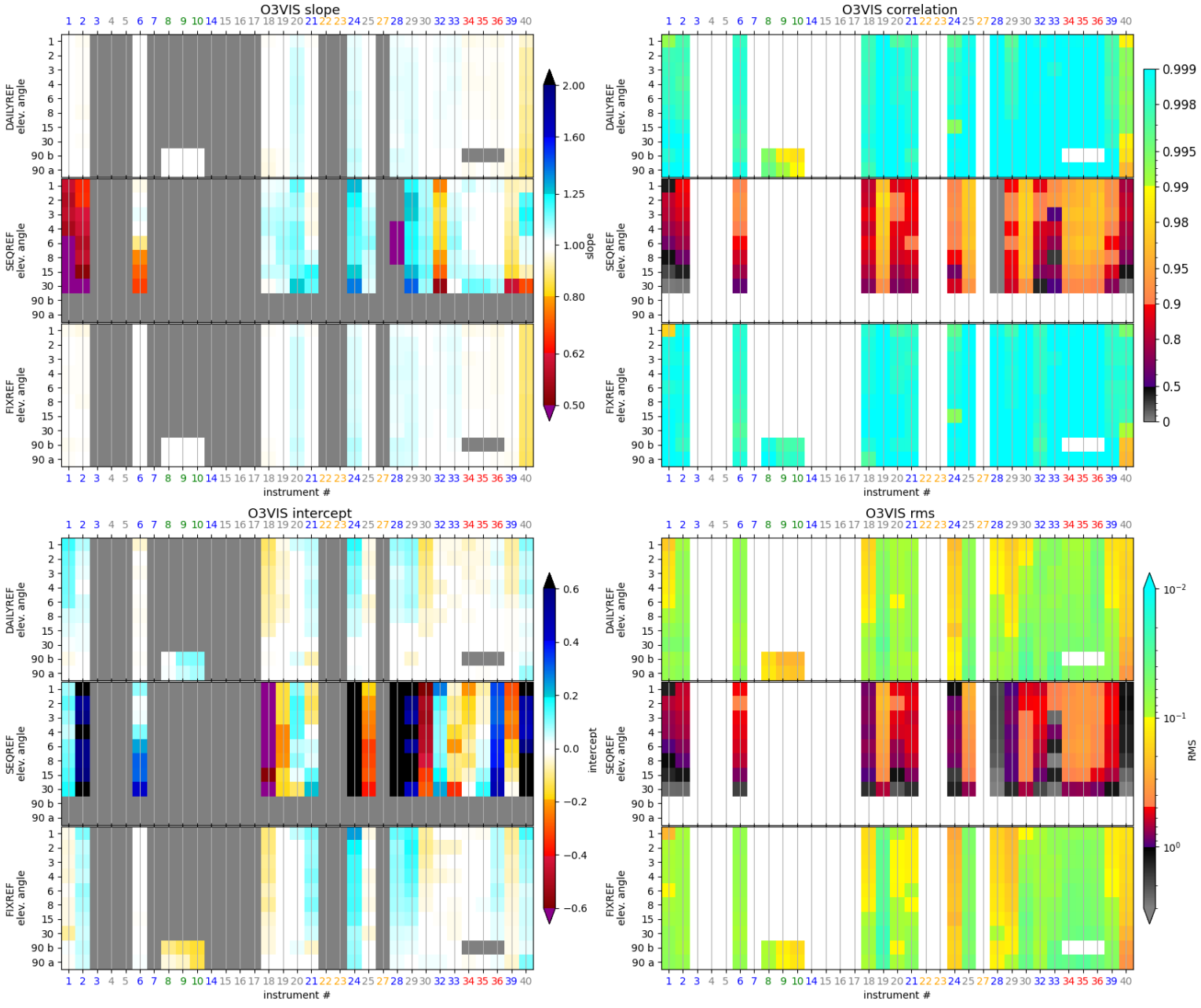

Figure 62: As Fig. 50 but for O3VIS

6.3.7 O3UV

The regression analysis for O3UV is shown in Fig. 63. There are a number of instruments that show less good correlation compared to the visible: #1, #2, #16 (no visible), #17 (no visible) and #18. #17 shows considerably higher correlation and lower RMS for FIXREF than for DAILYREF for lower VEA. Fig. 110 shows that the bad performance for low VEA for DAILYREF can be attributed to a very limited number of outliers, specifically belong to 2 scans at 18:39 – 18:39 on 5 June and 6 June. For #18, as previously discussed for O4UV, the large deviations from the baseline are mainly found in the evening scans after 17 UTC. Otherwise, the performance is very similar to the performance in the visible. However, instrument #29 has a high bias and #25 that has a low bias (visible in the slope), and hence both also have a larger RMS.

As for O_3 in the visible, the general agreement for SEQREF is poor.

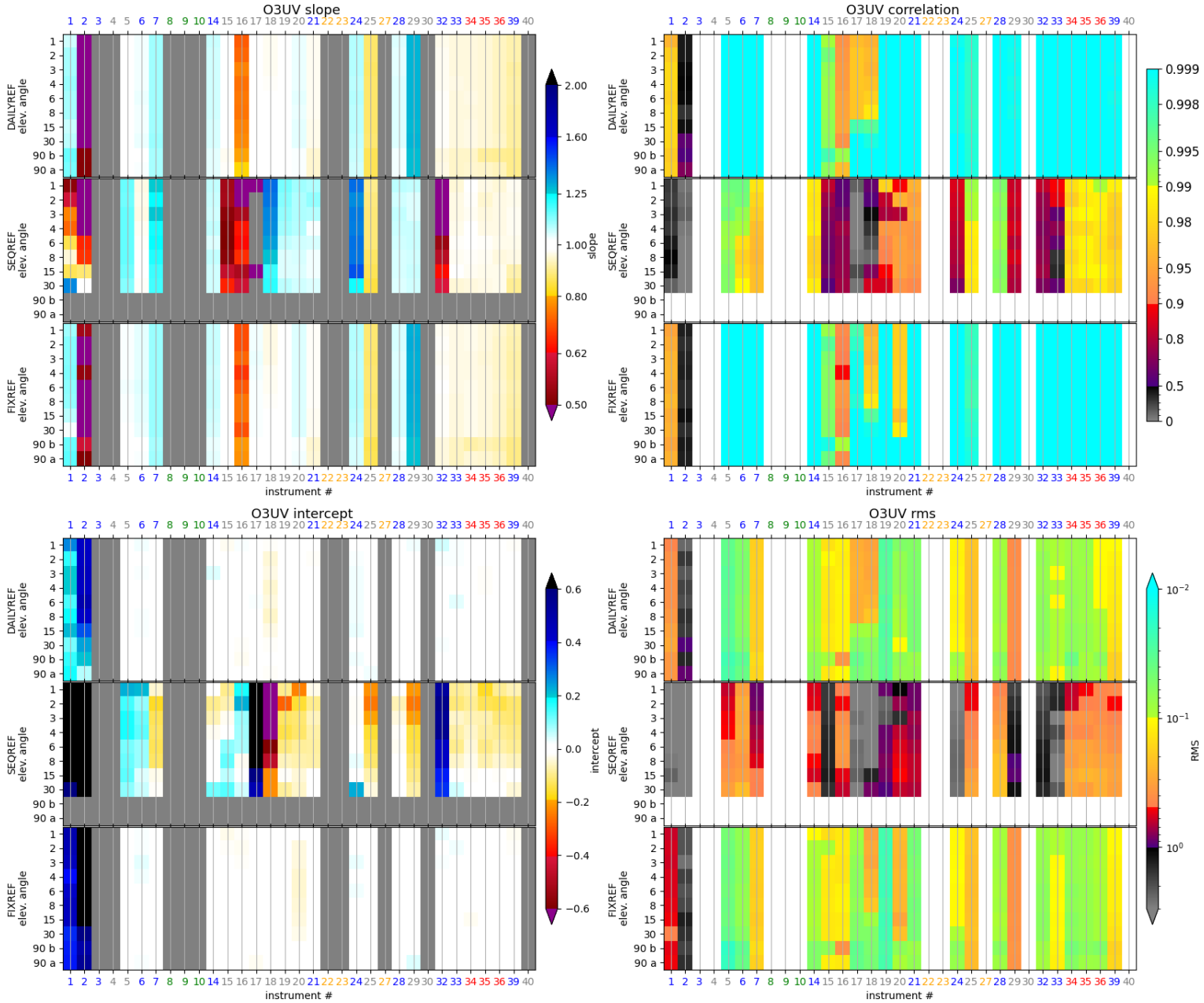

Figure 63: As Fig. 50 but for O4UV

6.3.8 HCHO

In Fig. 64 we show the DSCD regression analysis results for HCHO. The weather conditions were not favourable as described in D7.3.a. This affected the amount of formaldehyde since natural source of HCHO is linked to biogenic volatile organic compoundss (BVOCs) such as isoprene [e.g.][]2004JD005629 emitted by plants. This emission increases with temperature and brighter light conditions. CINDI-3 had mostly cloudy sky and for lower temperatures than CINDI-2.

It is immediately evident from Fig. 64, that the agreement is not good. Therefore, we also used the inverse RMS weighted mean as a baseline to test whether the bad agreement is partly due to a baseline affected by poorly performing instruments. Those results are presented in Fig. 65. While there are minor differences, the overall results do not change.

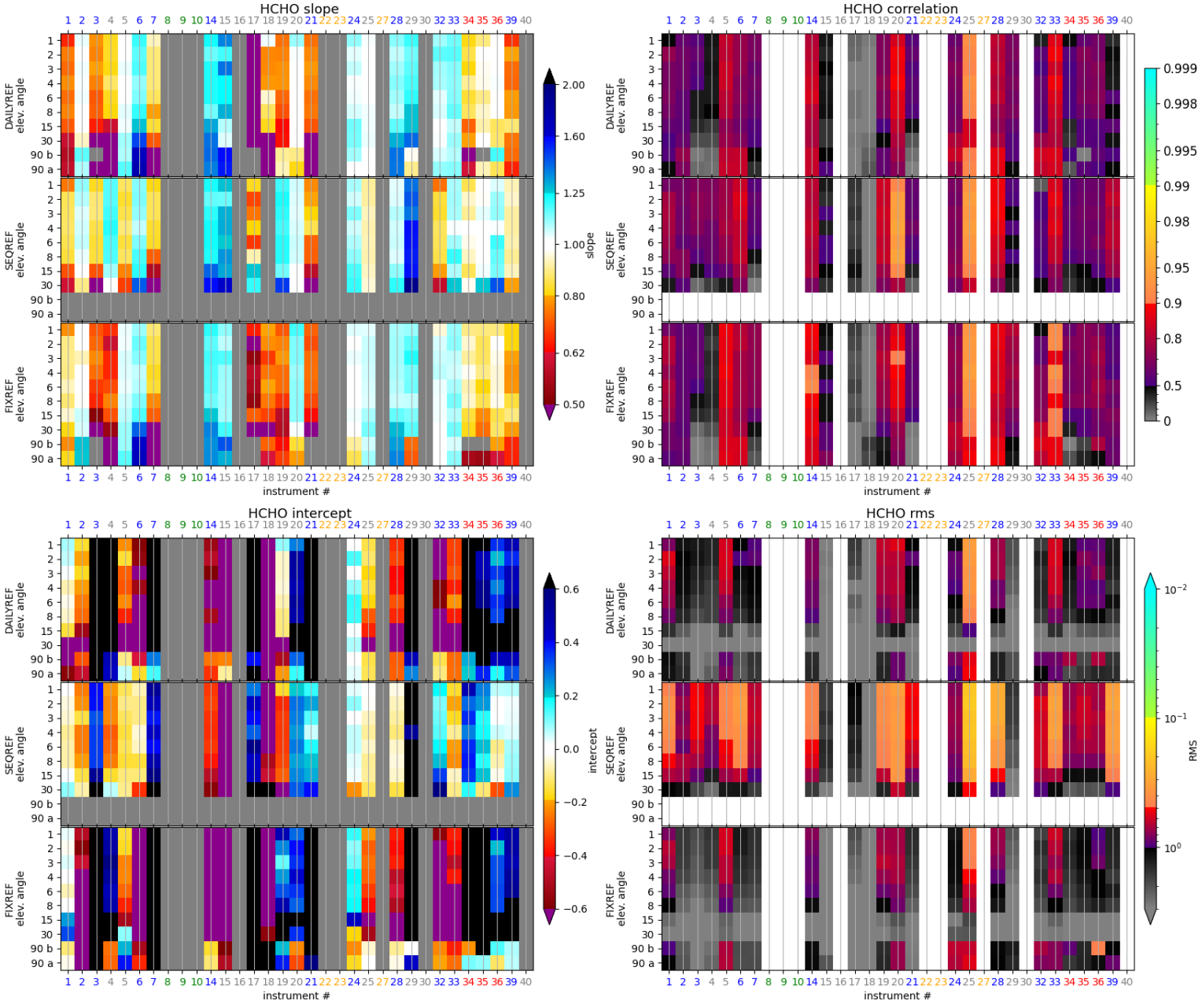

Figure 64: As Fig. 50 but for HCHO.

In Fig. 66 we show time series of median and inverse RMS weighted mean HCHO DSCDs for reference type DAILYREF for two selected VEA, 1° and 30°. The level of DSCD is very low. During CINDI-2, DSCD values at VEA 30° were as high as 5×10^{16} and at low elevation angles values above 13×10^{16} (Kreher et al., 2020). During CINDI-3, 1° DAILYREF DSCDs reached 5×10^{16} and 30° DAILYREF DSCDs reached only 1.3×10^{16}, so about a factor 3 lower. This contributes to the overall lower agreement among instruments.

Based on Fig. 66 we can decide to only consider days with a significant amount of HCHO during some parts of the day, such as 4th, 13th, 14th, 15th, 17th and 18th. We show the regression analysis with the baseline calculated as the inverse RMS weighted mean for those days only in Fig. 67. It

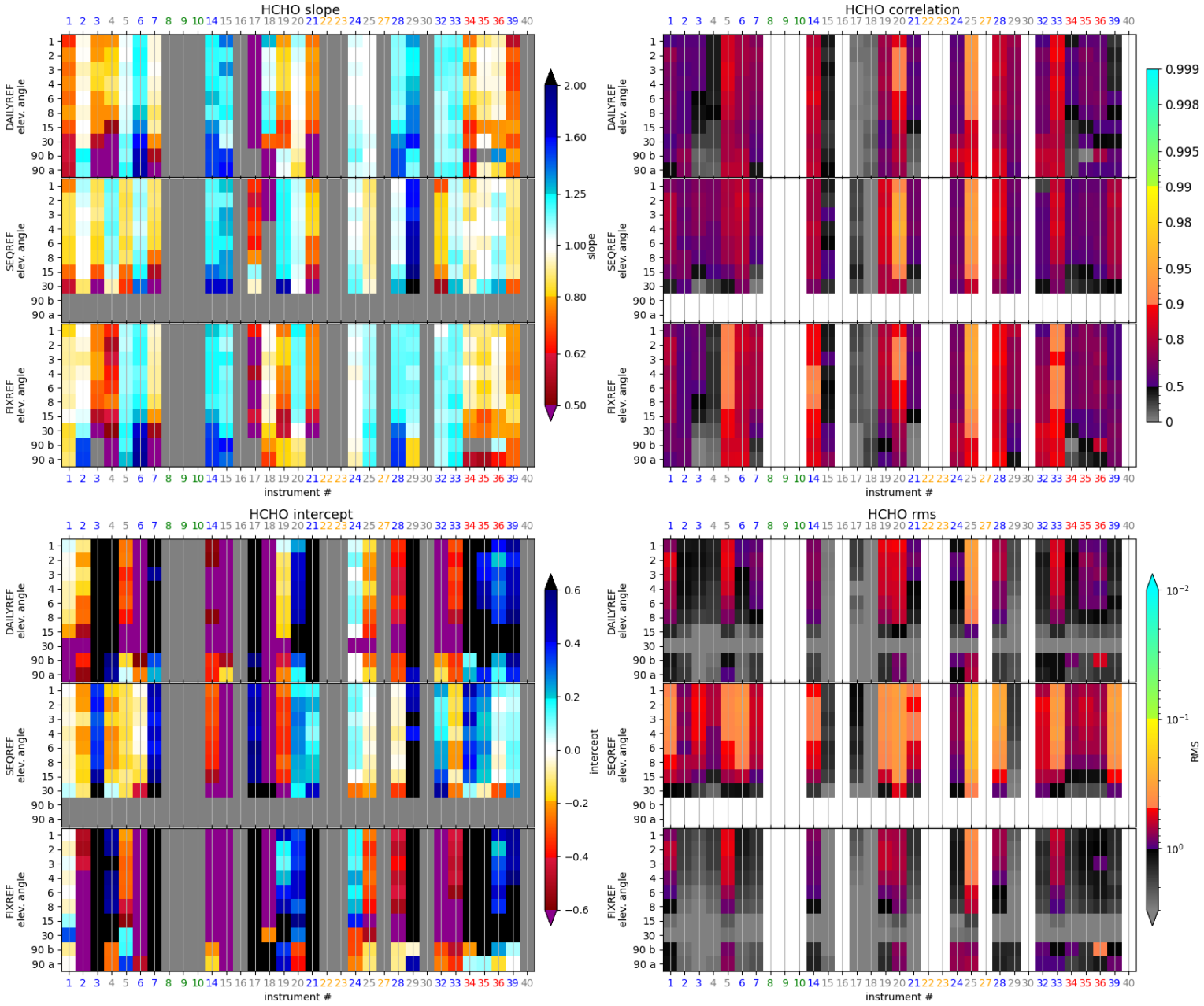

Figure 65: As Fig. 64 but for with the inverse RMS weighted mean as the baseline.

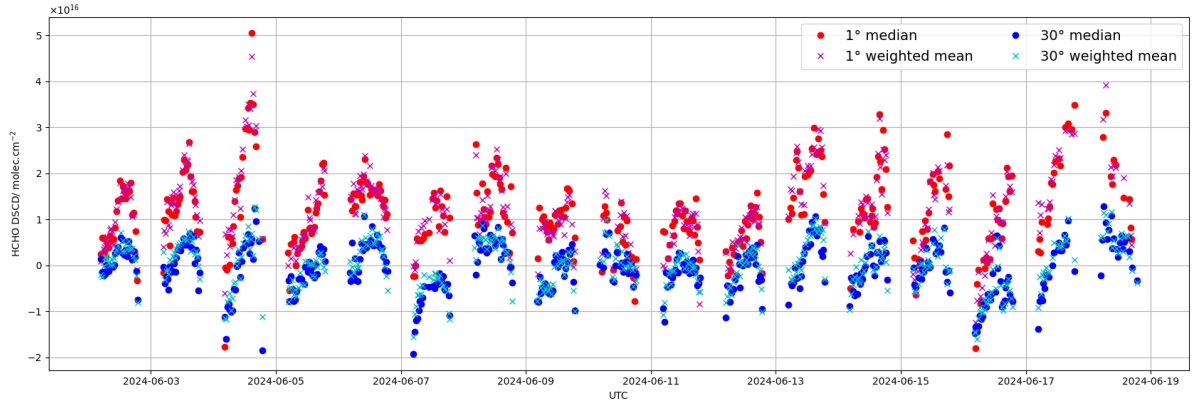

Figure 66: HCHO DSCD weighted mean and median time series at 1° and 30°.

is evident that the effect on the overall regression analysis is small. Considering the slope of the Theil-Senn fit, we can note that for a number of instruments (c.f. upper left panel of Fig. 65) it is still good or even excellent (#2, #5, #20, #24, #25, #32 and #33) for DAILYREF. However, regarding

the correlation R, only instruments #20 and #25 have $R > 0.9$ for most VEA.

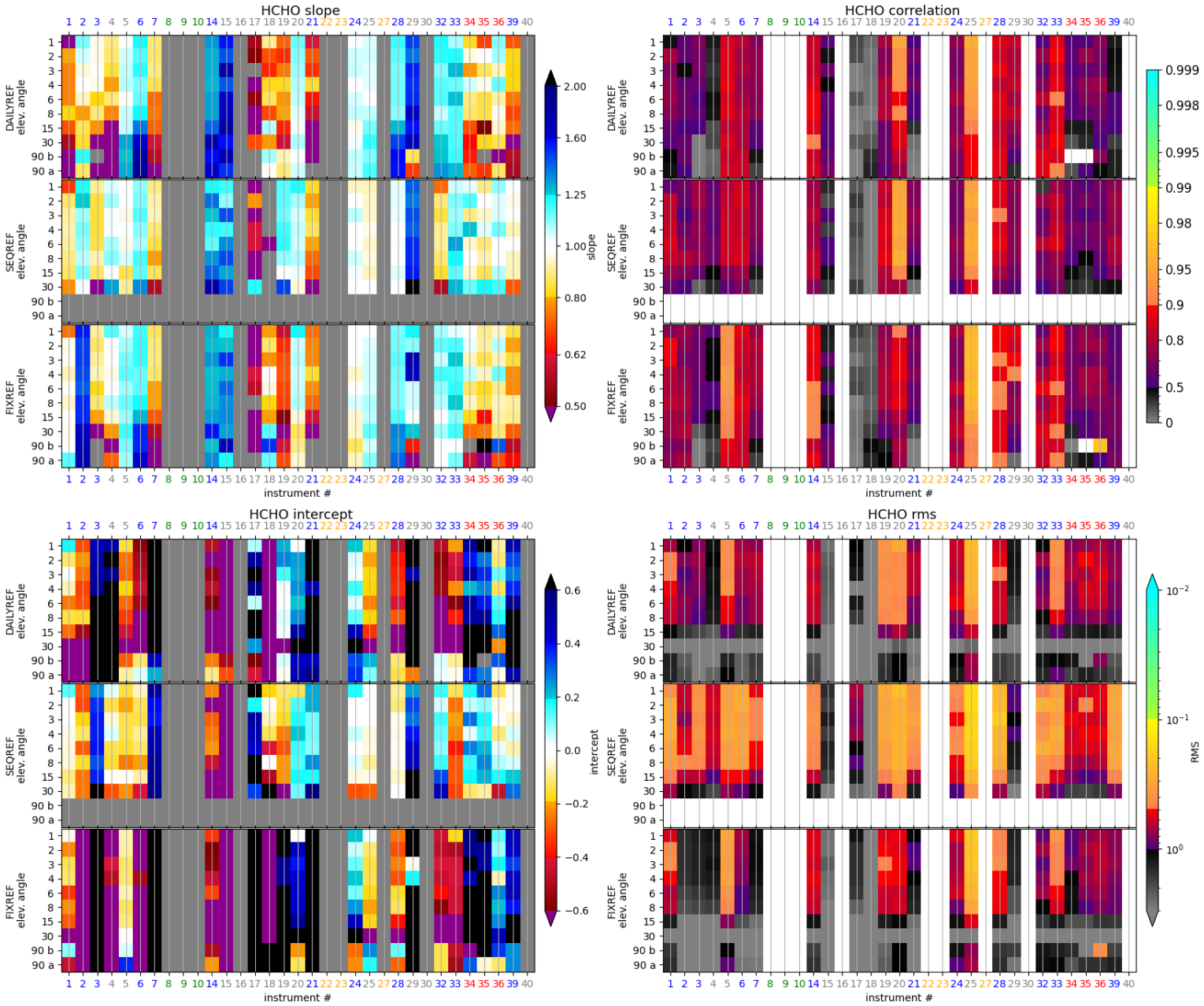

Figure 67: As Fig. 65 but only considering a subset of days: 4th, 13th, 14th, 15th, 17th and 18th June.

In order to exclude that the general low correlation with the baseline (independent of whether the median or the inverse RMS weighted mean is considered and independent of whether only days with higher HCHO load are considered) comes from several poorly performing instruments, we also considered the cross-correlation between different instruments. We show this for SEQREF in Fig. 75. In this plot, the instruments are ordered according to the correlation with the baseline, here the median. This correlation is also shown on the diagonal.

It can be seen that the correlation among the two instruments that showed the best correlation (for SEQREF) with the baseline, #20 and #25, is never better and often worse than each individual instrument's correlation with the baseline. The same is true for the gross of the instruments. Only occasionally, e.g. instruments #2 and #24 at 1° VEA have instruments a better inter-instrument correlation than their correlation with the baseline. We can therefore conclude that the general low agreement is not due to a low-quality baseline, but can rather be attributed to the overall unfavourable weather conditions. It also shows that the comparison to a baseline calculated from the median of all instruments is still a good choice.

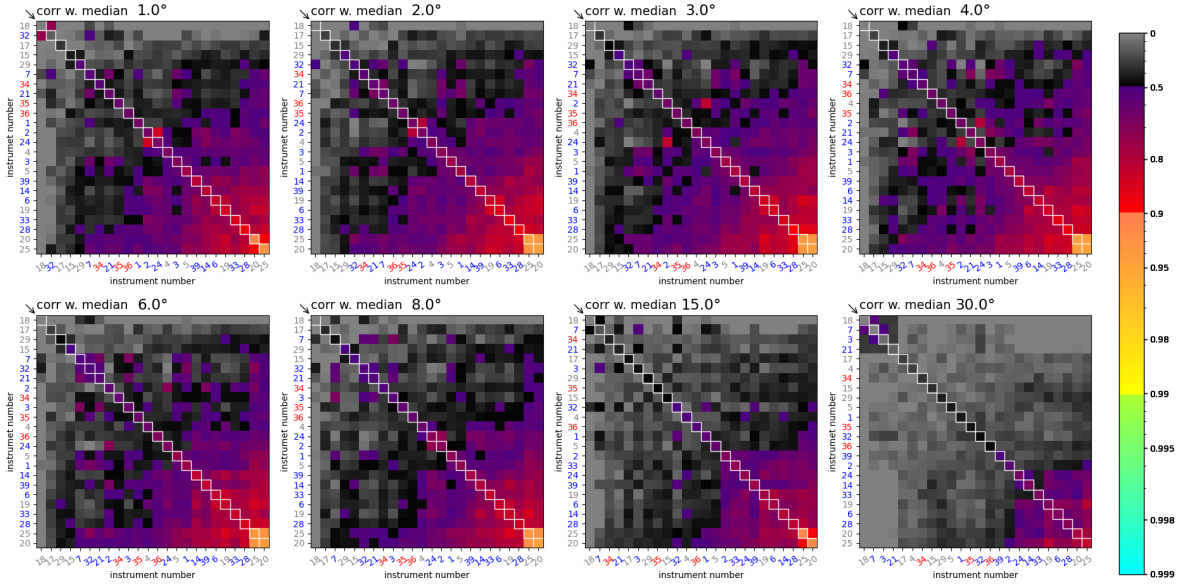

Figure 68: Correlation between different instruments for HCHO_SEQREF. Instruments are ordered according to the correlation with the baseline. This correlation is shown on the diagonal. Each panel corresponds to a different VEA.

6.3.9 HCHO-WIDE

For the wider HCHO fitting window, HCHO-WIDE, DSCD regression results are similar to the ones of the smaller fitting window, HCHO, as can be seen in Fig. 71. However, the correlation as well as the RMS look a bit better for most instruments. In fact, the RMS for the lower most elevation angles, up to 8°, is for many instruments in the good region.

An exception for which the correlation and RMS was in fact for almost all elevation angles in the good range for HCHO, is instrument #25. A closer look at the regression analysis plots for #25 (c.f. Fig. 69) reveals that for FIXREF, there are large differences between different days. It is mostly the early days in the campaign that seems to be problematic. However, also later days are significantly worse than for HCHO.

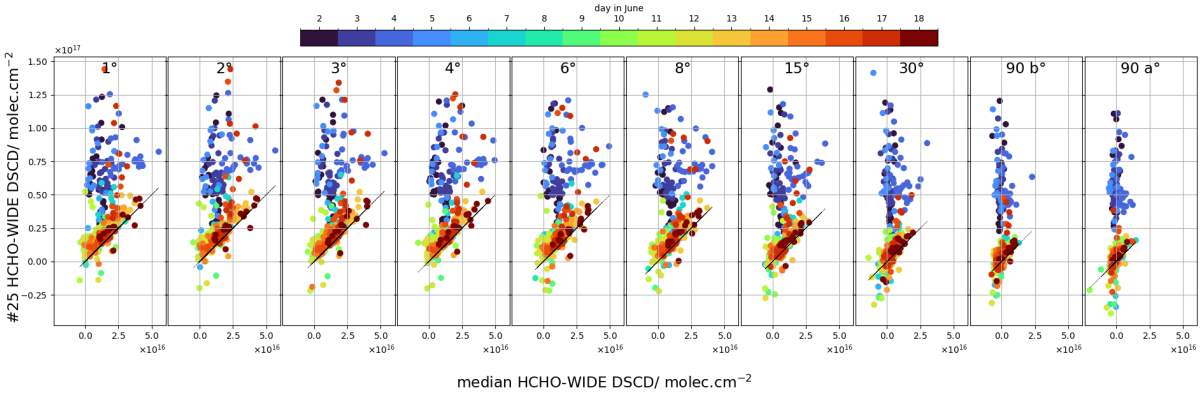

Figure 69: HCHO-WIDE_FIXREF regression plots for #25 against the median with color-coded days. The black line indicates the 1:1 line.

Another exception, however less extreme, is instrument #34. However, as can be inferred from Fig. 70, here it is daytime (approaching twilight, i.e. lower light conditions) that has an influence on the performance of the instrument. It should be noted that instrument #35, also a Pandora instrument, shows a better correlation and a lower RMS in the wider window, HCHO-WIDE. It was already established in Sect. 5 that instrument #34 performs worse under low light conditions compared to the other two Pandora instruments.

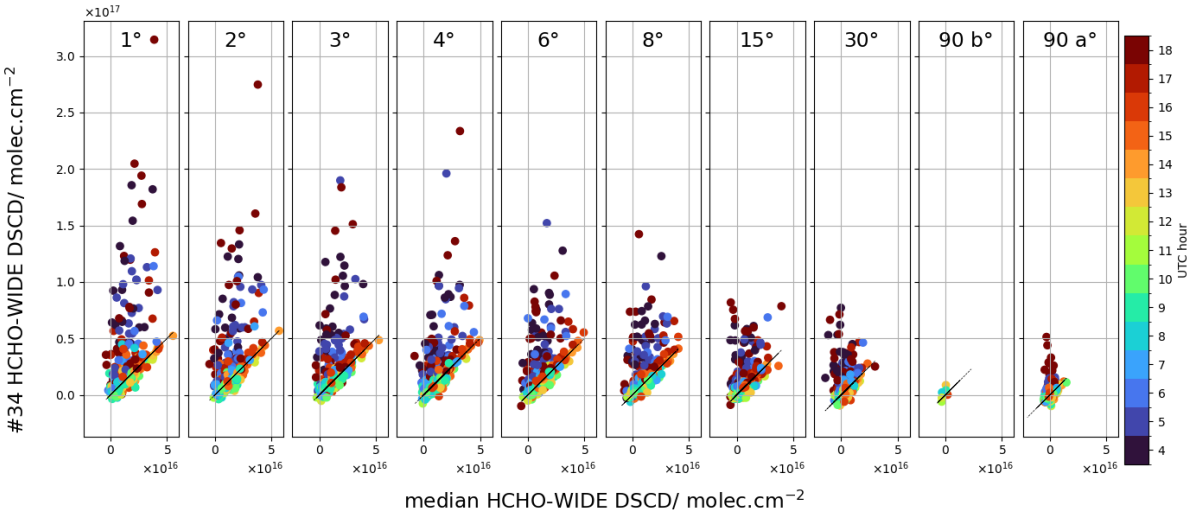

Figure 70: HCHO-WIDE_FIXREF regression plots for #34 against the median with color-coded UTC. The black line indicates the 1:1 line.

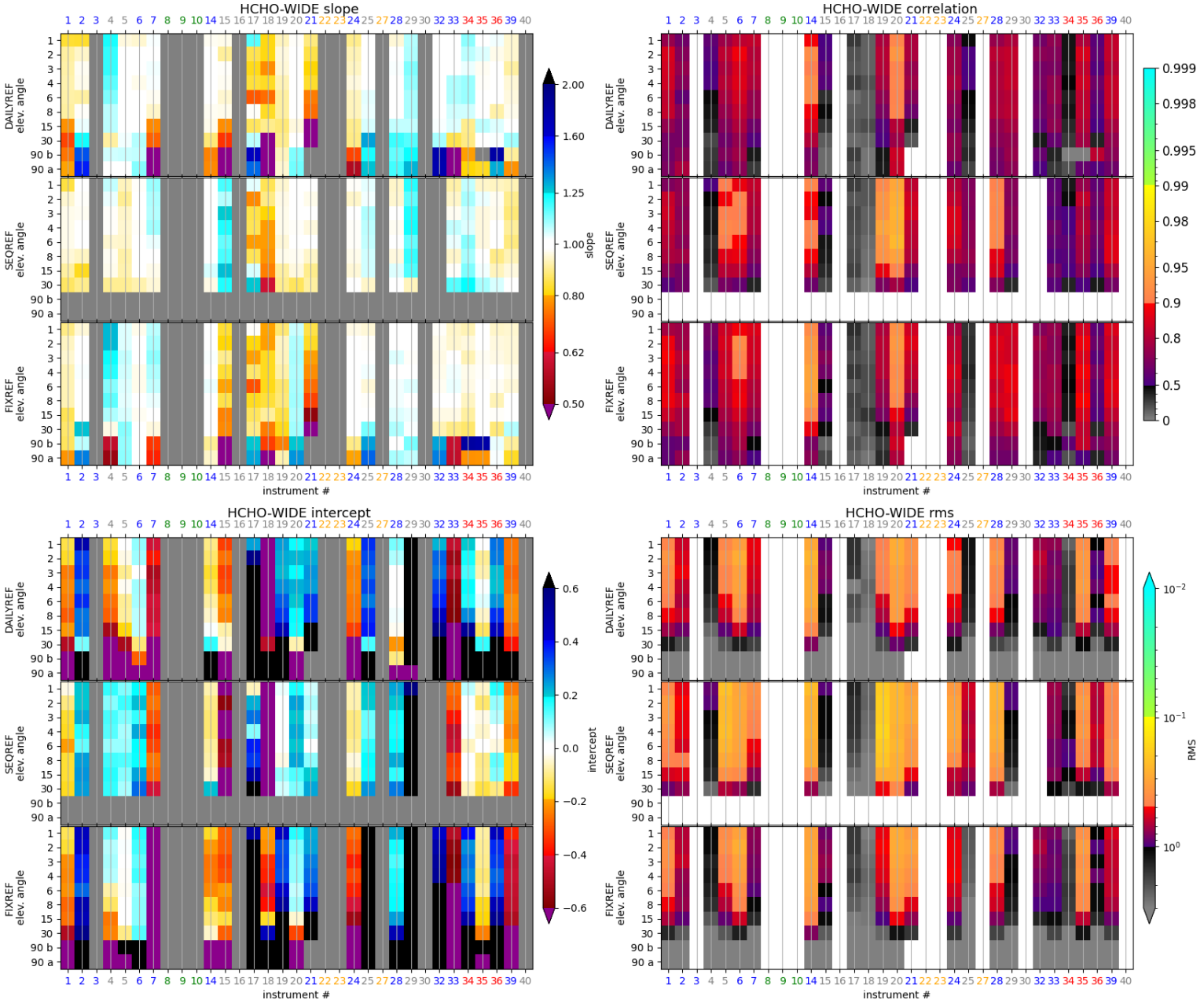

Figure 71: As Fig. 50 but for HCHO-WIDE

6.3.10 HONO

The regression analysis results for HONO are shown in Fig. 72. The overall agreement is very low. This is visible in the correlation results which are below 0.9 for all instruments at all elevation angles except for #28 at 1° elevation angle for DAILYREF, and additionally for 1° and 3° for FIXREF. Considering the intercept and the RMS almost all instruments at almost all elevation angles are in the bad region.

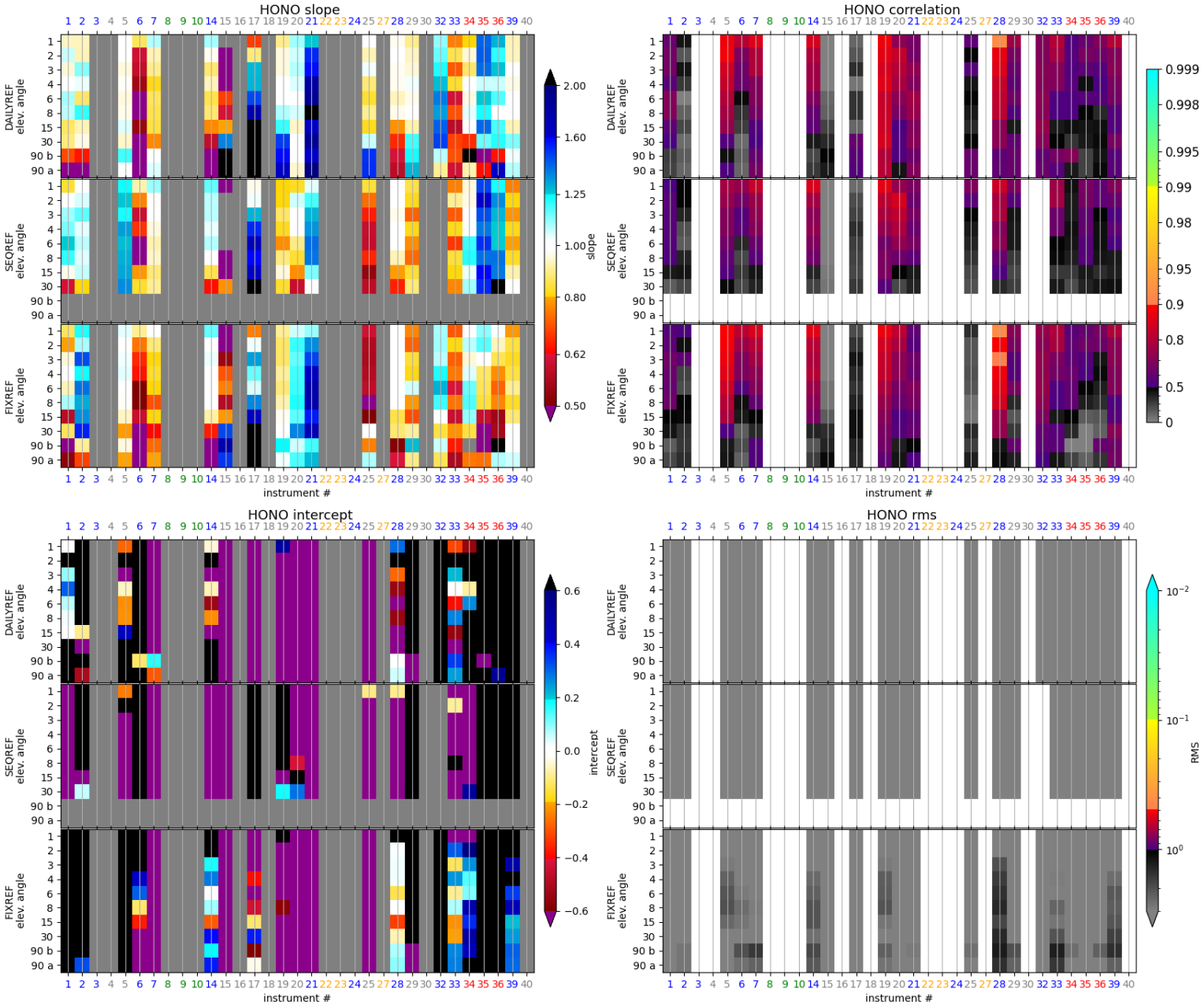

Figure 72: As Fig. 50 but for HONO

If we consider a time series of both baselines (i.e. the median and the inverse RMS weighted mean) for HONO at two different elevation angles, 1° and 30° (c.f. Fig. 73) it is visible that there are a number of mornings, 4,6, 8, 10, 13,14,17,18 June 2024, for which there might be some HONO confined to the surface present.

If we restrict the regression analysis, similar as was tested for HCHO, to those days in the mornings (4–11 UTC), we see that the regression analysis improves considerably in terms of correlation and RMS, especially for lower VEAs, c.f. Fig. 74. We can therefore conclude that the lack of agreement for HONO is due to its absence during most of the inter-comparison time period.

As for HCHO, we can look at the inter-instrument correlations for HONO. This is shown in Fig. 75. It is evident that the correlation with the median is mostly better than the correlation among instruments. For the lowest VEA, there are 7 instruments (#19, #25, #20, #7, #39, #14, and #28) which have a good correlation with the baseline.

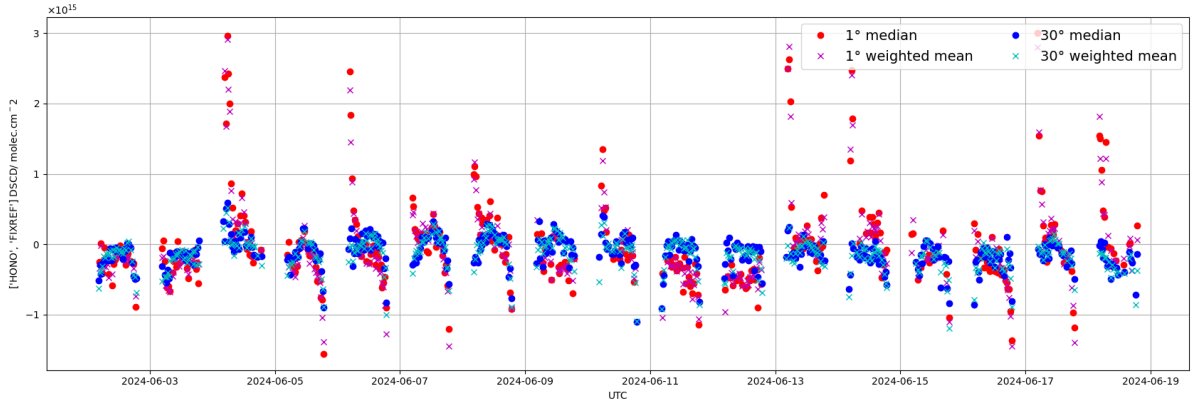

Figure 73: Time series of HONO median and inverse RMS weighted mean DSCDs at 1° and 30° VEA.

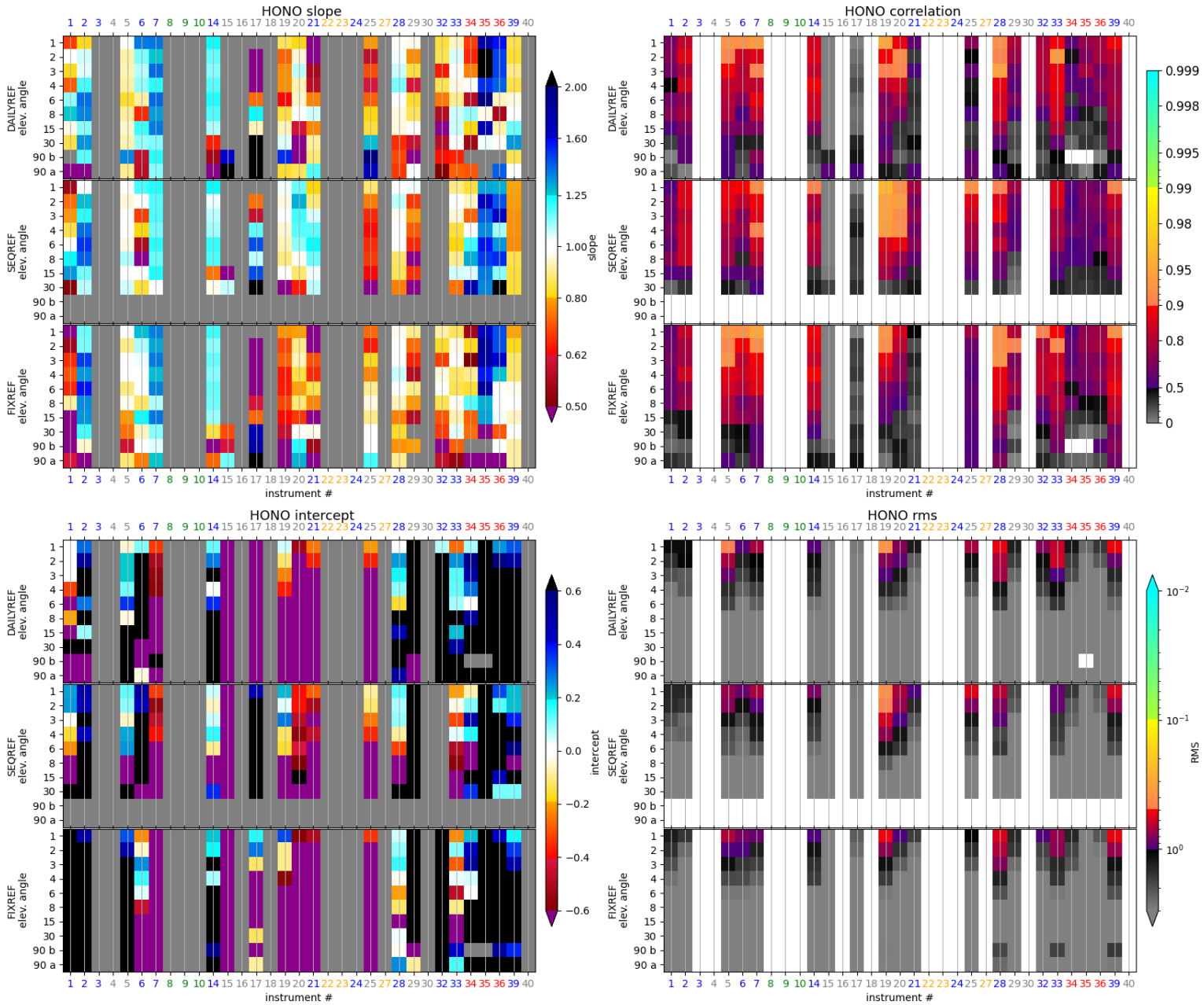

Figure 74: As Fig. 72 but only considering morning hours (4–11 UTC) on a subset of days: 4,6, 8, 10, 13,14,17,18 June.

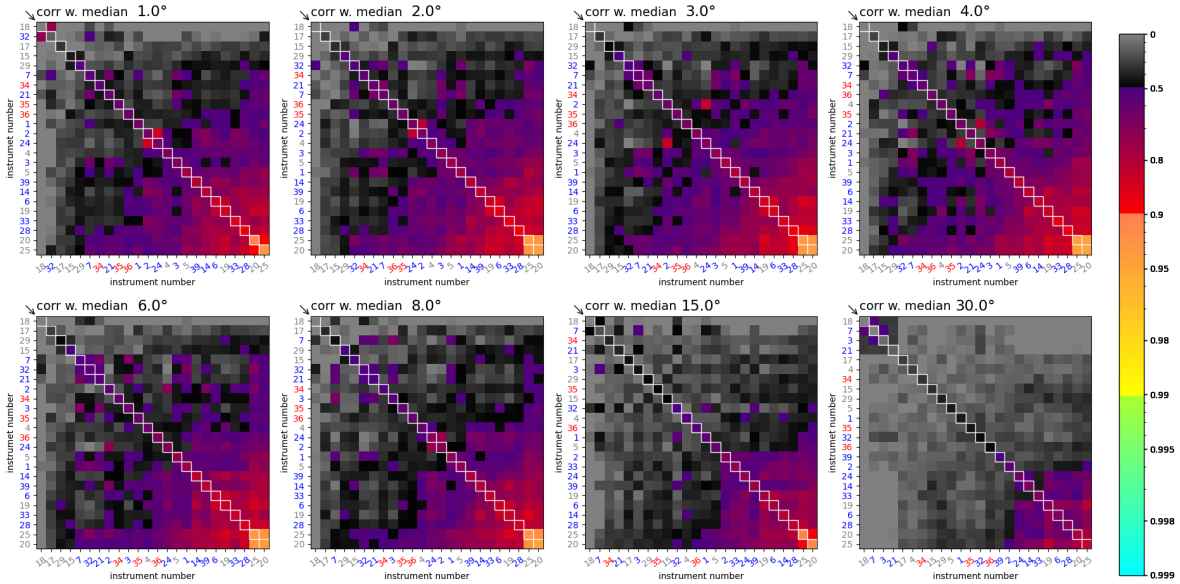

Figure 75: Correlation between different instruments for HONO_SEQREF considering only mornings of selected days, see main text. Instruments are ordered according to the correlation with the baseline. This correlation is shown on the diagonal. Each panel corresponds to a different VEA.

6.4 Final remarks

The introduction of a fixed reference FIXREF improved in many cases the agreement between the different instruments despite the larger influence of instrument instabilities. This might be due to the stricter prescription of which spectra to include exactly in the reference (e.g. for FIXREF, first 9 spectra of the reference spectra taken on 6 June). It might be therefore recommendable to enforce stricter rules for the selection of daily reference spectra as well.

The alternative wavelength window for NO_2, NO2VIS-SMALL not only showed very consistent DSCD results to NO2VIS, but also comparable spread between the instruments. It is therefore a very valuable addition for those instruments that do not cover the whole wavelength range of the retrieval window for NO_2 in the visible.

The HCHO, HCHO-WIDE and HONO inter-comparison greatly suffered from the poor weather conditions with low temperatures and mostly cloudy skies.

There are several contributing factors for an instrument comparing worse to the baseline, among those:

- bad time matching: most likely the predominant cause of the low correlation for NO2VIS-SMALL_FIXREF for instrument #23 that is inherent by the working mode of this imaging instrument; maybe a contributing cause for the lower performance of instrument #40.
- bad vertical alignment as seen for instrument #3
- large thermal noise for un-cooled detectors, especially visible for UV products in the evening hours for instrument #18
- larger FOV compared to other instruments, such as higher VEAs for #27

7 Comparison to FRM4DOAS centralized processing

16 instruments (in 23 channels) delivered L1 spectra to the FRM4DOAS centralized processor. The L1 were checked for compliance and DSCDs were fitted with the default FRM4DOAS settings, with the exception of using the zenith after for the HCHO-WIDE window (the default being to use a time interpolation of the zenith measurements bracketing the scan), in order to agree with the settings during CINDI-3 used by the PI. However, other details differed such as the window used for O_4 in the visible which is [425, 490] nm in the CINDI-3 DSCD from the instrument PIs and [450, 540] nm from the FRM4DOAS processing.

Since the default FRM4DOAS DAILYREF uses only 1 spectrum, namely the zenith spectrum at the smallest SZA, we do not compare the DSCD directly, but the Δ_{DSCD}, see Eq. 4 in order to eliminate the day-to-day shift introduced by the differences in the daily reference.

Of the 16 instruments, 2 instruments were corrected in the "own" submission: #3 provided interpolated DSCD at low VEA. Instrument #21 applied a stray-light correction to their spectra in the visible channel before submitting the final version of DSCD visible products. However, they have not provided updated L1 spectra files in the visible to the FRM4DOAS centralized processor. Hence, Both instruments are excluded for the NO2VIS comparison. Further, #32 did not submit SEQREF analysis and is hence also excluded from the HCHO-WIDE analsyis. We summarize which instruments where compared for which product-reference type combination in Table 9.

Table 9: Instrument numbers entering the comparison for each product-reference type.

product-reference type	**instrument # entering the comparison**
Δ DSCD NO2UV_DAILYREF	19, 20, 24, 25, 32, 33, 34
Δ DSCD NO2VIS_DAILYREF	18, 19, 20, 24, 25, 28, 30, 33
Δ DSCD NO2VIS-SMALL_DAILYREF	4, 5, 7, 14, 28
Δ DSCD O4UV_DAILYREF	7, 14, 19, 20, 21, 24, 25, 28, 32, 33, 34
Δ DSCD O4VIS_DAILYREF	19, 20, 24, 25, 28, 30, 33
DSCD HCHO-WIDE_SEQREF	4, 5, 7, 14, 18, 19, 20, 21, 24, 25, 28, 33, 34

In Subsect. 7.1 to Subsect. 7.6, we will discuss the differences between the results from the instruments PI DSCD and the centrally FRM4DOAS fitted DSCD. The former is referred to as "own", the latter as FRM4DOAS. The main figure in each subsection shows time series of own and FRM4DOAS DSCDs or Δ DSCDs, depending on the reference type, and the number of contributing instruments at three different VEA in the first row. The selected VEAs are 1°, 4° and 15°. The second row of each main figure shows for those times where at least 3 instruments contributed, the difference between the medians of the two datasets, own and FRM4DOAS together with the standard deviation (σ) of each data set. The third row shows the difference of the σs: $\sigma_{\text{own}} - \sigma_{\text{FRM4DOAS}}$. Hence, if the difference is positive, the variation in the dataset containing the PI processed measurements has more variation than the centralized processed data. While the first row always covers the whole y-data range, the second and third row show 99% of all data (covering the 0.5th percentile to the 99.5th percentile). This was chosen due to some extreme outliers.

7.1 NO2VIS_DAILYREF

For NO2VIS, we find that the difference between the medians of the two datasets, own and FRM4DOAS is always much smaller than the σ of each dataset (c.f. middle row of Fig. 76). This indicates that the differences in the datasets are likely really in the spectra and not in the way they are processed. We also see that the difference between the respective σ is small compared with their values (compare middle row to last row).

Further we see that the σ in the "own" dataset is mostly slightly larger than in the FRM4DOAS dataset until and including 16 June 2024 and vice-versa after. From the number of instruments contributing, we see that three instruments dropped out after that date, either in the "own" or in the FRM4DOAS dataset (only coincidence data is included here). These instruments are #25, #28 and #33. We will come back to these when we discuss the differences between the own and FRM4DOAS processed DSCDs individually for each instrument.

For each of the instruments included in the NO2VIS comparison, #18, #19, #20, #24, #25, #28, #30, #33, we show a regression plot of DSCD and RMS between own (on y-axis) and FRM4DOAS (on x-axis) in Fig. 77 to Fig. 85. The former includes the correlation R, Theil-Sen slope s, number of coincidences (the number in brackets is the theoretically possible number) and the RMS of the difference between own and FRM4DOAS DSCD. Note that the scale for RMS is logarithmic on both axes. Note that these include all VEA and show the Δ DSCD for DAILYREF products and DSCD

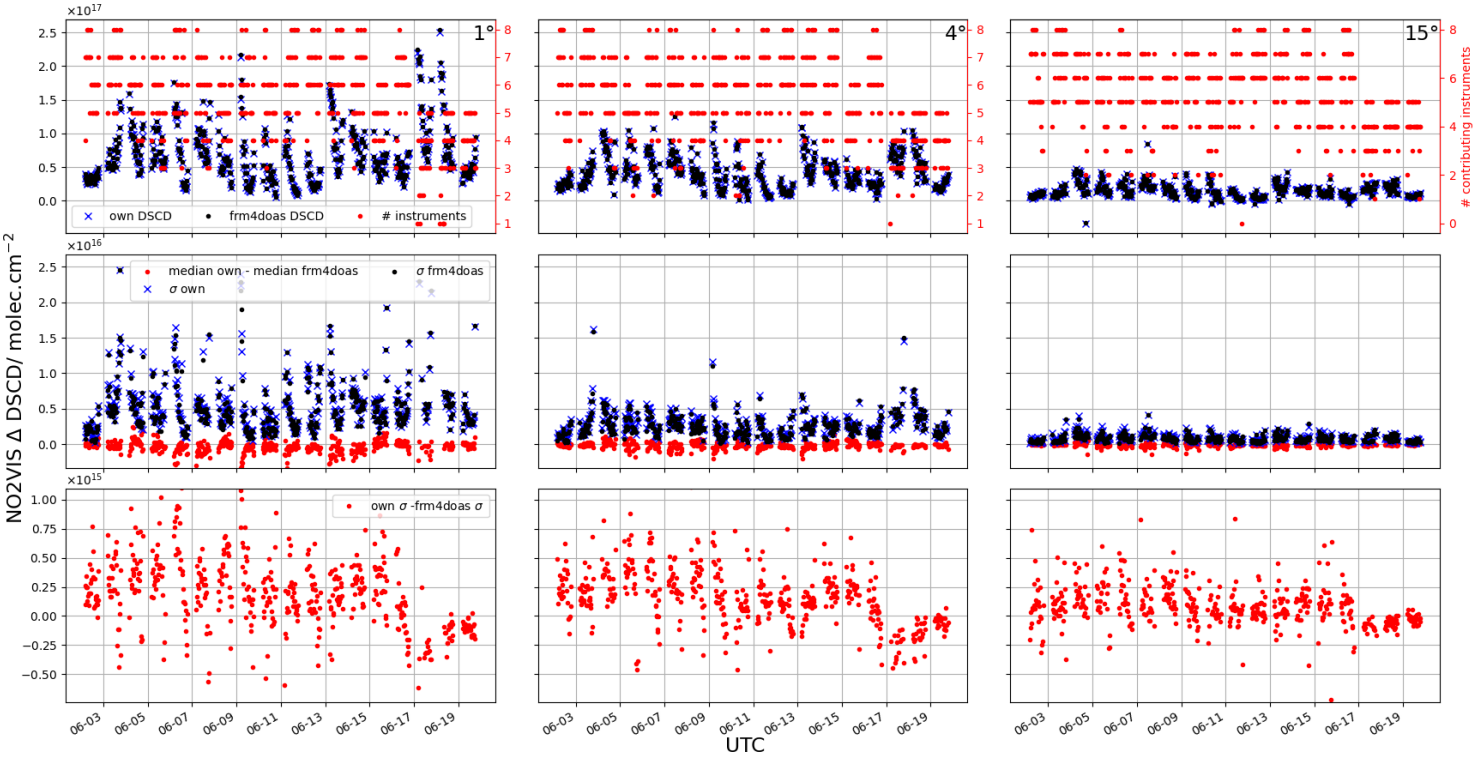

Figure 76: For 3 different VEA, left to right: Time series at VEA 1°, 4°, 15°: Upper panels: own and FRM4DOAS fitted Δ DSCD (left axis) and number of contributing instrument (right axis), Middle panels: difference between the corresponding medians for own and FRM4DOAS processed DSCDs and own and FRM4DOAS processed σ between different instruments. Lower panels: difference between the σs of own and FRM4DOAS processed DSCDs.

for SEQREF products. Not all of them will be shown for the other products. The complete list of plots is provided in Appendix C.

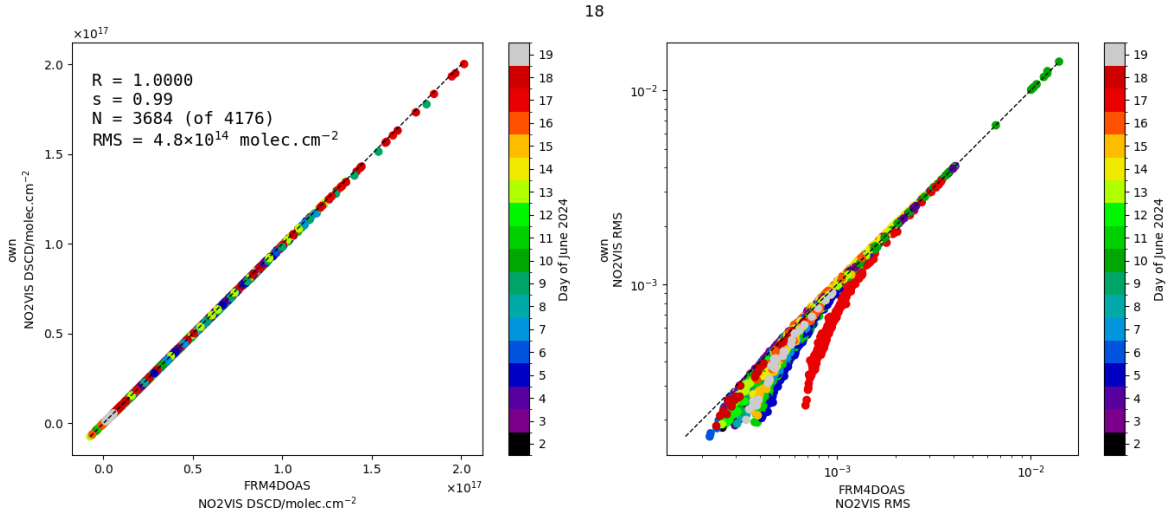

Figure 77: Regression analysis plot for own vs FRM4DOAS Δ_{DSCD} (left) and RMS (right) for instrument #18. Color coded is the day of June.

For #18 (compare Fig. 77) we see that the correlation between own and FRM4DOAS is excellent, as well as the slope and the RMS between the two datasets. It is noteworthy that there are day-dependent systematic differences in the fitting RMS, where the own RMS is smaller.

For #19 (compare Fig. 78) we also see excellent agreement between the two DSCD datasets (left panel). However, we see that the fitting RMS for the FRM4DOAS dataset is almost always slightly smaller than in the own processed one.

For #20 (compare Fig. 79) we see a good correlation, a good RMS and an excellent slope. The correlation and RMS are affected by two outliers. Regarding the fitting RMS, there are no systematic

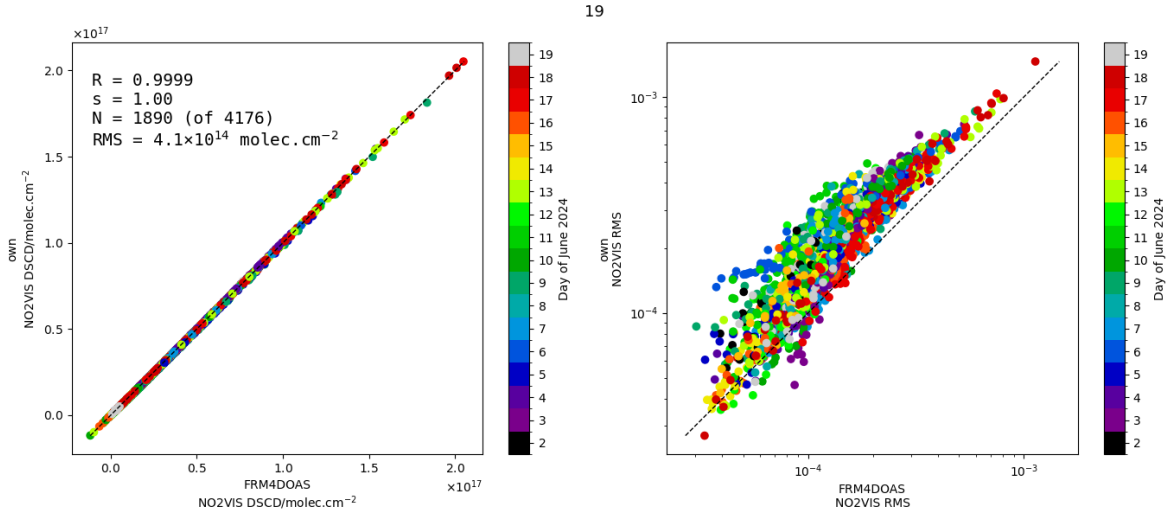

Figure 78: As Fig. 77 but for instrument #19

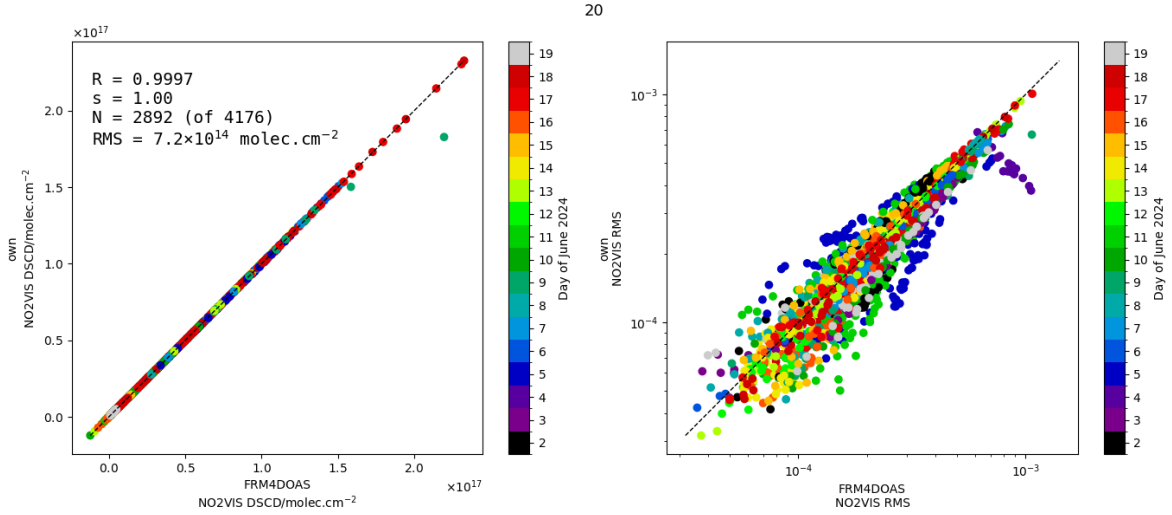

Figure 79: As Fig. 77 but for instrument #20

differences visible that depend on the day.

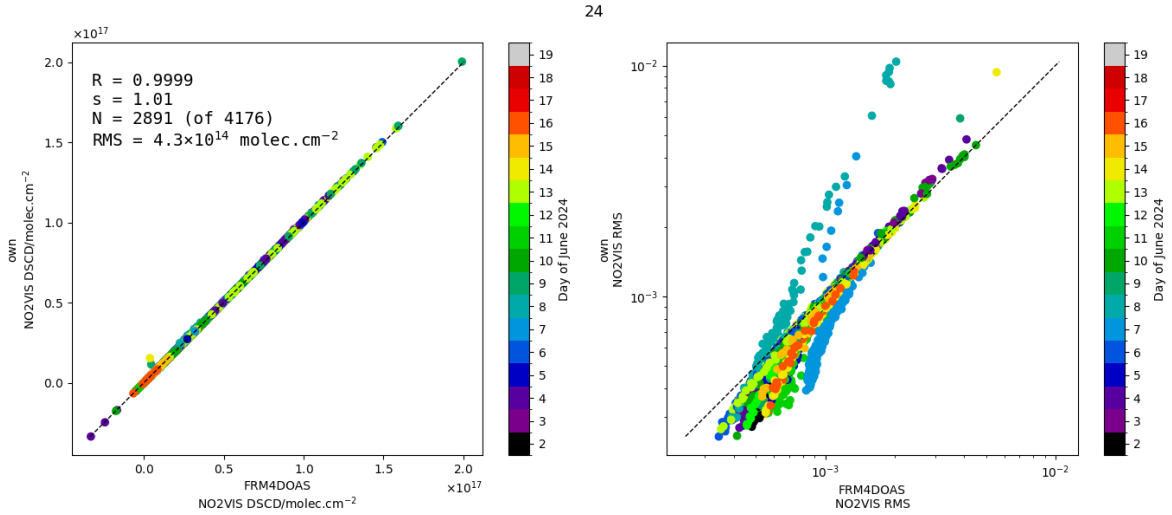

Figure 80: As Fig. 77 but for instrument #24

For #24 (compare Fig. 80) we see a very good correlation, slope and RMS between own and FRM4DOAS DSCDs. Regarding the fitting RMS, there are systematic differences depending on the day. E.g. On 8 June, own fitting RMS are substentially larger while on most other days, the own

fitting RMS is smaller.

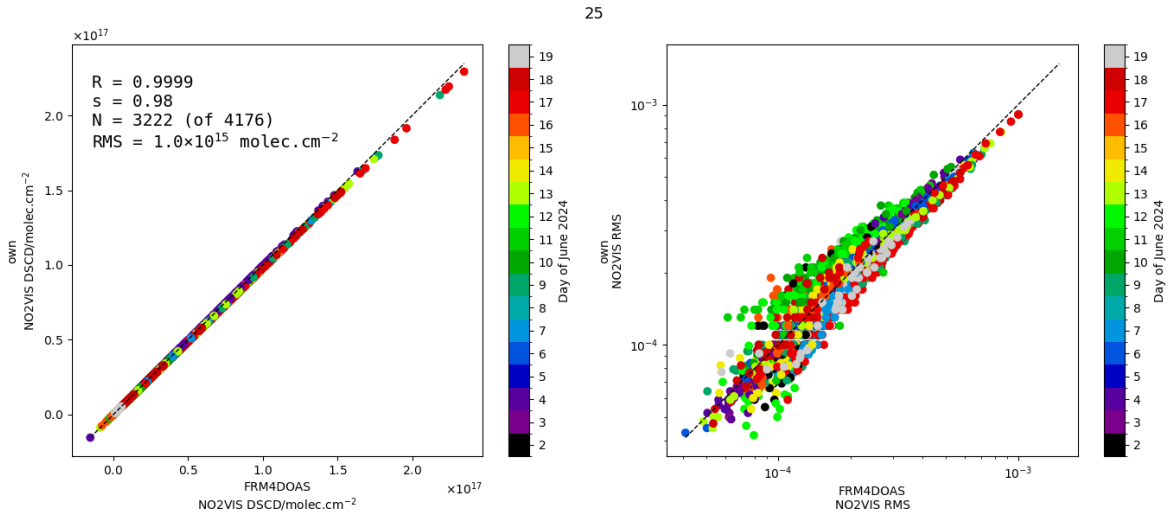

Figure 81: As Fig. 77 but for instrument #25

For #25 (compare Fig. 81) we see consistently slightly lower own DSCDs compared to FRM4DOAS DSCDs leading to a rather large RMS while showing very good correlation. Regarding the fitting RMS, there are some differences that seem to be only partly dependent on the day. No large systematic differences can be seen. In order to answer the question whether the FRM4DOAS Δ DSCD are too high or whether the own processing is too low, we need to compare first the two baselines, of own processing and FRM4DOAS processing to see if there is a systematic bias between the baselines. As evident in the left panel of Fig. 82, this seems not the case for NO2VIS. Next, we can check the processing for instrument #25, both own and FRM4DOAS processing, against the respective baselines. This is shown in the middle panel of Fig. 82. The slope and correlation for FRM4DOAS processing are slightly worse than for the own processing, likely indicating that the Δ DSCD from the FRM4DOAS processing are an overestimation.

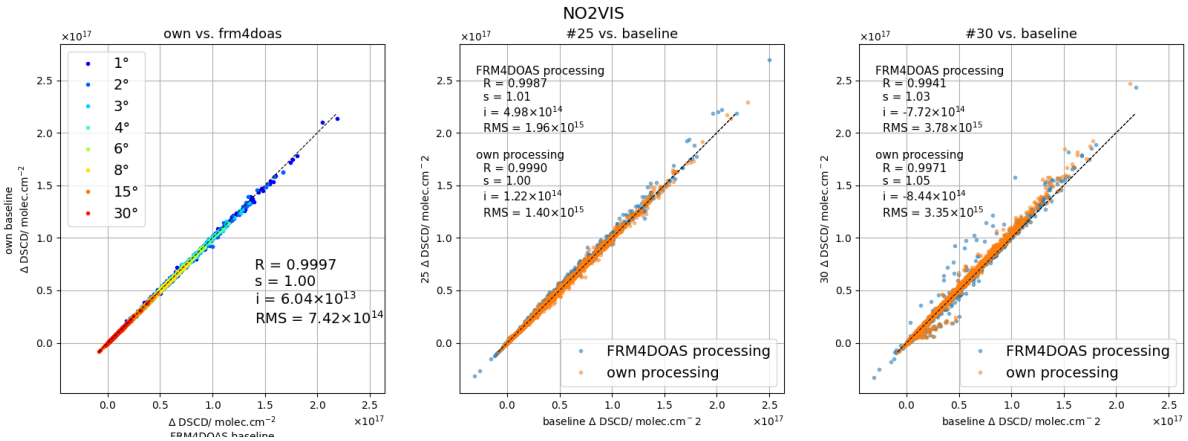

Figure 82: left: Median of own processing against the median of FRM4DOAS processing for NO2VIS. Middle panel and right panel: Comparison of own and FRM4DOAS processing against respective medians, middle for instrument #25 and right for instrument #30.

#28 (compare Fig. 83) shows some considerable scatter in the DSCD between own and FRM4DOAS seen in a lower correlation and a higher RMS. However, no systematic differences are seen, $s = 1$. The fitting RMS is systematically lower in the own processing. This is the second instrument that dropped out from either own or FRM4DOAS after 16 June. Given the considerable random differences between own and FRM4DOAS DSCDs, the drop-out of this instrument might also have contributed to the larger σ seen for own up to and including 16 June.

For #30 (compare Fig. 84) we see systematically higher DSCDs in the own processing compared to FRM4DOAS leading to a relatively larger RMS. The correlation however is excellent. For the fitting RMS, there are no huge systematic differences. Note that the fitting RMS does not vary greatly and hence the differences appear enhanced. As was already established, there is no systematic bias between the own and FRM4DOAS baseline, left panel of Fig. 82. Comparing Δ DSCD for #30 to the

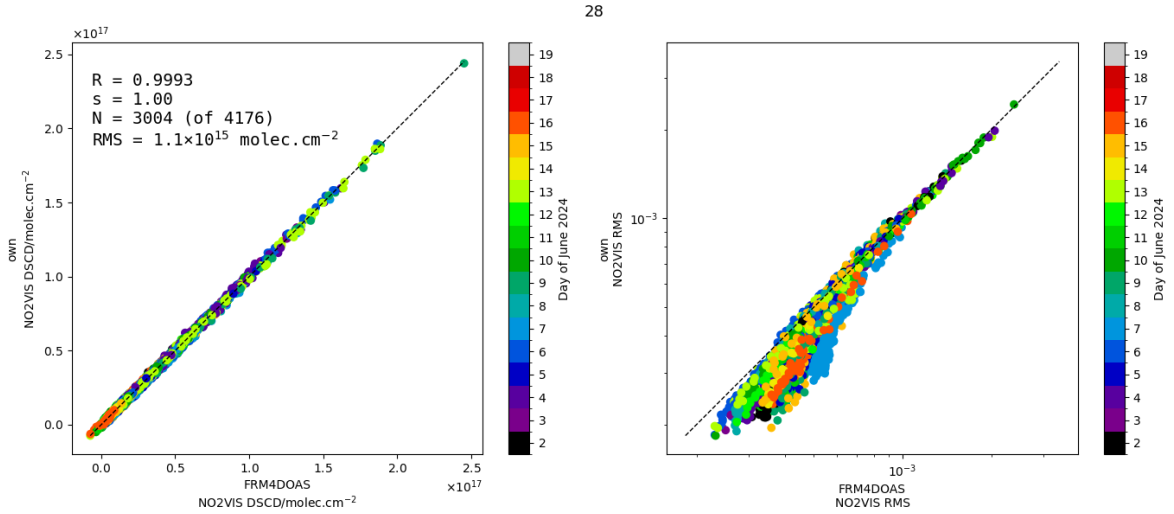

Figure 83: As Fig. 77 but for instrument #28

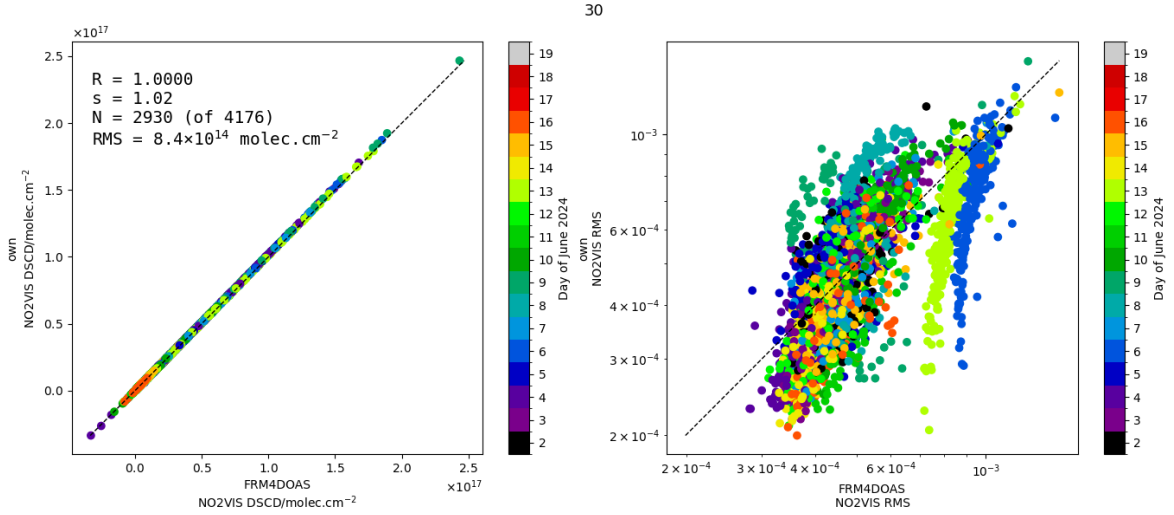

Figure 84: As Fig. 77 but for instrument #30

respective baselines for both processings (see right panel of Fig. 82) shows that #30 overestimates Δ DSCD in both cases, but less so in the FRM4DOAS processing, however on the cost of a lower correlation.

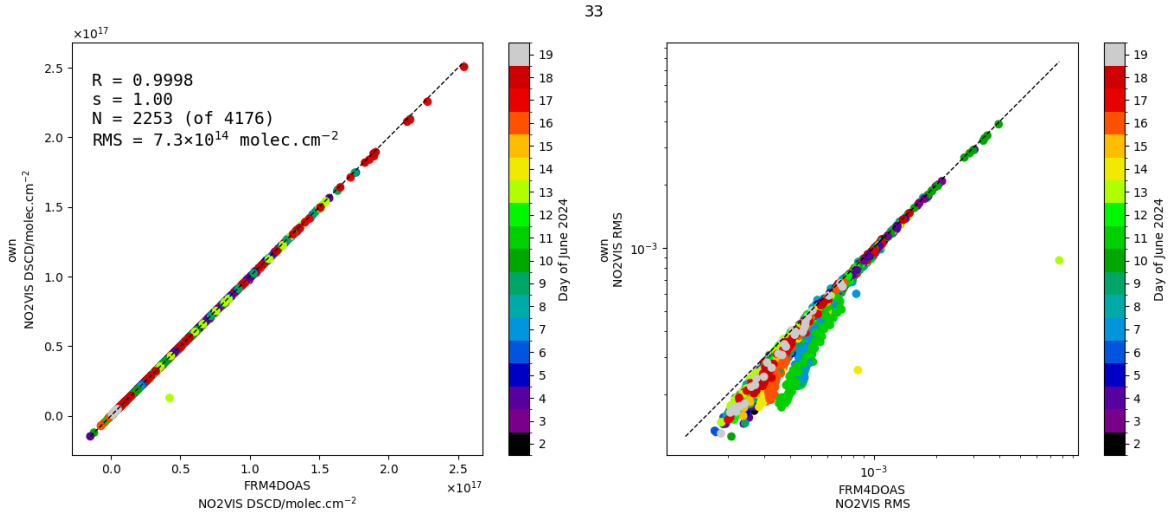

Figure 85: As Fig. 77 but for instrument #33

#33 (compare Fig. 86) shows good correlation and excellent slope and a somewhat larger RMS

between the own and FRM4DOAS DSCDs. This is mostly due to a single outlier. The fitting RMS are always slightly smaller for the own processing. Although this instrument was also one of the 3 dropping out after 16 June (either in the own or in the FRM4DOAS dataset), it is more unlikely that the drop in σ in the own dataset comes from this instrument.

As mentioned before, #3 and #21 were removed from the comparison shown in Fig. 76 for different reasons: #3 was removed since the DSCDs submitted at the lower elevation angles sometimes constituted an interpolation. In Fig. 86 we show the comparison of own and FRM4DOAS DSCDs and RMS. The larger scatter for some of the days comes from this interpolation.

The reason for excluding instrument #21 from the comparison of own and FRM4DOAS datasets was different. Here, the L1 spectra were changed after they were submitted to FRM4DOAS. The own DSCDs were produced from spectra that were straylight corrected, see Appendix D. This leads to the high slope seen between own and FRM4DOAS DSCDs. The stray-light consistutes a constant offset on the intensity of both the spectrum I and the reference spectrum I_0 which will hence result in an underestimation of DSCD. This is clearly visible in Fig. 87.

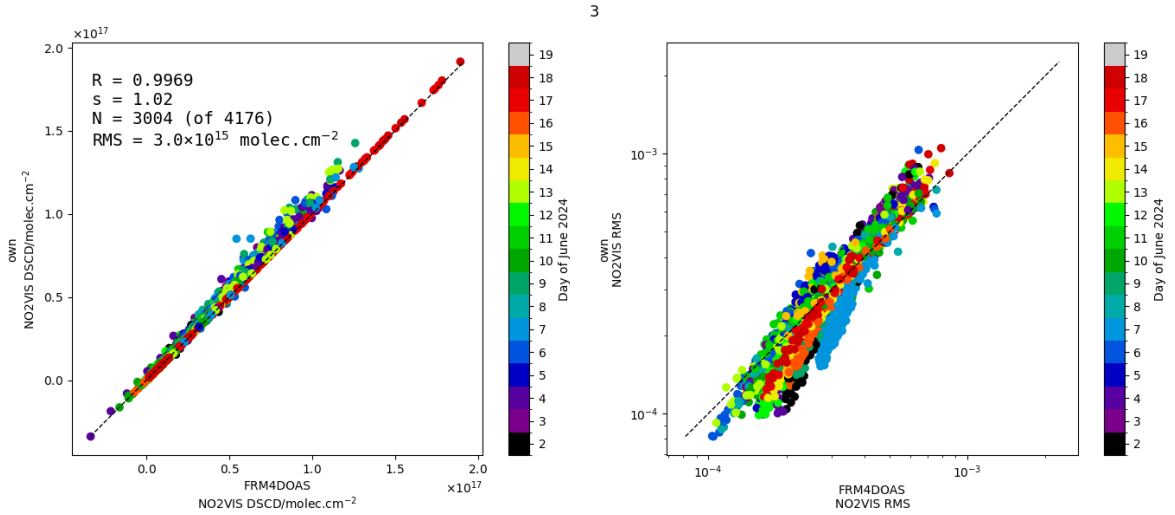

Figure 86: As Fig. 77 but for instrument #3

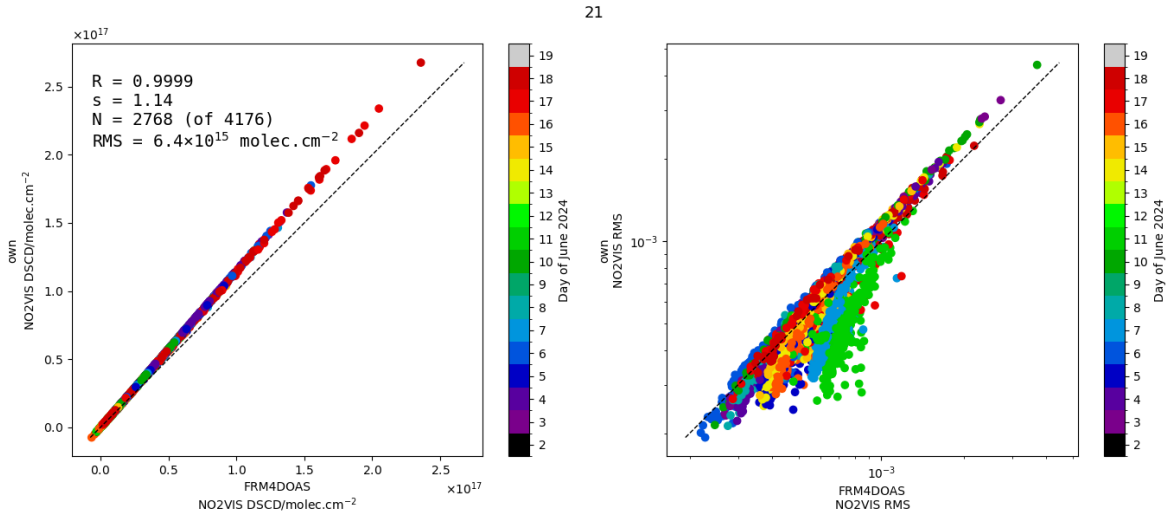

Figure 87: As Fig. 77 but for instrument #21. Note the large slope, indicating that the FRM4DOAS DSCDs are systematically underestimated due to starting from L1 spectra that were not stray-light corrected.

7.2 NO2VIS-SMALL_DAILYREF

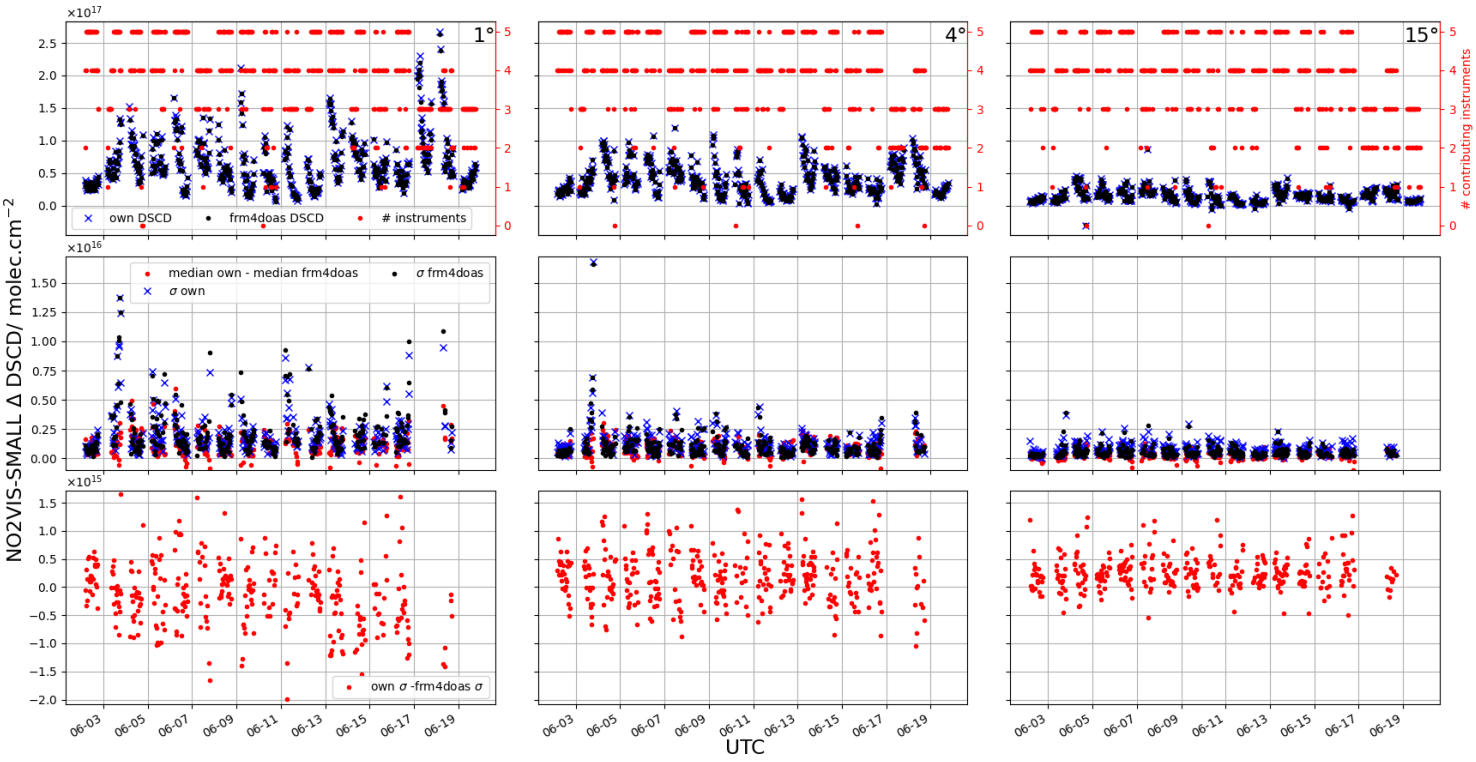

Figure 88: As Fig. 76 but for NO2VIS-SMALL_DAILYREF.

The total number of instruments for which NO2VIS-SMALl was processed within FRM4DOAS is unfortunately small. This is due to the configuration settings which were kept as close as possible to the operational settings. In those, NO2VIS-SMAll is only processed in the UV channel for instruments that have no separate visible channel.

In Fig. 88 we show the comparison between the datasets "own" and FRM4DOAS for NO2VIS-SMALL. As opposed to NO2VIS, where the differences between the medians of both datasets was always smaller than the individual σ, this is not the case for NO2VIS-SMALL (compare middle panels in Fig. 88) where the two are often comparable.

It is also visible that the differences in σ are generally larger for NO2VIS-SMALL than for NO2VIS, but there are no large systematics for low VEA, sometimes σ of own is smaller than σ FRM4DOAS, and sometimes vice-versa. At larger VEA, see lower right panel in Fig. 88, the σ is mostly larger in the own processing, indicating larger inter-instrument differences.

7.3 NO2UV_DAILYREF

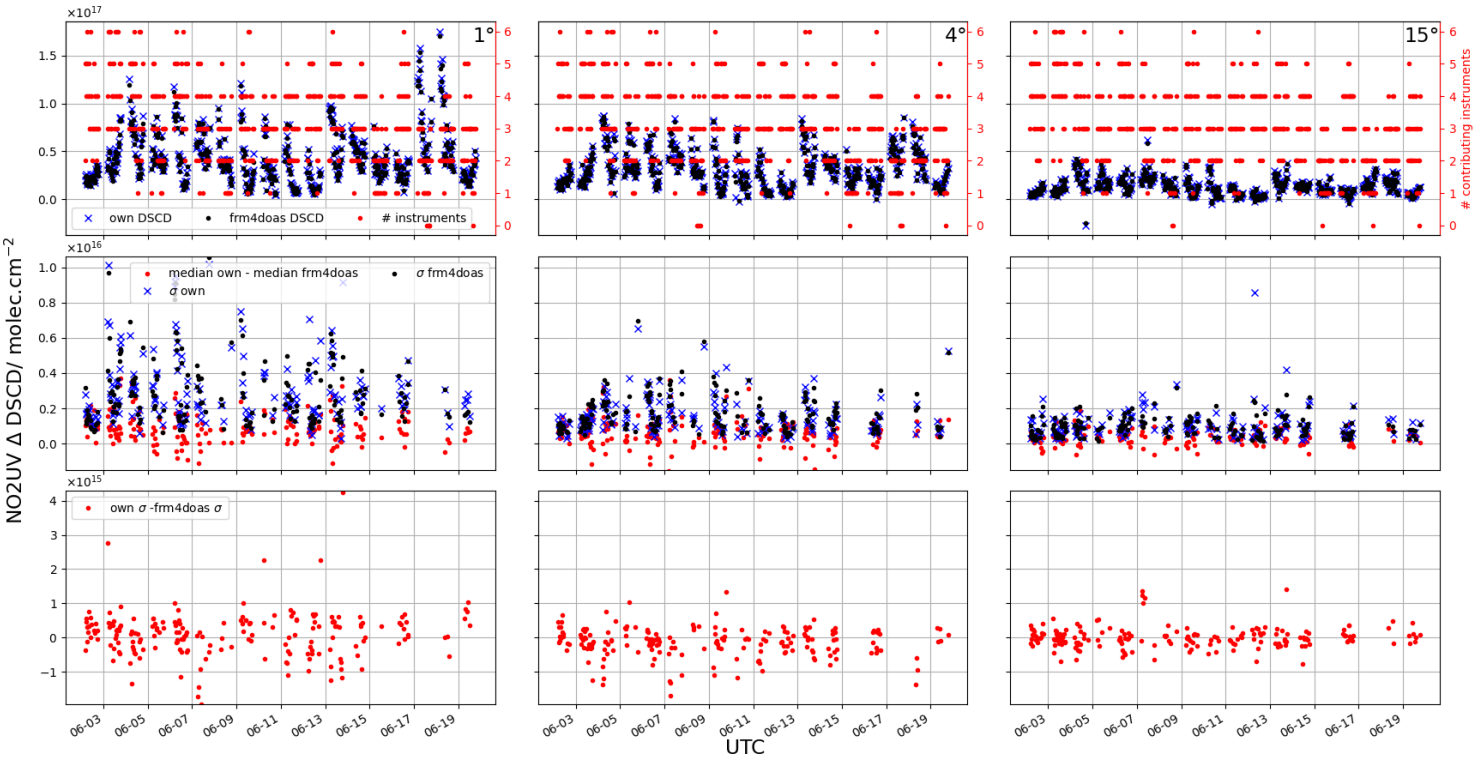

Figure 89: As Fig. 76 but for NO2UV_DAILYREF.

The total number of instruments for which NO2UV was processed within FRM4DOAS is also smaller than theoretically possible. This is attributable to the same reasons as for the low number of NO2VIS-SMALL: For those instruments where the UV channel allowed for NO2VIS-SMALL to be processed and that do not have a separate visible channel, no NO2UV processing is performed operationally.

In Fig. 89 we show the comparison between the datasets "own" and FRM4DOAS for NO2UV. In terms of comparisons of the difference between the medians and the respective σ of "own" and FRM4DOAS datasets (middle panel), NO2UV is somewhere between NO2VIS and NO2VIS-SMALL: The differences are most often smaller than the σ but not by much.

Comparing the σ of both datasets (third row in Fig. 89) show that there is no systematic difference between them, both datasets have a smaller and larger spread, varying from one time to the other. However, it is clear that there are some days where the overall σ in both datasets is larger than on others, and likewise, the difference between the two σ.

7.4 O4VIS_DAILYREF

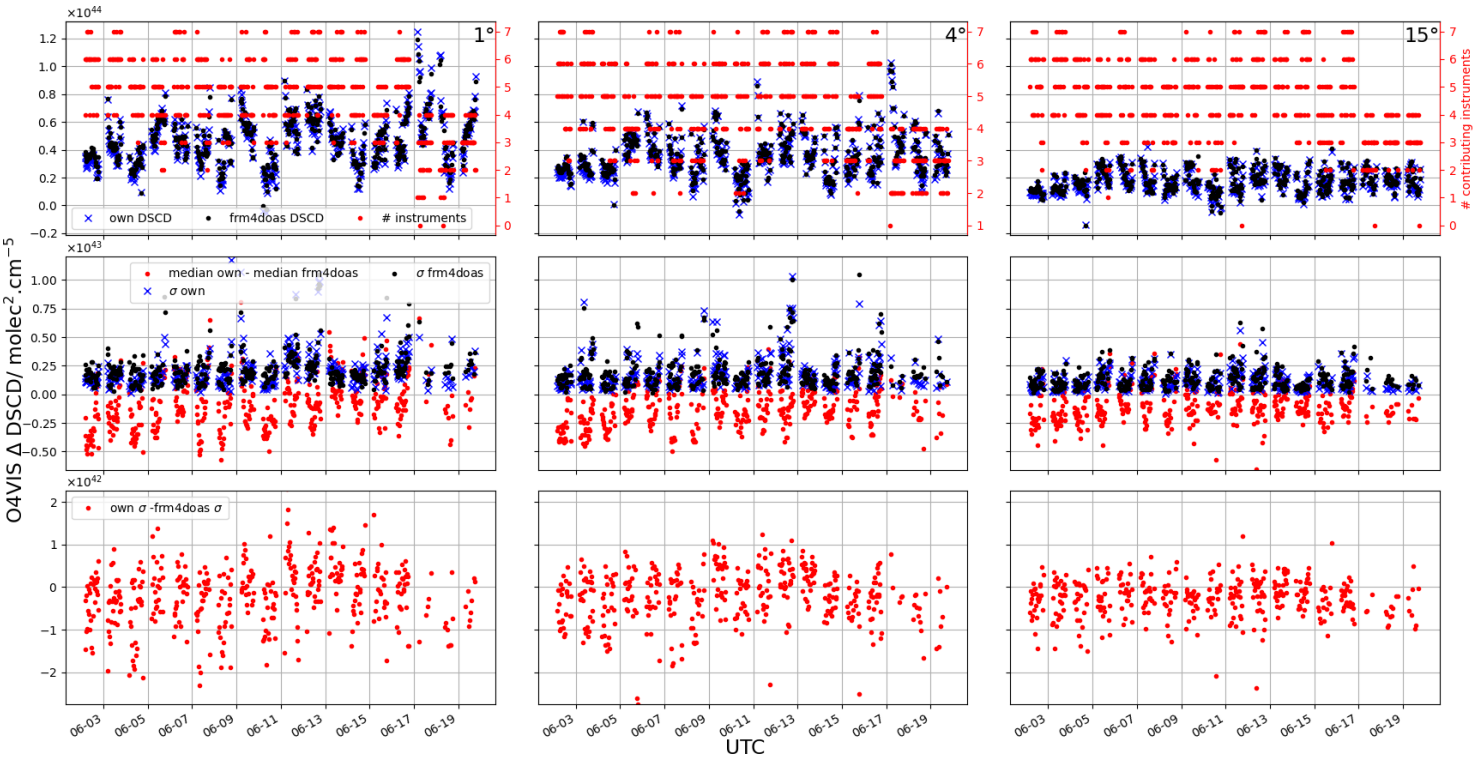

Figure 90: As Fig. 76 but for O4VIS_DAILYREF.

As previously mentioned, the wavelength window for O4VIS differ between the own processing ([425, 490] nm) and the FRM4DOAS processing ([450, 540] nm). Both Rayleigh scattering and scattering from normal (i.e. with positive Ångström exponent) is higher at lower wavelength, we expect more scattering (and hence lower DSCD) in the window at lower wavelength, so for the own processing. Comparing the magnitudes of DSCD between O4VIS and O4UV, this reasoning is confirmed. However, Fig. 90 and Fig. 164 to Fig. 170 indicate that for O4VIS, the own processing results in higher DSCD compared to FRM4DOAS only for the highest DSCDs. While the individual slopes comparing own (y-axis) to FRM4DOAS (x-axis) DSCDs (Fig. 164 to Fig. 170) are always either 1 or larger than 1, Fig. 90 shows that for small DSCD such as at high VEA, but also for low DSCD at low VEA, the effect is reversed.

Considering the middle panels of Fig. 90 shows that the absolute values of the differences between the respective medians for own and FRM4DOAS datasets is often larger than the individual σ. Further it can be seen that both are fairly large compared to the actual Δ DSCD. Comparing the σ of the own and FRM4DOAS datasets shows no systematic difference between the two.

7.5 O4UV_DAILYREF

As opposed to O4VIS, the wavelength window for O4UV is the same in the two datasets, own and FRM4DOAS. Since the same O_4 window is used in FRM4DOAS for NO2UV and NO2VIS-SMALL, the number of instruments for which the O4UV window was processed, is also larger, exceeding for some days 10 (c.f. Fig. 91).

From the same figure, the middle panels show that the absolute difference of the two medians is mostly smaller than the individual σ of the two datasets for both small and large VEA. The lower panels of the same figure indicate that the σ of the FRM4DOAS dataset is often slightly smaller than the σ of the own dataset, possibly pointing to differences in the processing, either because they were not specified or because the instructions were not strictly followed.

7.6 HCHO-WIDE_SEQREF

Fig. 92 shows the comparison of the own and FRM4DOAS dataset for HCHO-WIDE_SEQREF. After the CINDI-3 WS, the instructions about which reference to take as the reference for SEQREF were specified as using the reference after each scan. However, comparing the σ between the two datasets still shows that the FRM4DOAS centralised processing results in a smaller spread in DSCD for the different instruments compared to the own processing (lower panels).

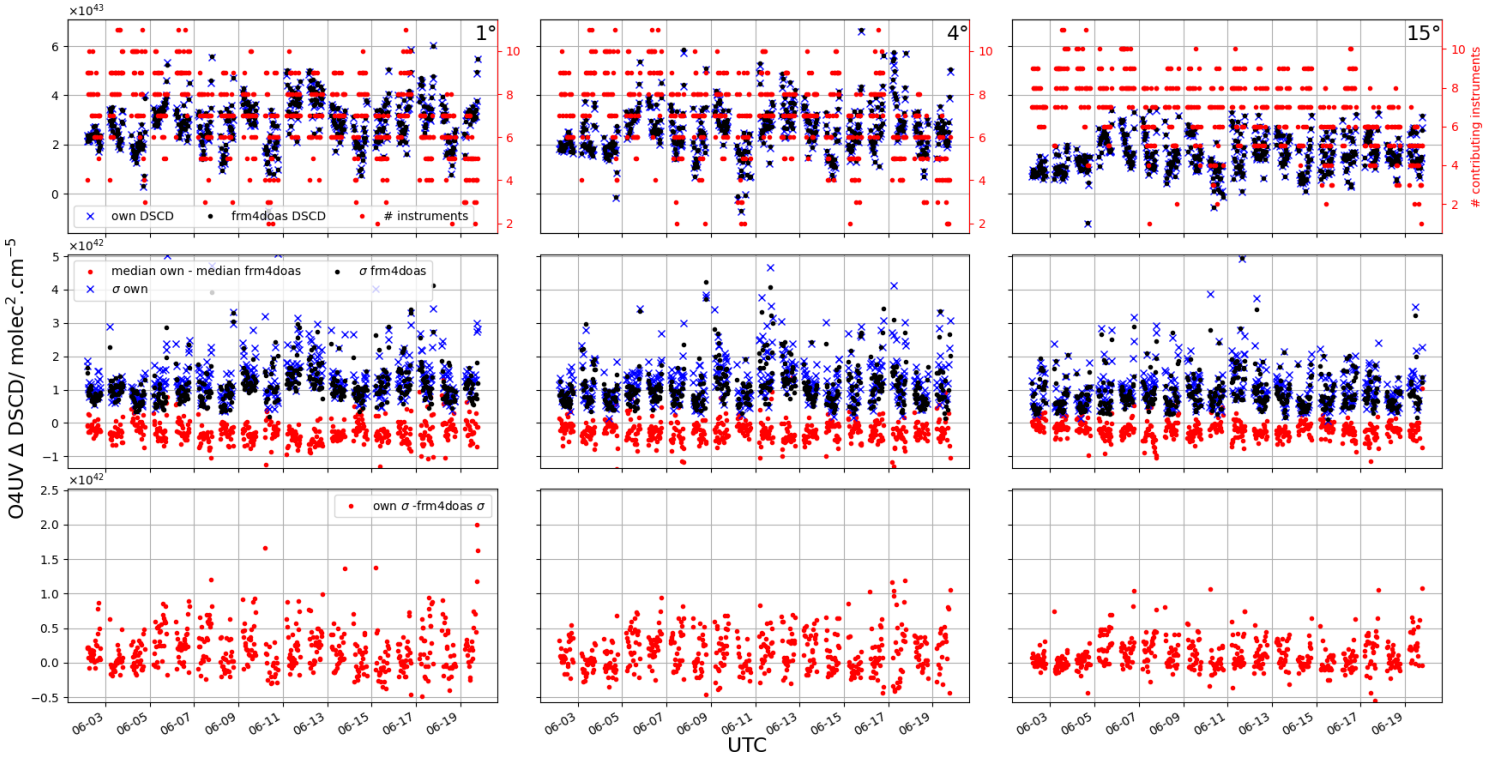

Figure 91: As Fig. 76 but for O4UV_DAILYREF.

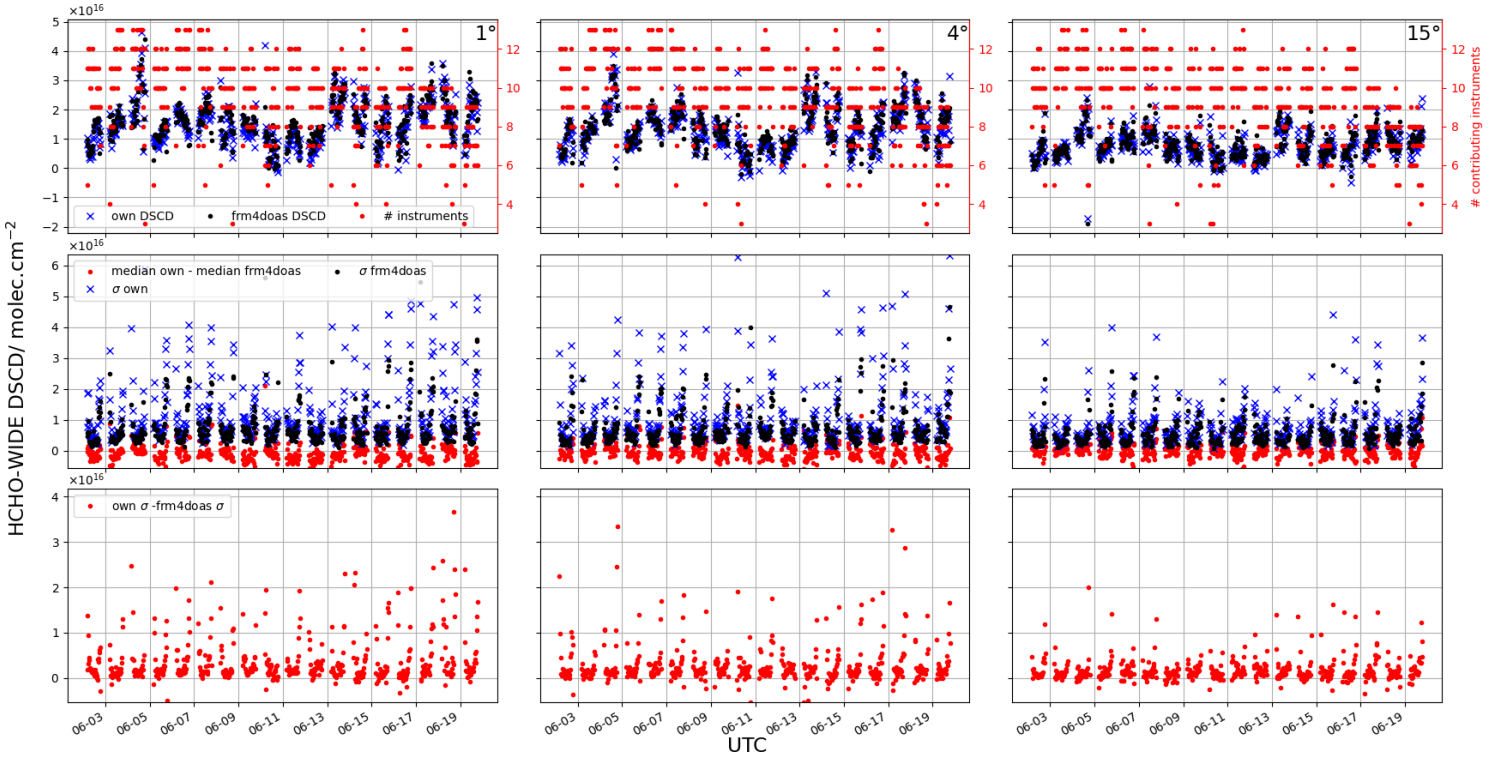

Figure 92: As Fig. 76 but for HCHO-WIDE_SEQREF.

Compared to the individual σ of the two datasets, the difference between the medians of the two is smaller (middle panels of Fig. 92). As previously found, the general inter-instrument differences are fairly large (compare the individual σ in the middle panels to the individual median DSCD (upper panels)).

We can now also make inter-instrument comparison plots as were performed in Subsect. 6.3.8. We do this separately for the own processing (see Fig. 94) and FRM4DOAS processing (see Fig. 95). However, we only take data into account that are present in both data sets, both in terms of instruments as a whole, as well as specific times of an instrument. This decision was taken to account for

the fact that obvious outliers (detected by high RMS) were filtered in the own processing, but are not filtered in the centralized FRM4DOAS processing at this step of the processing[8] This however also means that differences are to be expected for the correlation between the own measurements and the baseline from the own measurements as presented in Subsect. 6.3.9 and the ones in Fig. 94.

Comparing Fig. 94) and Fig. 95, it can be seen that not only the correlation with the respective baseline (shown on the diagonal) improved for all instruments, but also the inter-instrument correlation improved. Using all instruments of the own processing, it could be seen in Fig. 71 that only 2 instruments (#19 and #20) hat a correlation with the baseline at VEA ¿ 15° better than 0.9 for SEQREF. This is still the case if limiting the baseline calculation to the times and instruments at which FRM4DOAS data is available, see Fig. 94. Hence, there is no large selection effect. Considering the FRM4DOAS dataset, that number grows to 6, among which is #25 which had the best correlation with the baseline for HCHO but showed a large quality drop in HCHO-WIDE. However, since the centralized processed DSCD for this product show good correlation with the baseline, the problematic dates highlighted in Fig. 69 are due to bugs in the processing and not due to problematic spectra.

We show the full regression analysis for all instruments against the two baselines (first the two baselines against each other, upper left panel) in Fig. 93. While the correlation of the baselines is not great, it is still above 0.96. Although there is a non-negligible deviation of slope from 1 (FRM4DOAS below own DSCDs), the RMS is relatively small, 2.04×10^{15}. Regarding the instrument-baseline comparisons, it can be seen that for six instruments (#4, #7, #18, #25, #33, #34), the FRM4DOAS DSCDs are in considerable better agreement with the baseline. For further four instruments (#5, #14, #20, #28), the FRM4DOAS DSCDs are slightly better. For two instruments (#19, #21), the FRM4DOAS processing yields slightly lower agreement with the baseline. Only for 1 instrument, the own processing yields considerably better agreement (#24). These statements are based on correlation R and RMS.

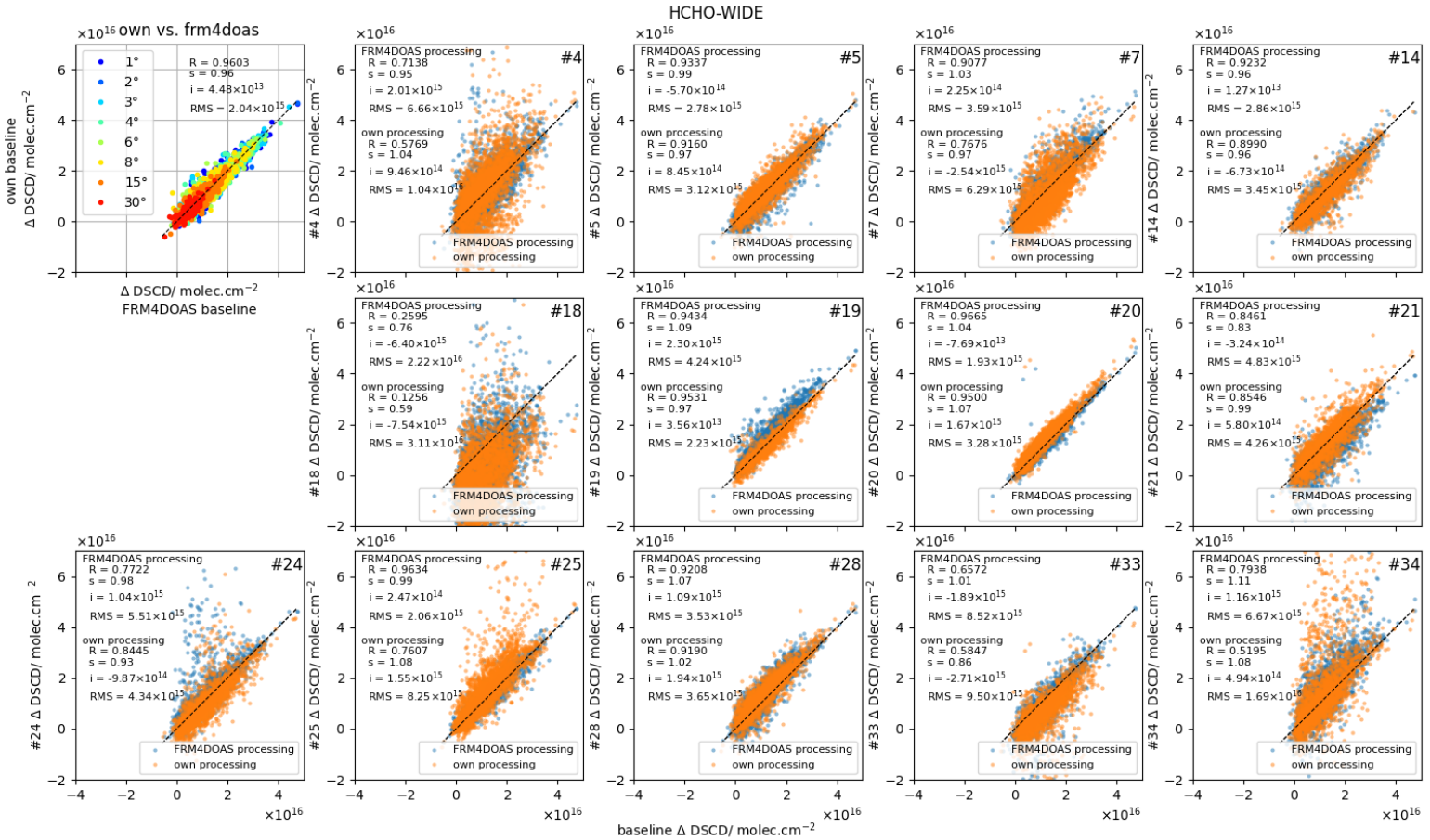

Figure 93: As Fig. 82 but for HCHO-WIDE

[8] In case of very large RMS, the corresponding DSCD fitting error will most likely also be high which will eventually result in a low Degrees of Freedom (DOF) in profile retrieval and hence it is likely that scans dominated by such values will be flagged as invalid.

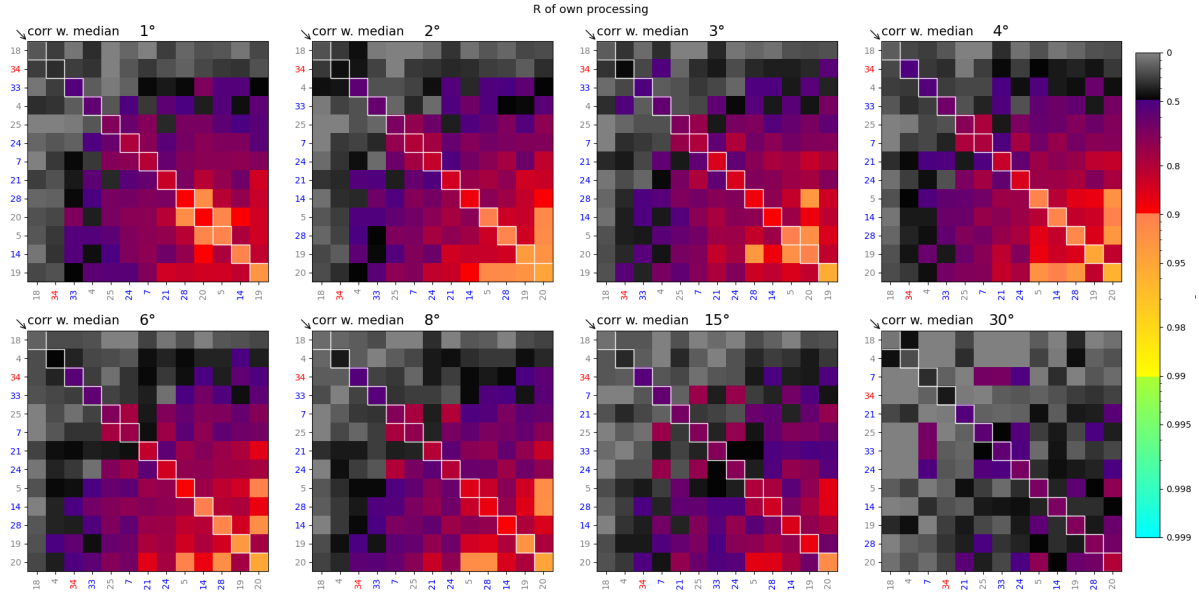

Figure 94: As Fig. 75 but limited to instruments that also have FRM4DOAS processing and limited to times where FRM4DOAS processing for that instrument is available. Note that this also affects the median, which is only calculated for times when at least 70% of the instruments (i.e. 9 here) contribute.

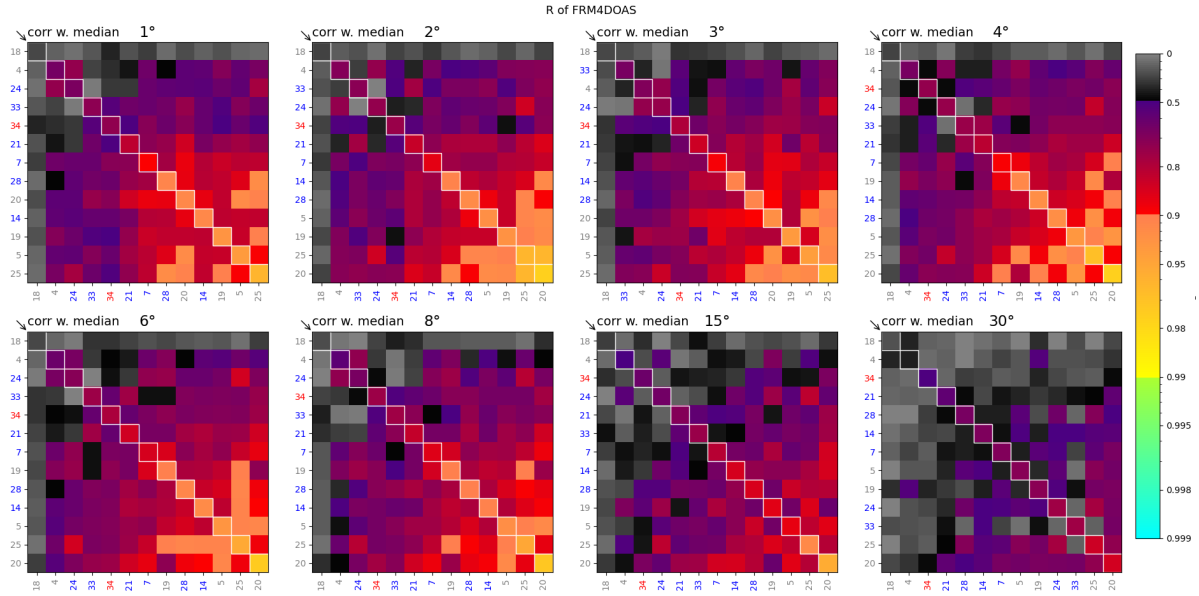

Figure 95: Same as Fig. 94 but for the FRM4DOAS dataset.

7.7 Final considerations

For most products, the spread around the median is very comparable whether considering own processing or centralised processing. As instrument #21 shows, since the centralized processing starts off the corrected L1 spectra, if no correction of spectra were performed although a high offset (or other effect such as stray-light that acts as a type of offset) is present, then the fitted DSCD will suffer the effects.

However, this exercise also shows that e.g. for HCHO-WIDE, even though instructions for retrieval settings were detailed, a central processing is still beneficial. Compared to the respective baseline, 10 instruments showed an improvement and only 3 a worsening.

8 Summary and conclusions

Most subsection had their own specific conclusions or final consideration section. Here, we summarize some specific comments to each instrument in Subsect. 8.1 and some comments on HCHO and HONO products that are not included in the comment for each instrument in Subsect. 8.2.

8.1 Comments to each instrument

In the comments to each instrument, we exclude HCHO, HCHO-WIDE and HONO when we refer to the performance of "all products". Instead, we add a separate section on those instruments in Subsect. 8.2

#1 This instrument delivered data for all UV products and O3VIS. The horizon scans indicate a fairly unstable pointing, however, as for all SkySpec instruments, this could be due to the recorded angle not corresponding to the real angle since the instrument pointing position will not move unless the difference in angle is larger than the angle uncertainty. Considering the median of horizon scans, they indicate that there is a general small offset of some fraction of a degree, such that the 1° VEA corresponds rather to real 1.4° VEA. In the elevation scan analysis, this might explain the generally seen slightly low slope for small elevation angles that might be responsible that #1 is among the worse ones for NO2UV, O4UV and O3UV, especially for low VEA. Since this viewing offset is not constant but has changes of more than 0.5° from day to day, this also affects the correlation. However, it does not explain the large scatter also seen at larger VEA for O3UV (much worse correlation than for O3VIS), hence it is likely not the only reason for the not satisfactory performance for this product. The high SZA zenith-only comparison suggests that in order to measure under low light conditions, such as above SZA 89°, a lower temperature for the detector stabilization might be recommendable. Considering the consistence between different reference types, there seem to be a small systematic difference between FIXREF and DAILYREF, while differences with SEQREF are not systematic and larger for most products.

#2 This instrument delivered data for all products except for NO2VIS-SMALL. The horizon scans indicate a smaller pointing instability than seen for #1, however, there is still a considerable day-to-day change. There is some systematic difference between up-ward and down-ward scanning; this might indicate that this instrument suffers from smaller backlash issues. It is hence recommended that scans are always performed in up-ward scanning during operational use. As instrument #1, this instrument seems not very well suited for measurements under very low light conditions. Daytime measurements of NO_2 and O3VIS are good; for O_4 they are acceptable. O3UV on the other hand show dissatisfactory results. They show very low correlation with the baseline. Since instruments #1 and #2 were processed by the same data provider, it is possible that the poorer performance for O3UV might be rather due to a bug in the O3UV fitting than to an issue in the spectra. The inconsistencies between the different reference types are of similar magnitude as for #1.

#3 This instrument delivered NO2VIS, NO2VIS-SMALL, HCHO and O4VIS. It suffered greatly from elevation pointing instability. Although the PI tried to minimize the effect of this by submitting interpolated DSCD for low VEAs, the effects of this pointing instability is obvious for small VEAs at 1° and 2°. Since no high SZA measurements were submitted, no statement about measurements at high SZAs can be made. The zenith analysis up to the SZA delivered (approximately 85°) however shows satisfactory results, however with a notably positive bias. There is a slight systematic inconsistency observed between the reference types for NO2VIS and NO2VIS-SMALL. Although very small, it might be worth investigating this systematics more. Apart from the pointing related issues at small VEAs, the performance for NO_2 and O_4 is acceptable.

#4 This instrument delivered NO2VIS, NO2VIS-SMALL, HCHO, HCHO-WIDE and O4VIS. The horizon scans indicate a spread in pointing comparable to that of #2. For the zenith measurements, data was provided up to slightly higher elevation angles than for #3, approximately up to 90°. As for #3, there is a notable positive bias, overall the performance is marginally acceptable. For the elevation scan analysis, there is a considerable spread in agreement with the nominal measurement time, especially for the zenith measurements. Regarding the consistence between the different reference types, there is a fairly large spread which might be worth investigating. For NO_2 products, the performance is satisfactory, while the correlation for O4VIS is somewhat worse, especially for the DAILYREF reference type.

#5 This instrument delivers all products except for NO2VIS, O4VIS and O3VIS. It suffers from a known and consistent backlash affecting the downward scans, however, if scans are always performed in up-ward direction, this is not an issue. Except for an adjustment of the pointing after 8 June, the pointing accuracy is very high. NO2UV is the only UV product that was analysed at medium to high SZAs. Compared to the performance of other instruments for this product, the performance of #5 is good, even though in absolute terms only acceptable. From the SZA color-coded regression analysis plots it is visible that the performance at very high SZA starts to get a bit worse. The consistence between different reference types is very good. For all delivered products, for the elevation scan analysis, it is among the best performing instruments.

#6 This instrument delivers all products. Compared to other SkySpec instruments, the pointing stability during the first half of the campaign was very good, but a bit worse during the second half of the campaign. No notable differences between up-and down-ward scanning were detected. The data coverage of this instrument is excellent. The intermediate to high SZA zenith performance for O3VIS is among the best. The remaining products are acceptable except for a large negative offset for O_4. Although #6 has no cooled detector and is stabilized at 20°, it seems not to suffer too much from large thermal noise. The consistence between different reference types is very good. Regarding the elevation scan analysis, there is a notable positive bias for NO2VIS which is less prominent in NO2VIS-SMALL and not present in NO2UV. The other quality identifiers are very good (except for at large VEAs, which is attributable to the overall small median DSCD used to normalize the intercept and RMS). The performance for O_3 and O4UV is also very satisfactory. Only O4VIS is an exception, where the performance is a bit worse in terms of correlation.

#7 As instrument #5, this instrument delivered all products except for NO2VIS, O4VIS and O3VIS. In terms of pointing stability, it showed mostly a good stability during the campaign and fairly consistent for up-ward and down-ward scans. Regarding the NO2UV zenith DSCD products, it is visible that the performance gets worse with increasing SZA. Even limiting the SZA to below 85°, still results in an RMS below the threshold for acceptable values. The agreement between different reference types is good and shows no systematics and overall very small differences. The performance for elevation scan products of NO_2, O_3 and O_4 is very good except for a notable positive bias in O3UV that also affects the RMS.

#8 As instruments #9 and #10, this only delivers NO2VIS and O3VIS at 90° VEA for DAILYREF and FIXREF. As those instruments, due to the lower frequency of measurements (approximately every 2 min), there are larger average time differences between the observation time and the nominal time, and hence between the measurement time of these instruments and other instruments. Due to the different approach for the zenith only measurements at intermediate to high SZA, namely interpolating at fixed SZA, these are less important for the zenith only investigations compared to the elevation scan analysis in which these instruments were nevertheless included. For the former, for NO2VIS, there is a notable negative bias which also affects the RMS. For O3VIS, the performance is good. There is no obvious negative effect on the products at high SZA. The same bias for NO2VIS is also seen in the measurements throughout the day that were analyzed together with the elevation scans. No such bias is seen for O3VIS, and the performance is good.

#9 As instruments #8 and #10, this only delivers NO2VIS and O3VIS at 90° VEA for DAILYREF and FIXREF. For the intermediate to high SZA zenith inter-comparison, #9 performs a bit worse for O3VIS compared to #8, but better for NO2VIS without the negative bias that was present for #8. Considering the comparison throughout the day, from the elevation scan analysis: the correlation and RMS for O3VIS are slightly worse than for #8, but no systematic bias is observed, however a positive offset is visible for DAILYREF. There is however a distinct difference between DAILYREF and FIXREF, the latter having better correlation and RMS. This might be due to a random time shift of the DAILYREF reference. For NO2VIS throughout the day, there is a negative bias observed that is slightly larger than for #8 if considering DAILYREF. In general it can be seen that instruments #8, #9 and #10 show a better performance during the day with FIXREF. This might be related to a larger random time-shift of the reference spectra or a different approach to average the spectra taken during the reference time slot.

#10 As instruments #8 and #9, this only delivers NO2VIS and O3VIS at 90° VEA for DAILYREF and FIXREF. Regarding the intermediate to high SZA zenith comparison: The performance for NO2VIS and O3VIS is comparable and overall very satisfactory, only the intercept for O3VIS FIXREF is a bit worse.

#14 As instruments #5 and #7, this instrument delivered all products except for NO2VIS, O4VIS and O3VIS. The pointing stability during the campaign was good, but the horizon was likely seen at a slightly too low angle, such that a nominal angle of 1° might correspond to approximately 1.3° in reality. There are no backlash issues observed. The NO2UV zenith comparison at intermediate to high SZAs shows a slightly larger scatter for higher SZA (larger than 90°) and overall a positive bias. However, the results are still in the acceptable region. The consistence between different reference types is very good. Regarding the elevation scan analysis, the same positive bias is seen in NO2UV. It is absent in NO2VIS-SMALL for which the performance is very good. The positive bias in NO2UV also affects the RMS, although the correlation is as good as for NO2VIS-SMALL. A similar but somewhat smaller bias is also visible in O4UV, however the correlation remains very good. For O3UV, the bias is much smaller and the correlation excellent.

#15 This instrument delivered all products except for NO2VIS, NO2VIS-SMALL, O4VIS and O3VIS. The horizon scans indicate a small pointing instability (except for 6 June which is an outlier) but a clear difference between up-wards and down-ward scans. It is recommended that during operational use, only upwards scans are performed. The intermediate and high SZA zenith comparisons suggest that this instrument is not suited for measurements under low light conditions. Measurements above approximately 85° SZA showed no agreement with the baseline. Regarding the time compliance with the nominal measurement times, systematic (this instrument measured earlier) and larger random deviations were observed. This could potentially contribute to a worse agreement to the baseline DSCDs generally observed. There is a sub-optimal agreement between the three different reference types. There is no large systematic offset or bias detected for NO2UV, O3VIS and O4UV. Regarding the correlation, there is a better correlation for FIXREF than for the other two reference types. The performance for NO2UV, O3UV and O4UV is acceptable.

#16 This instrument only delivered data at NO2UV, O3UV and O4UV. Except for a single down scan horizon scan, only upward scans were delivered. These scans indicate a rather unstable pointing with large day-to-day variations. Most DSCD at SZAs larger than approximately 85° were filtered out. Therefore, the regression analysis results are of limited use, however indicate a rather poor performance in RMS and correlation. Overall, the number of submitted data is very low. There is a consistent small (-10 s) time offset from the nominal measurement times for all VEA. While DAILYREF and FIXREF seem very consistent, the scatter when comparing Δ $DSCD_{FIX/DAILY}$ to $DSCD_{SEQ}$ is a bit larger. The performance for O3UV, O4UV and NO2UV is among the worse ones. For O3UV, there is a notable negative bias. This is less pronounced for NO2UV and absent for O4UV. Considering the fitting RMS, it is among the greatest ones among the instruments. It is unclear why, since the nominal spectral resolution is very high (0.35°) and the instrument has a cooled detector (-40° C). It might be recommended to let the spectra be analysed by a group with more experience to figure out whether the poor performance is due to a sub-optimal data processing or due to the quality of the spectra.

#17 As instrument #15, this instrument delivers all products except for NO2VIS, NO2VIS-SMALL, O3VIS and O4VIS. No downward horizon scans were submitted. The upwards scans indicate a poor pointing stability. Furthermore, the horizon seem to be seen on average a bit too high, meaning that the 1° VEA corresponds rather to 0.8°. The intermediate to high SZA zenith DSCD comparison indicates that this instrument is not suited for measurements at SZA greater than approximately 85°. This might be related to the rather high temperature (25°) at which the detector is stabilized. Apart from the large scatter at high SZA, a clear positive bias is also seen. In terms of time compliance, there is a somewhat larger spread of time agreement. For NO2UV and O4UV a positive bias and negative offset as well as a relatively low correlation are observed. For O3UV FIXREF (only), the performance is good. The bad performance for DAILYREF is solely due to outliers from two single scans at 18:39 – 18:39 on 4 June and 5 June, which might even simply be a formatting error in the submitted files.

#18 This instruments delivered all products except for HONO. The horizon scan analysis indicates a large day-to-day variation in the vertical pointing, no differences between up- and down- scans and generally a small offset such that the 1° results might really correspond rather to about 1.3°. In general, the performance in UV is not satisfactory for NO_2 and O_4. For NO2UV, this is both seen in the zenith only analysis as well as in the elevation scan analysis: The correlation is generally very low, the RMS high. The same is seen for O4UV in the elevation scan analysis. Regarding the intermediate to high SZA comparison, the visible products look better, but it is obvious that the quality decreases sharply at about 88° SZA. The elevation scan analysis shows that for the visible products, although there is generally a good correlation, for low elevation angles, there is a systematic

negative bias. While this fits together with observed slightly too high pointing, it does not really explain the difference still seen at intermediate elevation angles at 8° and 15°. The agreement between different reference types is not systematic and small.

#19 This instrument delivers all products. However, since it switches between the UV and visible channel, each channel only measures half the time. The horizon scan indicate a very stable and accurate pointing. The intermediate to high SZA comparison shows acceptable results for all products, it belongs to the best performing instruments. The consistence between different reference types is good. Regarding the elevation scan analysis: The performance of this instrument is good for all products and among the best ones.

#20 As instrument #19 and #6, this delivers all products. The intermediate to high SZA zenith only comparison shows good performance, except for the NO2UV product, where the performance is only acceptable in absolute terms, however also here it belongs to the best performing instruments. The horizon scans indicate a very stable and accurate pointing. Regarding the elevation scan analysis, it shows good performance for all products. It should be noted however, that the FIXREF is often slightly worse than the DAILYREF. This can be pinned down to measurements on 17 June, probably indicating some instrumental change on that day.

#21 This instrument submitted all products and showed some instability in the pointing, but a good median elevation and no systematic difference between up- and down- scans. Regarding the intermediate to high SZA comparison: O3VIS and NO2VIS show a good performance even at high SZA. A contributing factor to the good results might be the stabilization of the detector to 10°C. For NO2UV, a slightly worsening above around 91° can be observed, but is still in the acceptable range. For O4VIS, there is considerable difference between DAILYREF and FIXREF, where the former shows less scatter. In the visible, there is a systematic difference seen between DAILY/FIXREF compared to SEQREF. The general performance considering the elevation angle comparison is good throughout all products. Since a stray-light correction to the L1 spectra were performed after submission to the centralized processing facility, it is recommended to submit the corrected spectra.

#22 This instrument is an imaging instrument that arrived after 6 June to the campaign and as such does not deliver FIXREF. It only delivers NO2VIS and O4VIS. The total number of delivered DSCD is small (e.g. for 1° DAILYREF, there are only 2 coincident measurements). For NO2VIS the performance is good, however there is a negative bias at higher elevation angles. For O4VIS, it is notable that the RMS is smaller for SEQREF than for DAILYREF, even in median baseline units. For O4VIS, there is both a positive bias (so opposite to NO2VIS) as well as a positive offset in the DAILYREF data, less visible in the SEQREF. The few points delivered at high SZA however indicate that there are no issues under low light conditions at high SZA for NO2VIS; for O4VIS, the positive bias already noted is equally visible.

#23 As #22, this instrument is an imaging instrument. It only delivered data for NO2VIS-SMALL at FIXREF for elevation angles up to 15°. While the correlation and hence RMS are not great, it is understood that the time coincidence is considerably worse due to the mode of measuring (1 complete image at a single wavelength per second and hence all VEA measured during 5 min). There is some systematic offset (positive for all VEA except for at 15° where it is negative), possibly due to pixel-angle mapping.

#24 As #18, this instrument submitted all products except HONO. The horizon scan analysis shows that while the median pointing is accurate, there is considerable change from day-to-day and unsystematic considerable difference between up-and down- scans. From the intermediate to high SZA zenith only comparison it is visible that there are issues at lower light conditions. DSCD up to about 87° show a good agreement for O3VIS and NO2VIS. The performance for O4VIS is acceptable while NO2UV shows large scatter also for lower SZA. The agreement between different reference types is mostly excellent except for some outliers for SEQREF. The 1° VEA shows often larger RMS across NO_2 and O_4 products, likely related to the pointing instability. For O3VIS, a moderate offset and RMS are observed for all reference types, while correlation remains high.

#25 This instrument delivered all products. The horizon scans indicate that there was an adjustment made after 14 June. The stability before and after is moderate. The pointing before this adjustment might have been slightly too high. For the moderate to high SZA zenith only comparisons, the performance for O3VIS is excellent. For other products it is acceptable or good. The agreement between different reference types is very good. The elevation scan comparison shows that

the performance for the NO_2 and $O4_2$ products as well as for O3VIS is very good. O3UV shows a slight negative bias that also affects the RMS but with a good correlation.

#27 As #22 and #23, this instrument is an imaging instrument. It delivered NO2VIS and O4VIS for all reference types. The zenith only comparison at moderate to high SZA shows a good correlation for NO2VIS but a fairly large positive bias . O4VIS does not show such a bias, but, as all instruments, a lower correlation. Both products are acceptable. There is relatively large disagreement between the different reference types, but not systematic between FIX/DAILY compared to SEQREF. However, some structural difference can be seen between DAILYREF and FIXREF. It might be possible that a different binning was used for the 90° when used as the reference. As previously mentioned, the FOV at higher VEA is large compared to other instruments. There is also a positive bias observed for NO2VIS in the higher elevation angles that is not seen in the lower elevation angles. The performance is overall acceptable for NO2VIS. For O4VIS, the bias at large angles is smaller, but at lower elevation angles, a slight negative bias is seen. The correlation for O4VIS and RMS is worse than average.

#28 This instrument covered all products. Except for 15 June, the instrument had a fairly good pointing stability. There is a fairly large disagreement between the different reference types DAILY/FIX and SEQREF in the visible (only), but non systematic. The intermediate to high SZA zenith only comparison shows for O3VIS very good agreement up to about 91° SZA after which a slight degradation can be observed, but the overall performance is still good for all (DAILYREF) or almost all (FIXREF) quality identifiers even up to 92°. NO2VIS and NO2UV on the other hand show a large positive bias and are hence outside the acceptable range. O4VIS is still acceptable. Regarding the elevation scan analysis, no positive bias for NO_2 products is observed and the performance is good. The performance for O_4 is good. For O3VIS, there is a positive offset that also affects the RMS, however the product is still acceptabe. O3UV does not show such and offset and the product is good.

#29 This instrument covered all products. There was a small adjustment in the vertical pointing after the first two days and the pointing stability is otherwise very good. There is a somewhat larger inconsistency between DAILYREF and SEQREF, less visible between FIXREF and SEQREF. The deviation from the nominal measurement time is consistently 20 s earlier. Regarding the intermediate to high SZA zenith only comparison, it is seen that both NO_2 products show a low bias, while O3VIS shows a slightly high bias. The O4VIS product is good, the O3VIS product is only acceptable. Close investigation in O3VIS FIXREF shows that there are two branches, both very well correlated but with a slightly different slope. This might indicate a minor change in the instrument. Regarding the elevation scan comparison, NO2VIS, O4VIS, NO2VIS-SMALL and O4UV products are good, NO2UV is slightly worse at higher elevation angles due to a larger negative offset. O3UV has a positive bias, as already seen in the zenith only comparison, but the overall performance is still acceptable. The bias seen for O3VIS is even larger for O3UV having a large effect on the RMS.

#30 This instrument only covered NO2VIS, O4VIS and O3VIS. Except for a pointing adjustment after 12 June, the pointing was very stable. The agreement between different reference types is very good. The intermediate to high SZA zenith only comparisons show excellent results for O3VIS. The other products are only acceptable. Regarding the elevation scan analysis, the NO2VIS and O3VIS product is good while the O4VIS product is a bit worse, but still acceptable.

#32 This instrument delivered all products, but for HCHO-WIDE and HONO, no SEQREF data was submitted. The reference type SEQREF shows considerable difference compared to FIX-/ DAILYREF. No DSCDs were submitted above approximately 85° SZA. While the instrument is included in the intermediate to high SZA zenith only comparison, the use of this is limited. However, it can be seen that the O3VIS product up to the intermediate SZA delivered is good, while the remaining products are only acceptable. The horizon scan analysis show large day-to-day variations in the elevation pointing, but no systematic offset. The NO_2 products are all good except for NO2VIS-SMALL_SEQREF which shows an insufficient correlation and RMS at intermediate and high VEA. O_4 products are also good, except for SEQREF with is slightly worse but acceptable. O_3 products are good.

#33 This instrument delivered all products. The horizon scans show a moderately stable pointing that might be slightly too high. It shows good agreement between the different reference types. The intermediate to high SZA zenith only comparisons show that all products are at least acceptable and that O3VIS has all quality indicators good or better. For the elevation scan comparison, it shows good to very good performance for all products without any noteworthy offsets or bias.

#34 This instrument is one of three pandora instruments that participated. They all delivered all products. As opposed to the other two pandora instruments, it was operated by KNMI and used a noon spectrum from a different day for the wavelength calibration. As the other two pandora instruments, there were larger deviations from the nominal measurement times, the pandora instruments measuring earlier on average. As the other two pandora instruments, the pointing was very stable and the agreement between different reference types is good. The pointing of #34 was very accurate. From the intermediate to high SZA comparison it can be seen that the overall performance of #34 starts getting bad above a SZA of about 88°. This is visible for all products. However, the performance below 88° is still acceptable. The performance for all all products during the elevation scans is very good, especially for low VEA. It clearly profits from the very stable VEA pointing. For NO2UV at high and intermediate VEA the performance is slightly worse, but still good. There are often no or only very few zenith measurements before the scan, likely due to the time match criterion.

#35 Many things that were discussed for #34 are also valid for #35. Only differences are highlighted: Compared to #34, the pointing is a bit higher for #35 and also for #36. Considering DSCD at high SZA, a larger deviation from the baseline is only observed above SZA about 90°. However, compared to #34, the scatter at lower SZA is higher for NO2VIS. Except for NO2UV, the performance is acceptable. Regarding the elevation scan analysis, the performance is good for NO2VIS, O4VIS, NO2VIS-SMALL, O3VIS and O3UV but slightly worse for NO2UV and O4UV, even slightly worse than for #34. However, still acceptable. Differences are larger for DAILYREF than for FIXREF which might point to the issue being rather related to the less good time coincidence.

#36 Many things that were discussed for #34 and #35 are also valid for #36. Only differences are highlighted: At high SZA, it performs slightly better than #34 but worse than #35 for O4VIS, O3VIS and NO2VIS and much worse than both for NO2UV. The visible products are still acceptable. Regarding the elevation scan comparison, #36 performs equally good as #34 and #35 for NO2VIS, O4VIS, O3VIS, O3UV and NO2VIS-SMALL but worse than both for NO2UV, but still acceptable, especially at the lowest VEA where the agreement with the baseline is still good. For higher VEA there is however a systematic low bias in the DAILYREF. O4UV is better than for #35 but slightly worse than #34 but still good.

#39 This instrument delivers all products except for NO2VIS-SMALL. The horizon scan analysis shows a fairly large day-to-day variation in pointing around the correct elevation. There are no high SZA measurements delivered above approximately 85°, hence the results from the zenith-only comparison are of limited use. Up to the delivered SZA, the performance for O3VIS is excellent for all (FIXREF) or almost all (DAILYREF) quality identifiers. The other products are acceptable. There are small but systematic difference between all reference types. Regarding the elevation scan comparison, a high bias for low VEA is observed for NO2VIS and O4VIS and a lower correlation, leading to fairly large RMS which makes this instrument one of the worse for NO2VIS. A smaller negative bias for low VEA is observed for NO2UV together with a lower correlation in DAILYREF and SEQREF, but slightly better for FIXREF. O4UV shows a lower correlation for SEQREF and DAILYREF, and again better for FIXREF. Only a very small negative bias is seen. O4UV and NO2UV are still acceptable. The O_3 products show a good correlation but O3VIS shows a slight negative offset for low VEA and O3UV a slight negative bias for all VEA. Nevertheless, these products are still acceptable.

#40 This instrument delivered all NO_2 and O_4 products and O3VIS. Unfortunately, horizon scans were only submitted during the second half of the campaign. They indicate a moderate pointing stability. Depending on the VEA, there was a fairly large difference between the actual measurement time and the nominal time. This difference is larger for smaller VEA and might contribute to a worse correlation seen. DSCDs were only submitted for SZA smaller approximately 85°, hence the information from the intermediate to high SZA zenith only analysis is only of limited use and shows that only the O3VIS and NO2VIS products are acceptable. Regarding the elevation scan analysis: For NO2VIS and NO2VIS-SMALL, there is a small negative bias seen for high VEA and a lower correlation, both contributing to a higher RMS, Nevertheless, the product is still acceptable. For NO2UV, a general low correlation, high positive offset and high negative bias, especially for high VEA is observed, leading to a high RMS. While lower VEA are still somewhat acceptable, high VEA are not. For O4VIS and O3VIS, it is among the worst performing instruments, especially at high VEA. The latter comes mostly from a low bias and a positive offset, and, compared to other instruments, a lower correlation. The former from a general lower correlation and a low bias. Since often the low VEA perform better than the high VEA, the worse correlation cannot really be explained by the worse time match since that is better for high VEA.

8.2 HONO, HCHO and HCHO-WIDE

It could be seen that the overall agreement for HCHO, HCHO-WIDE and HONO was not satisfactory. Much of this dissatisfactory results can be attributed to the fact that that there was not much HCHO and HONO in the lower troposphere during the campaign. However, the comparison study for HCHO-WIDE with the data from the centralized processing of FRM4DOAS showed that a part of it is still due to sub-optimal data processing and fit settings. It might hence be advisable for the participants to review their retrieval settings.

For HONO it could be seen that concentrating on days with some observable HONO increases the agreement at least for the lowest VEA. It had been interesting to have a similar centralized processing as was available for HCHO-WIDE to clarify whether the consistence can be improved on by harmonizing the fitting approach.

A Correlation scatter plots for all product-reference type combination for elevation scan analysis

Here we present the complete set of off-axis DSCD scatter plots of individual instruments (y-axis) against the baseline (x-axis). Unless otherwise stated, the baseline is the median as calculated outlined in Sect. 6.

Note that while all axes have an equal aspect ratio, only the plots for NO_2 and O_4 products have equal axis limits for both x- and y- axis. We opted for having the same axis limits for all instruments and tried to keep the same limits for a certain product accross different reference types. However, especially for O_3, this would have rendered the SEQREF plots unreadable. Especially for HCHO and HONO, several outliers were present for a number of instruments which are cut-off from the axes. This is noted in the Figure captions.

A.1 NO2VIS product

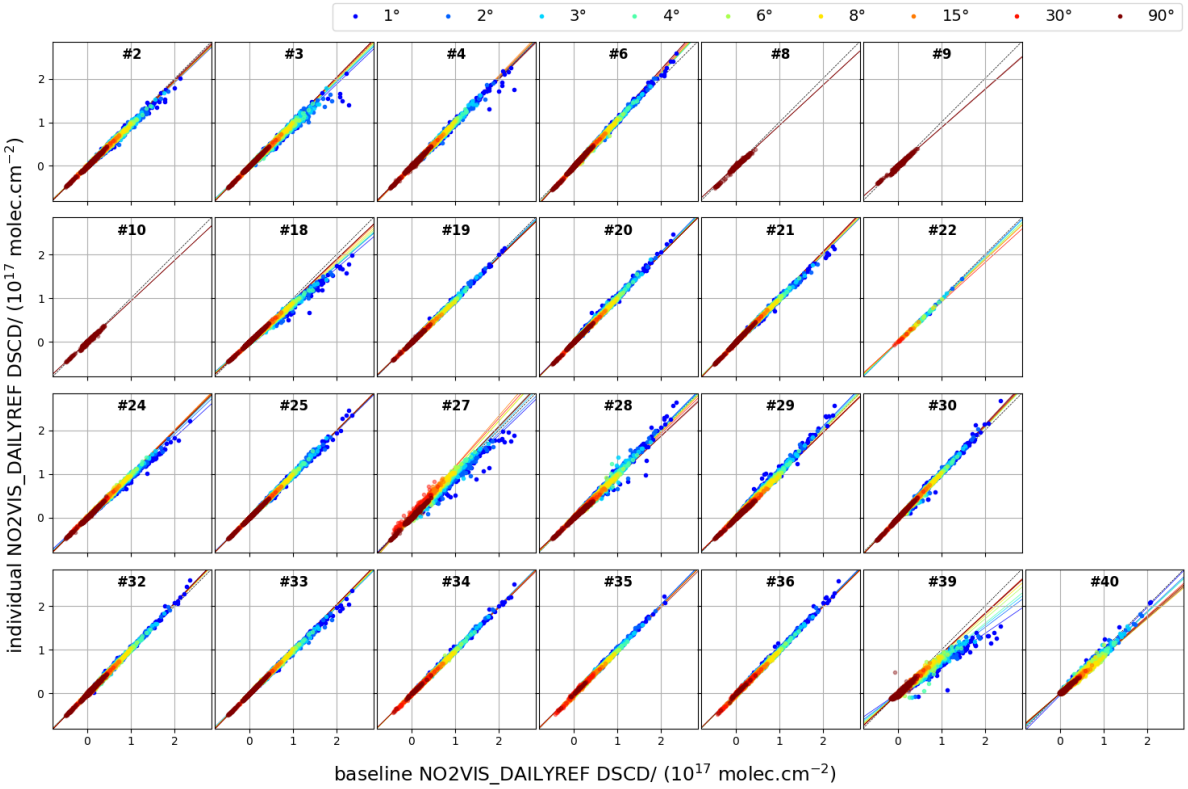

Figure 96: As Fig. 49 but for NO2VIS_DAILYREF

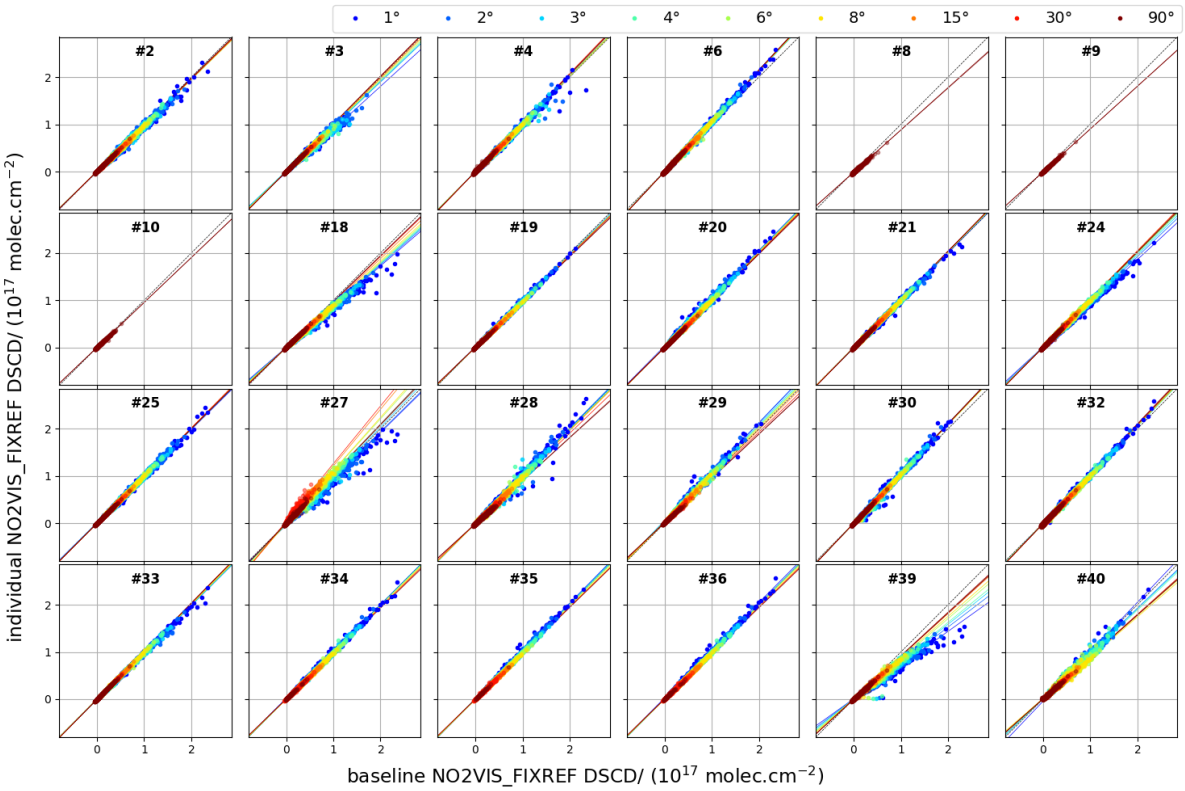

Figure 97: As Fig. 49 but for NO2VIS_FIXREF

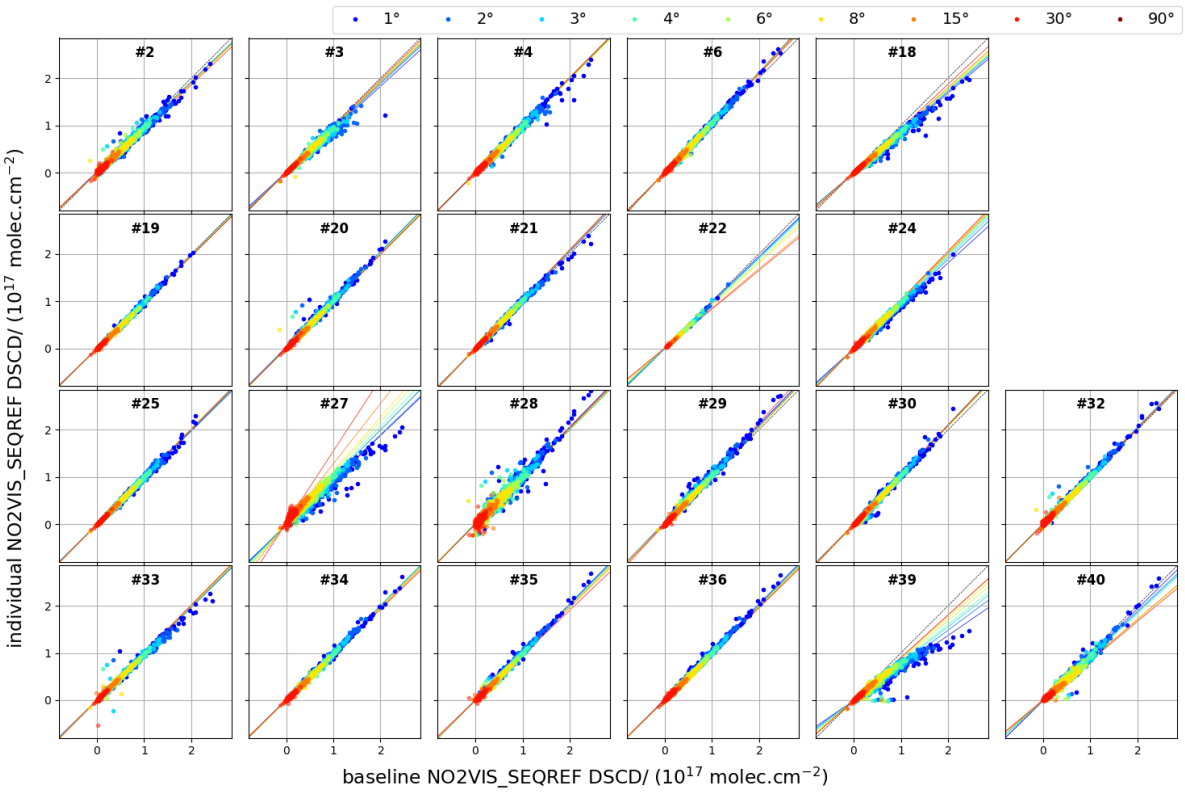

Figure 98: As Fig. 49 but for NO2VIS_SEQREF

A.2 NO2VIS-SMALL product

Note that the plot for FIXREF has been already shown in Fig. 49.

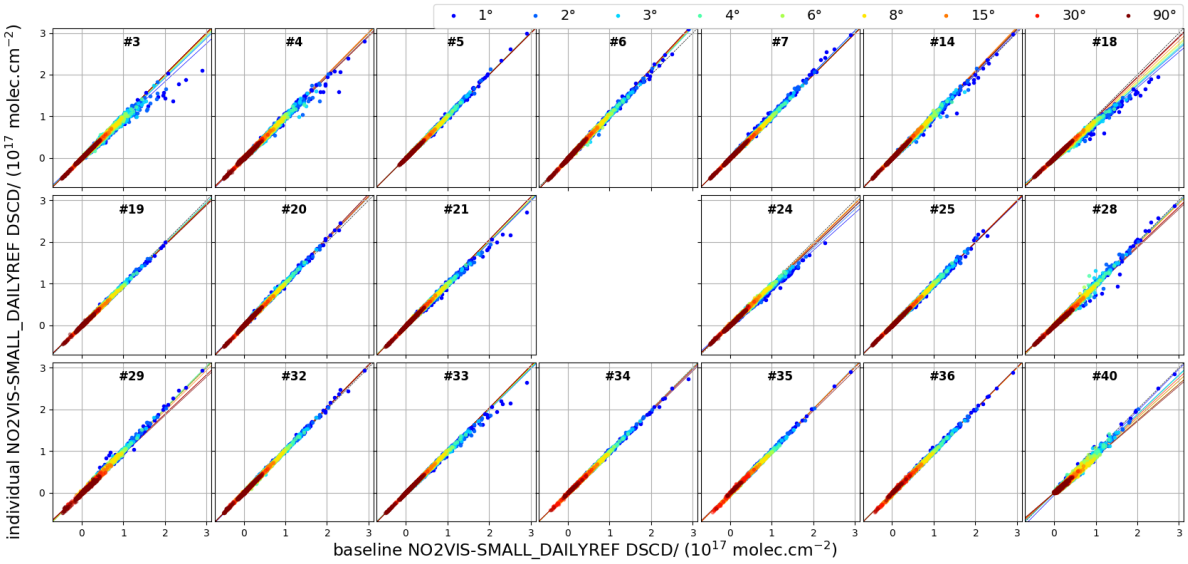

Figure 99: As Fig. 49 but for NO2VIS-SMALL_DAILYREF

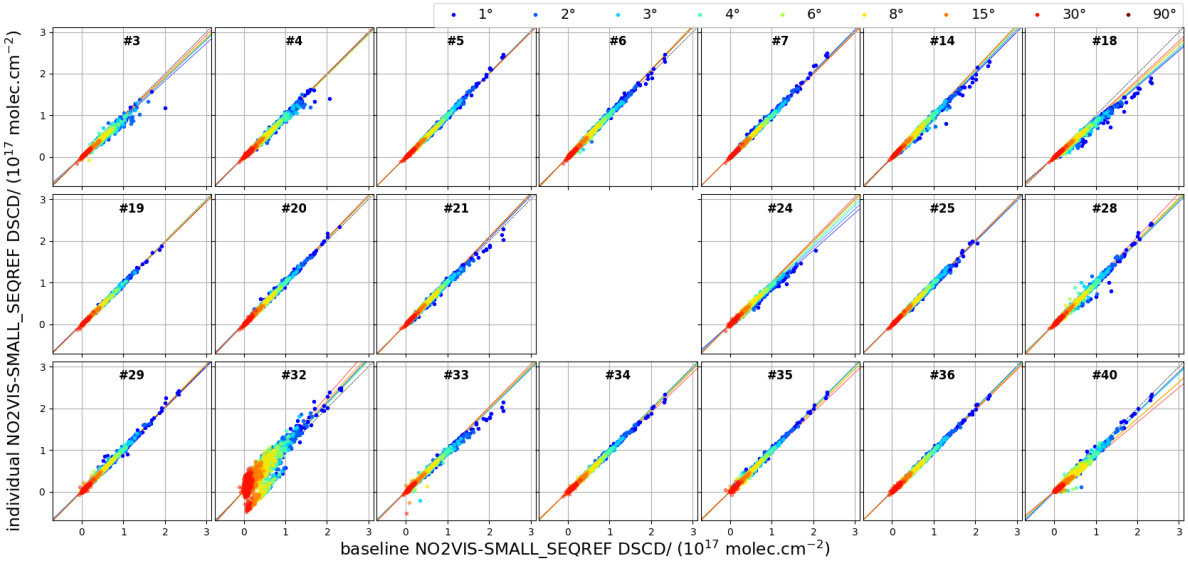

Figure 100: As Fig. 49 but for NO2VIS-SMALL_SEQREF

A.3 NO2UV product

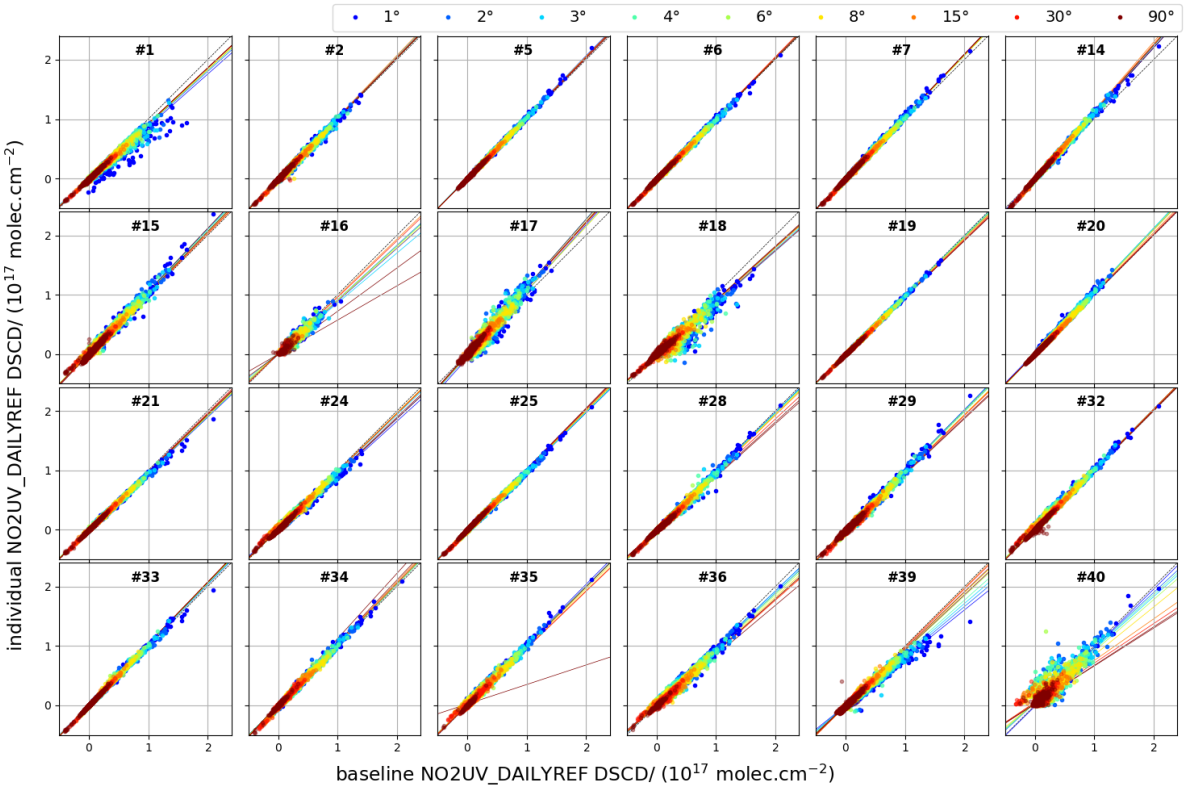

Figure 101: As Fig. 49 but for NO2UV_DAILYREF

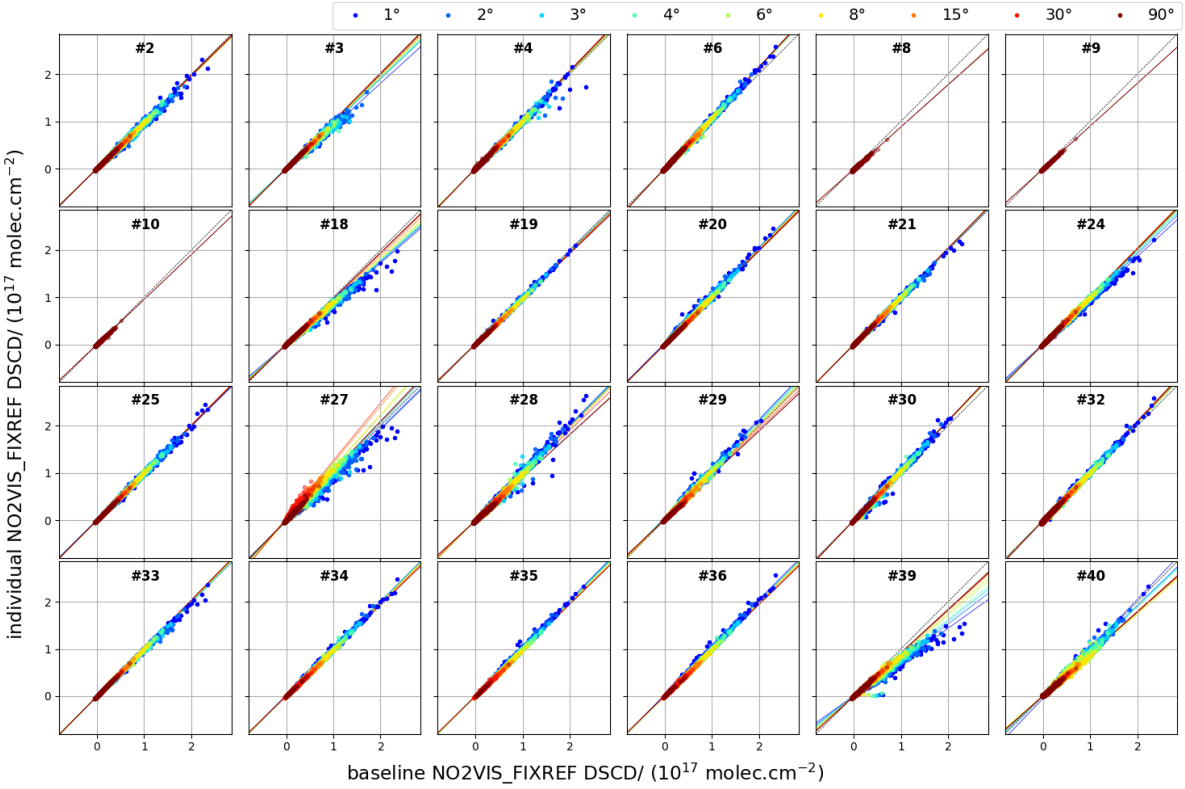

Figure 102: As Fig. 49 but for NO2UV_FIXREF

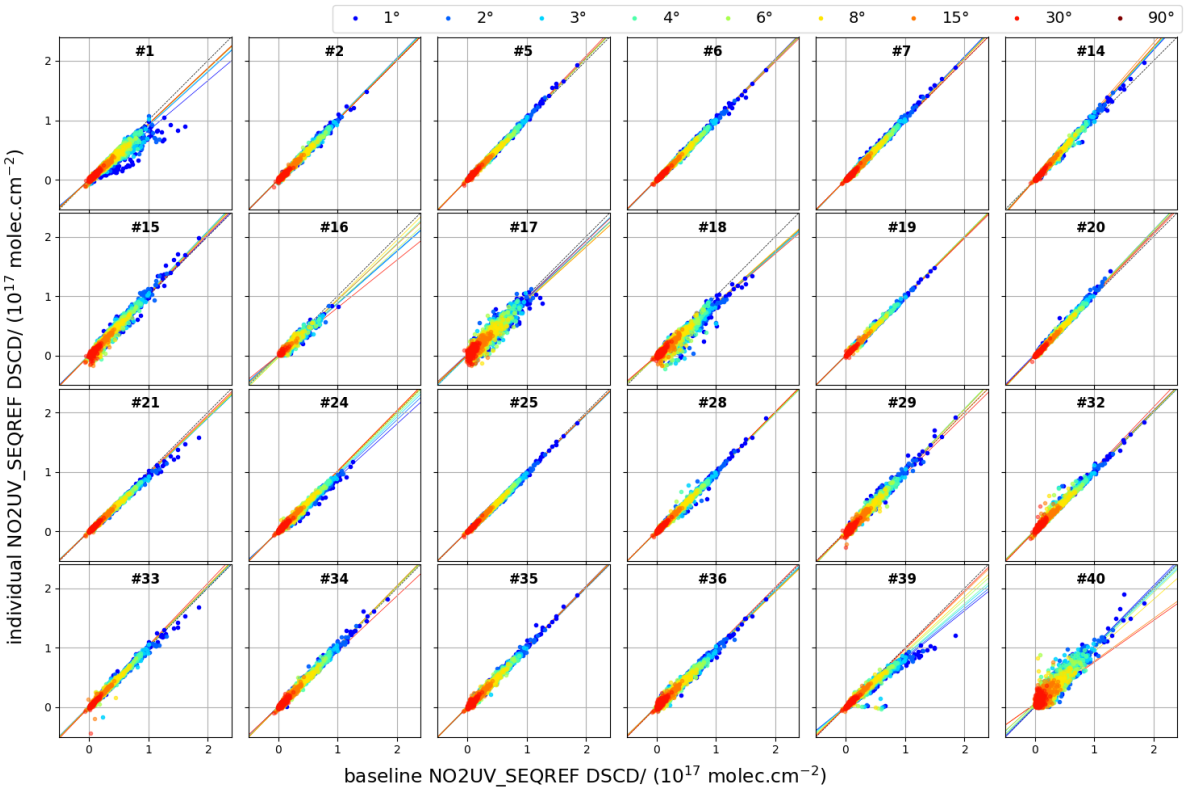

Figure 103: As Fig. 49 but for NO2UV_SEQREF

A.4 O4VIS product

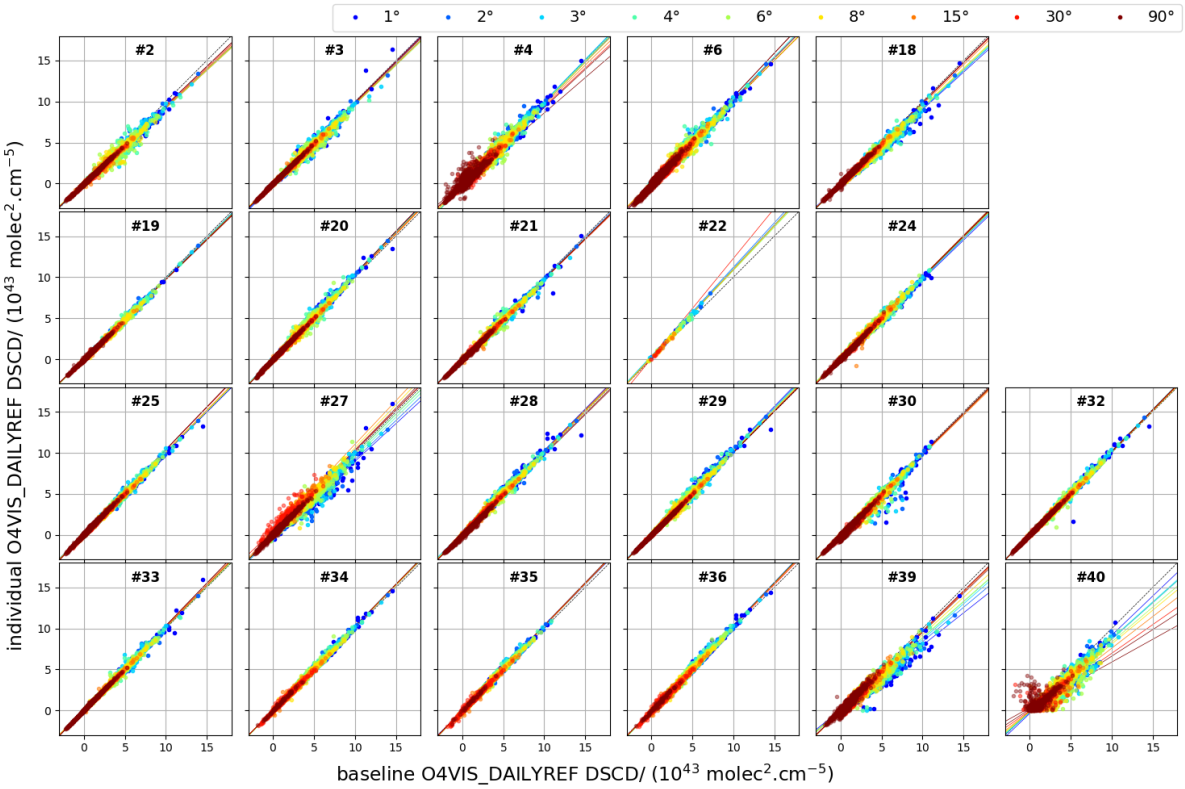

Figure 104: As Fig. 49 but for O4VIS_DAILYREF

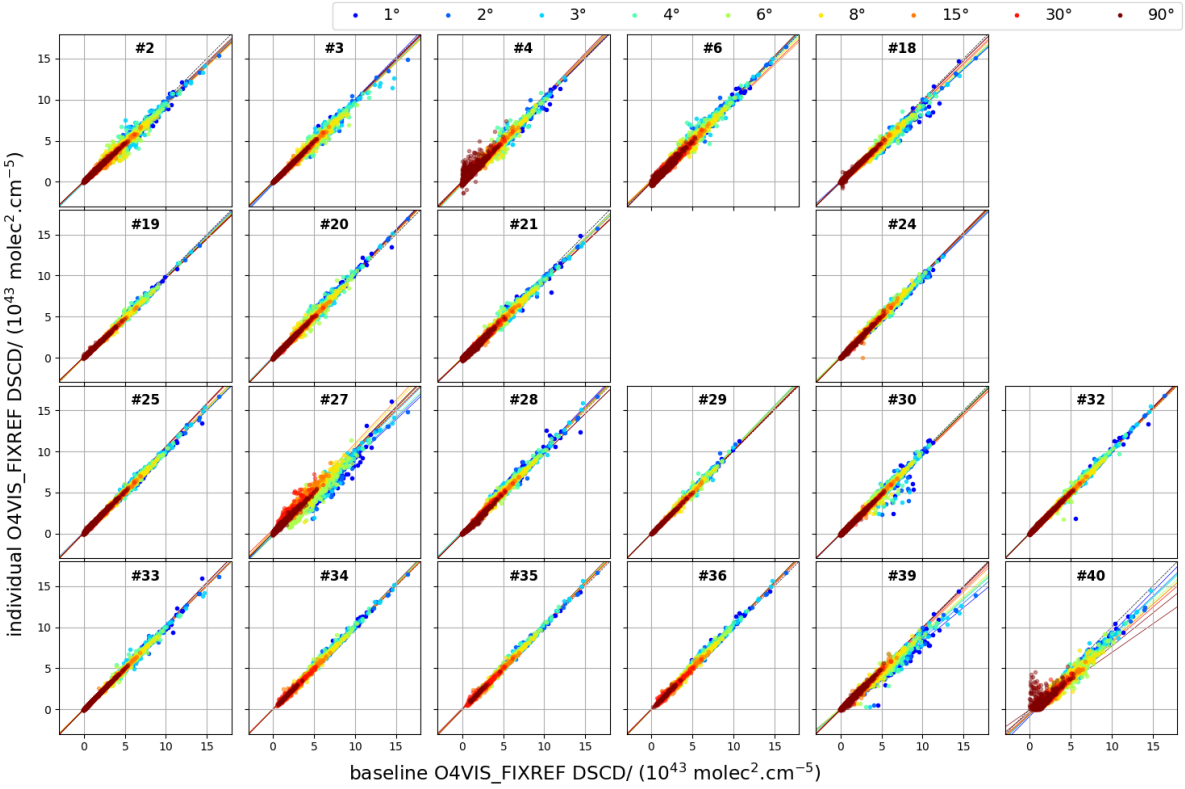

Figure 105: As Fig. 49 but for O4VIS_FIXREF

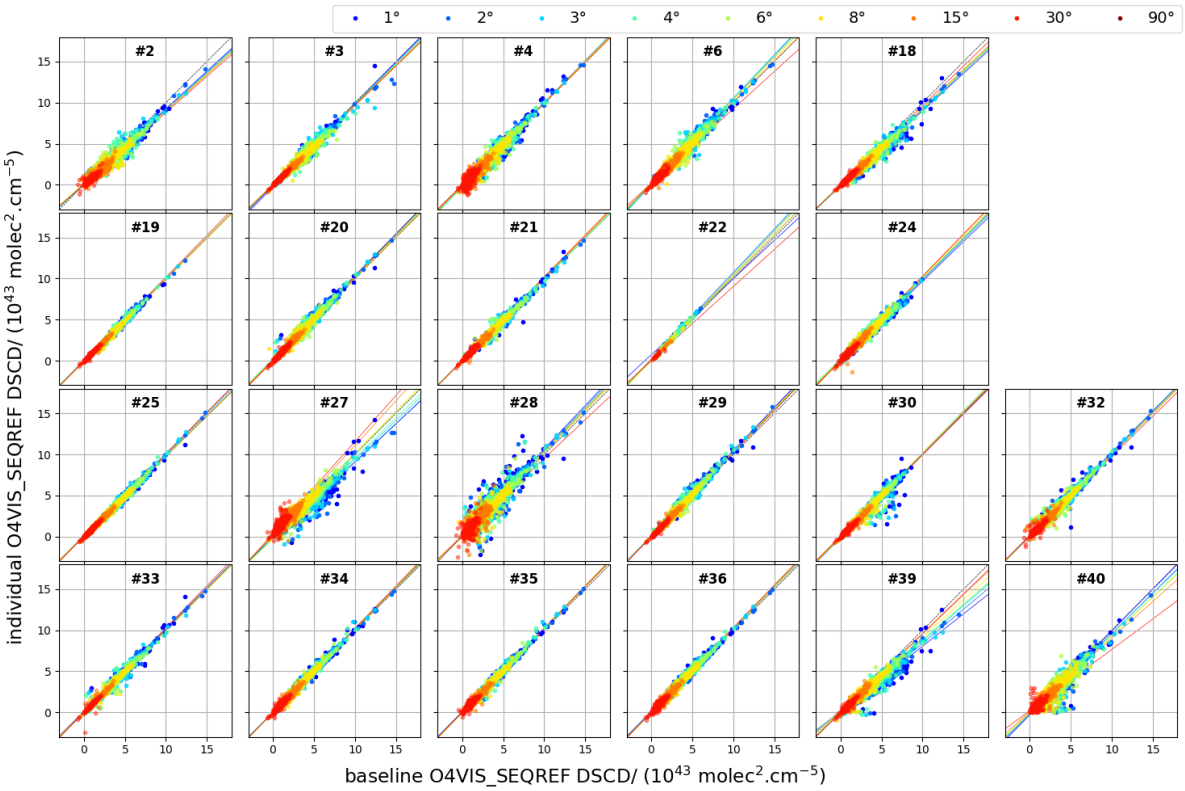

Figure 106: As Fig. 49 but for O4VIS_SEQREF

A.5 O4UV product

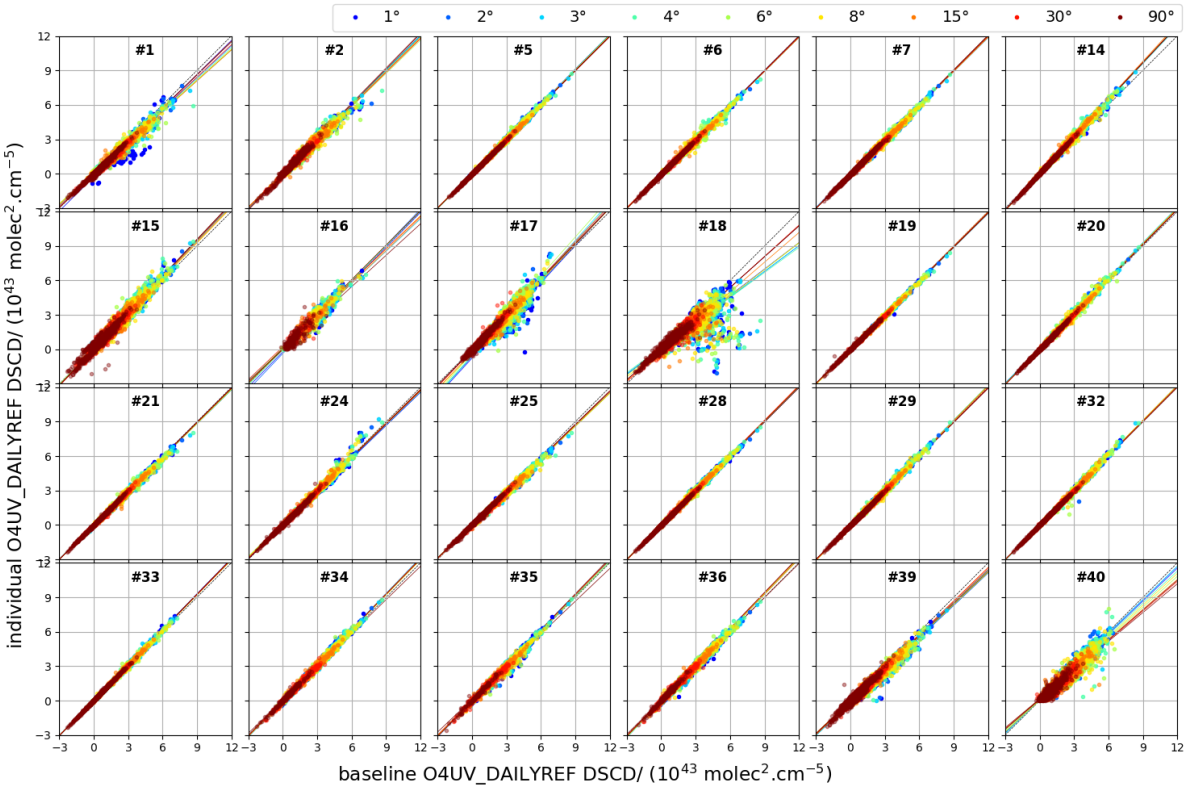

Figure 107: As Fig. 49 but for O4UV_DAILYREF

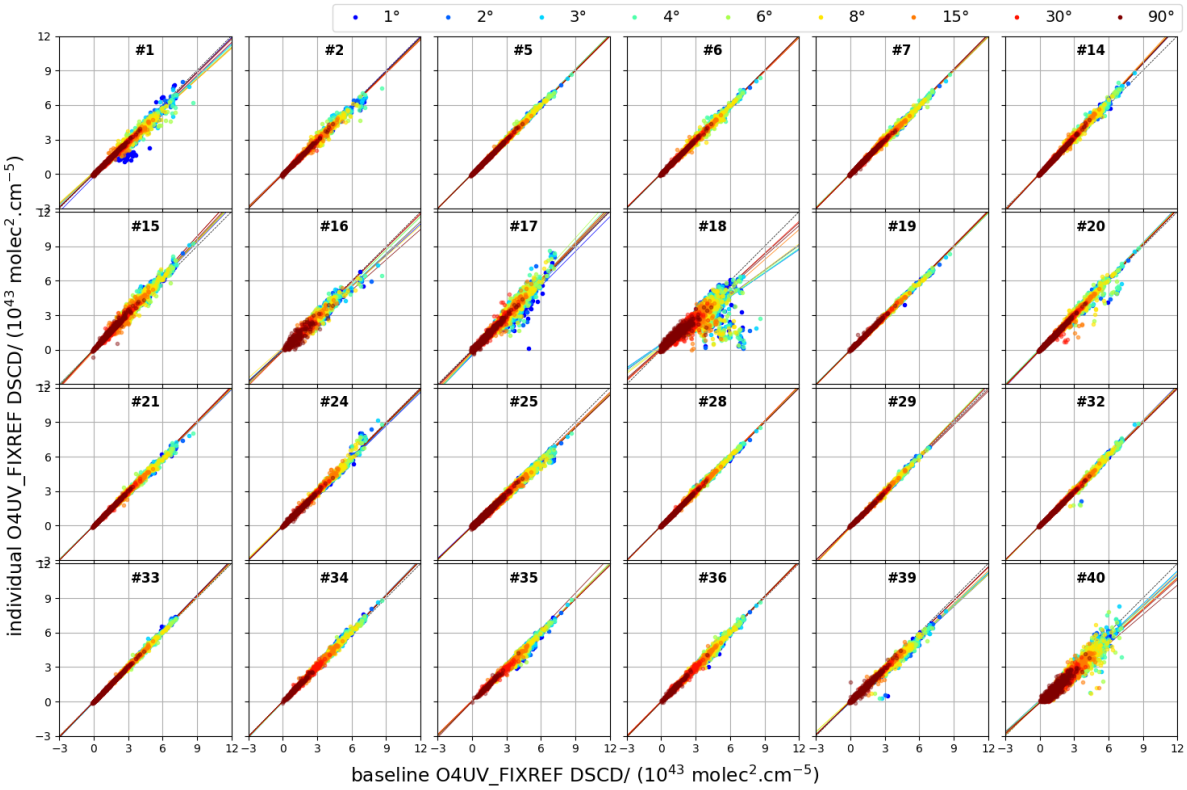

Figure 108: As Fig. 49 but for O4UV_FIXREF

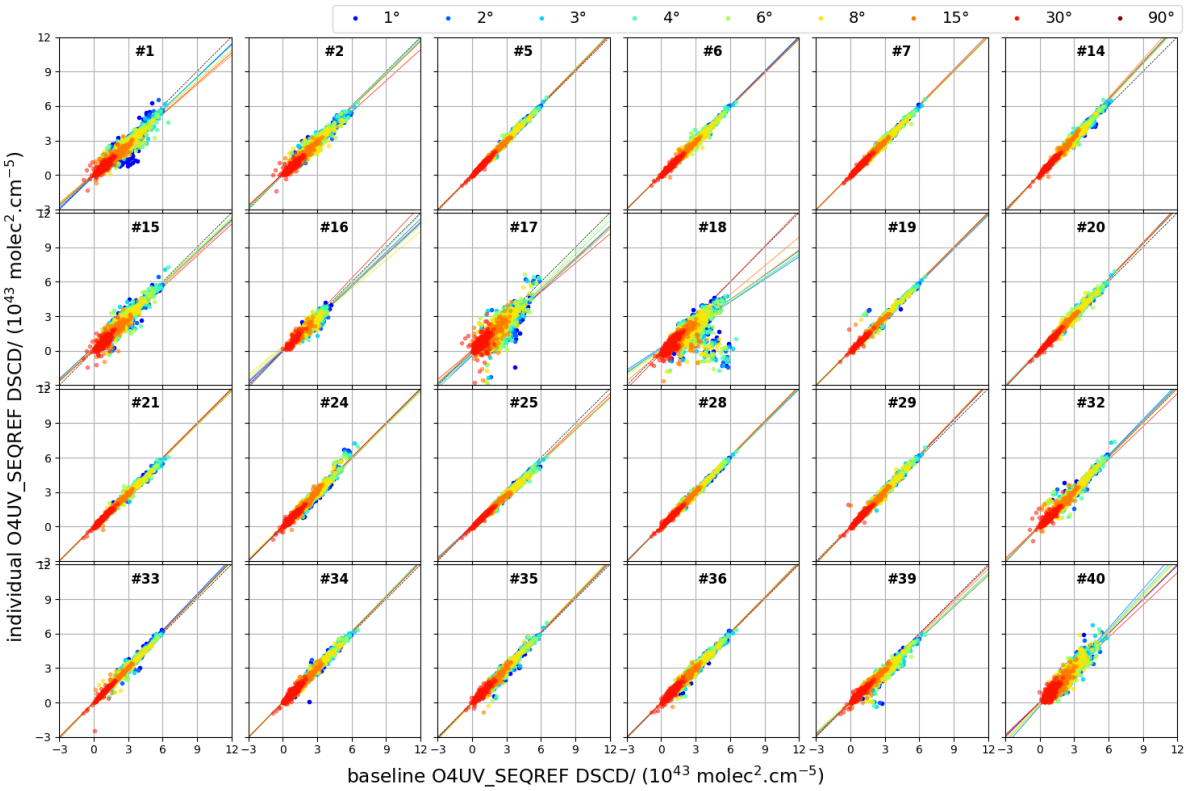

Figure 109: As Fig. 49 but for O4UV_SEQREF

A.6 O3VIS product

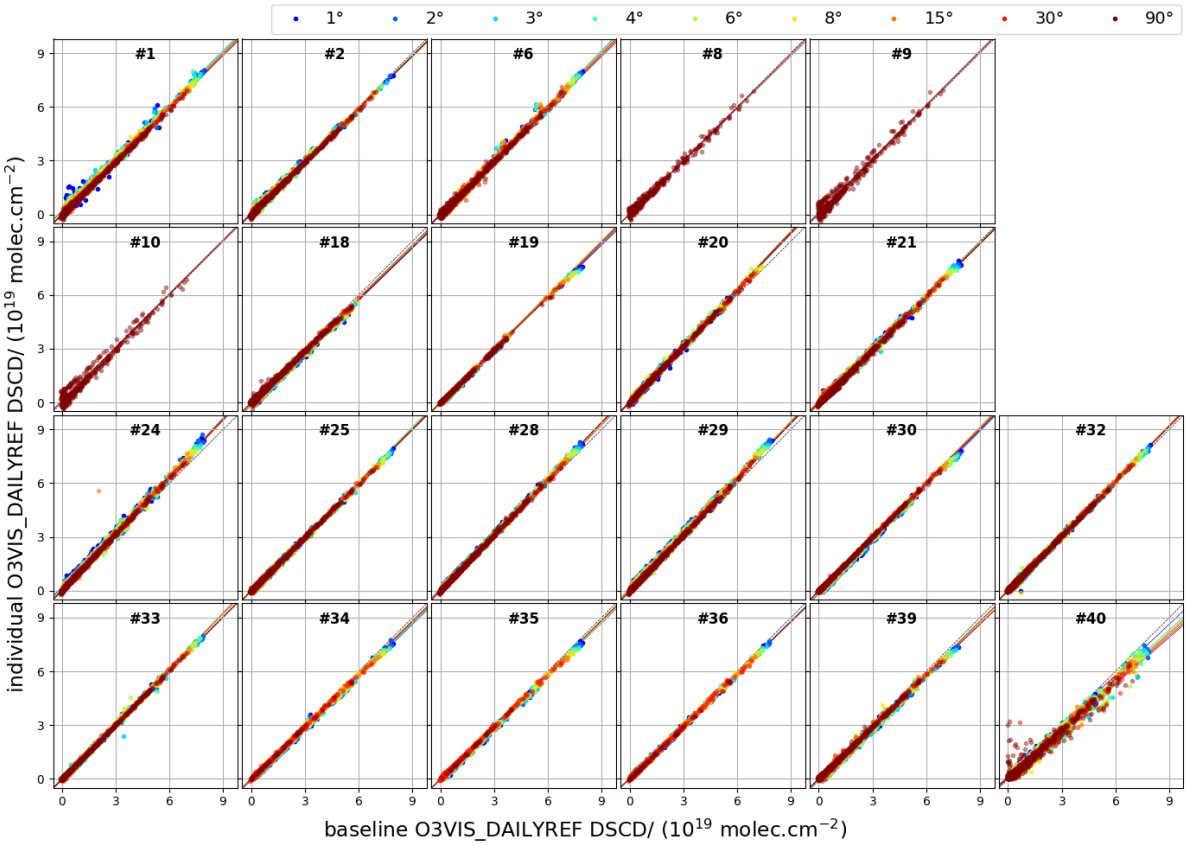

Figure 110: As Fig. 49 but for O3VIS_DAILYREF.

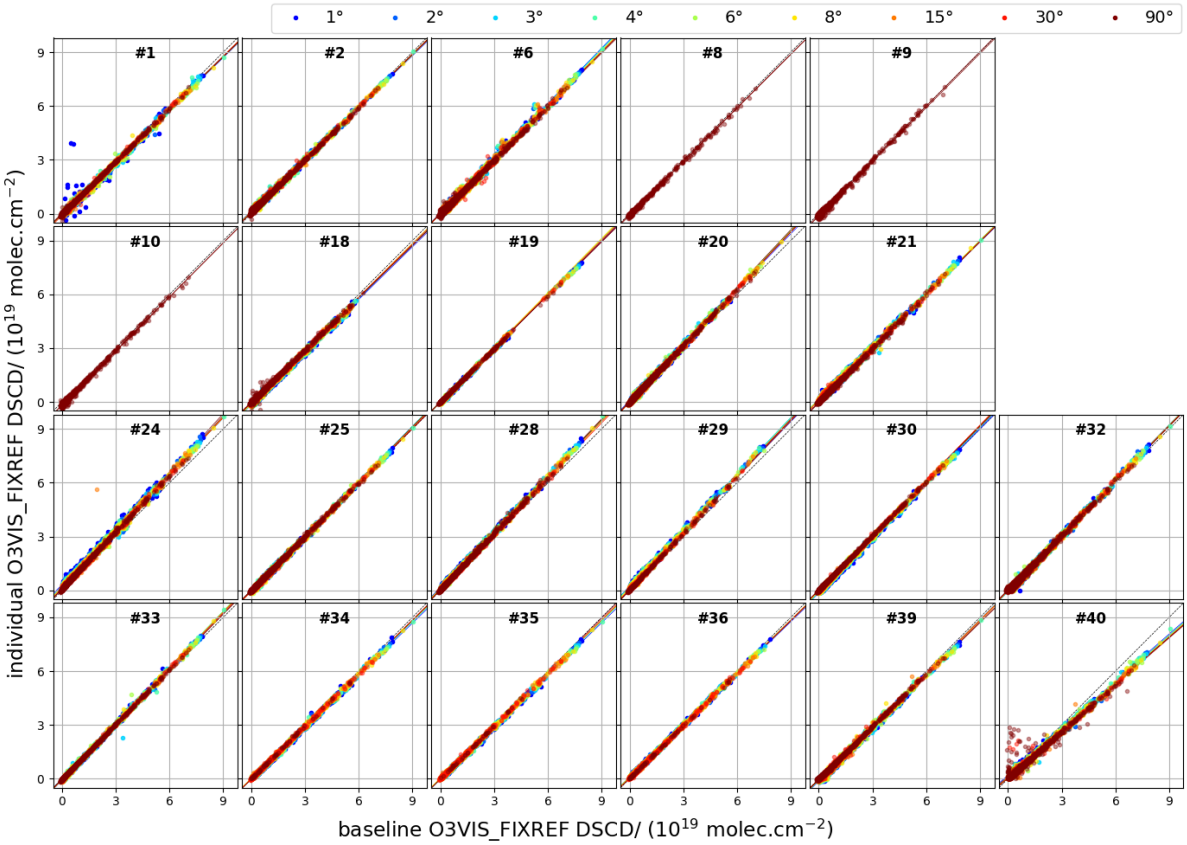

Figure 111: As Fig. 49 but for O3VIS_FIXREF

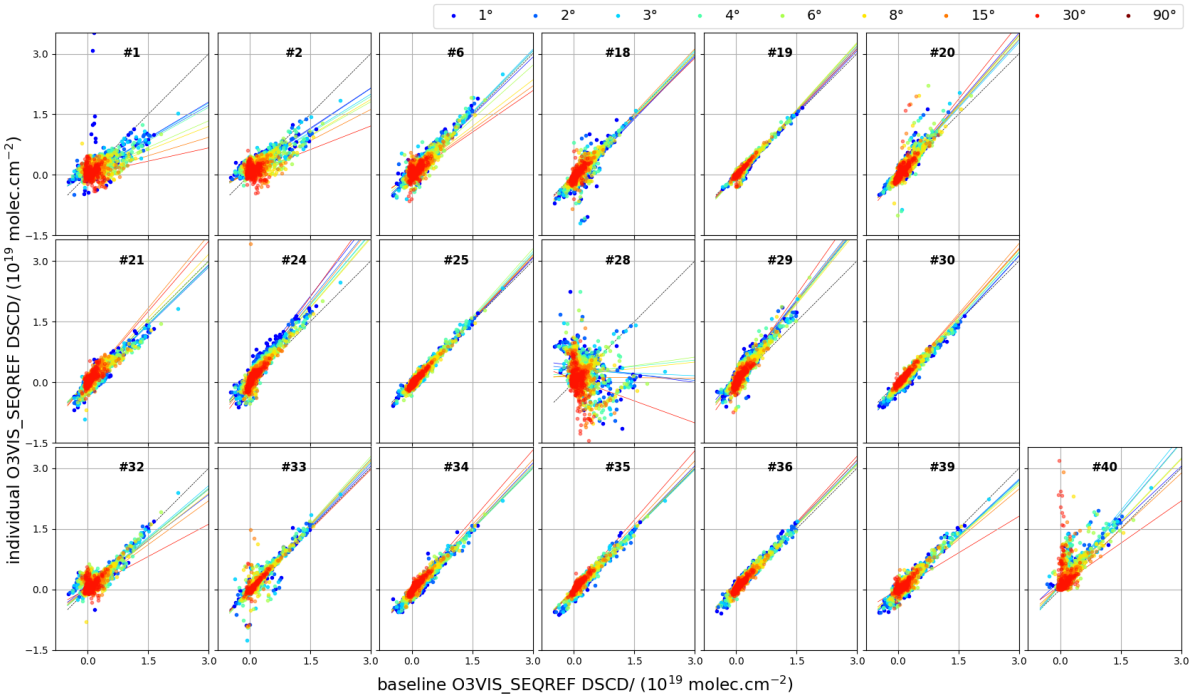

Figure 112: As Fig. 49 but for O3VIS_SEQREF. Note that not all values for instrument #3 are displayed

A.7 O3UV product

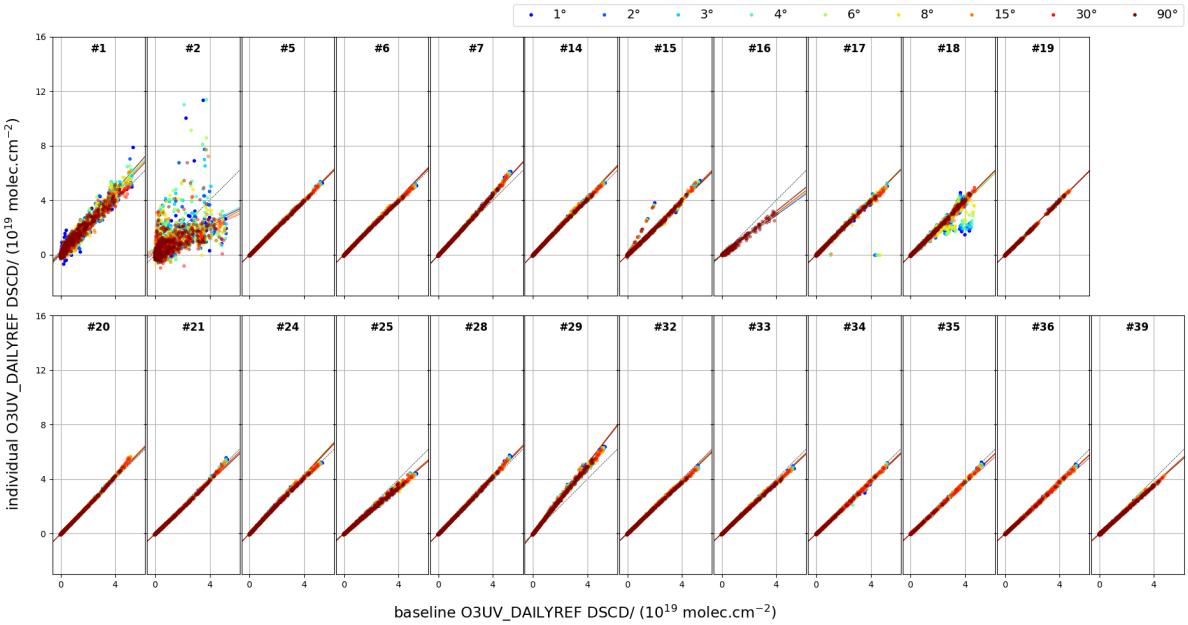

Figure 113: As Fig. 49 but for O3UV_DAILYREF. Note that not all values for instrument #2 are displayed.

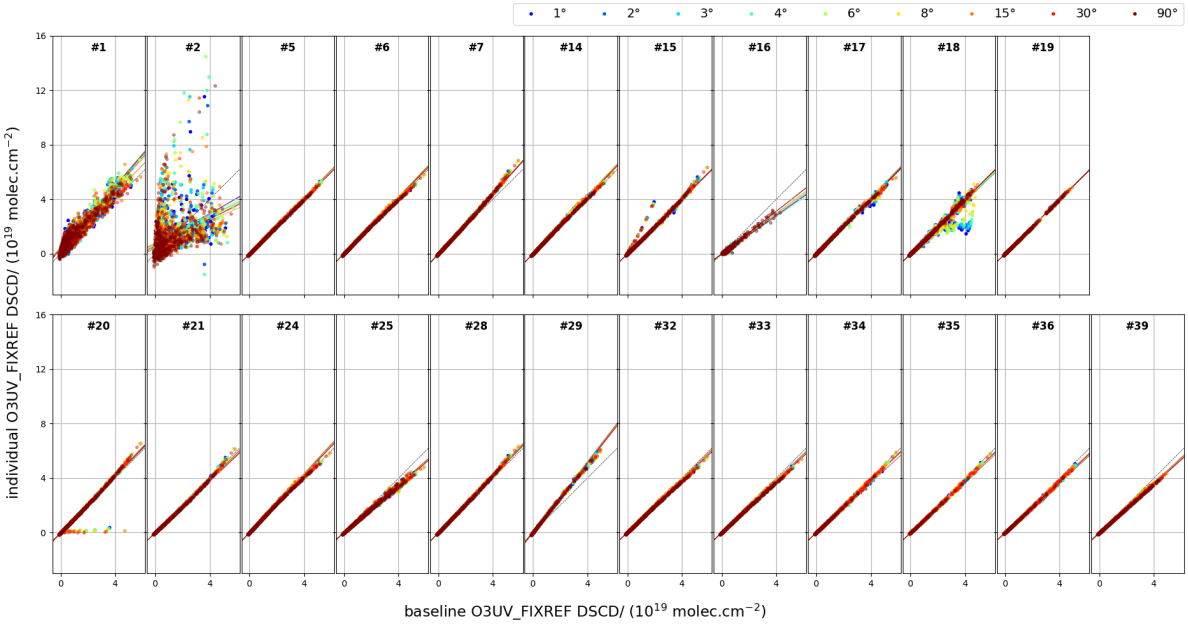

Figure 114: As Fig. 49 but for O3UV_FIXREF. Note that not all values for instrument #2 are displayed.

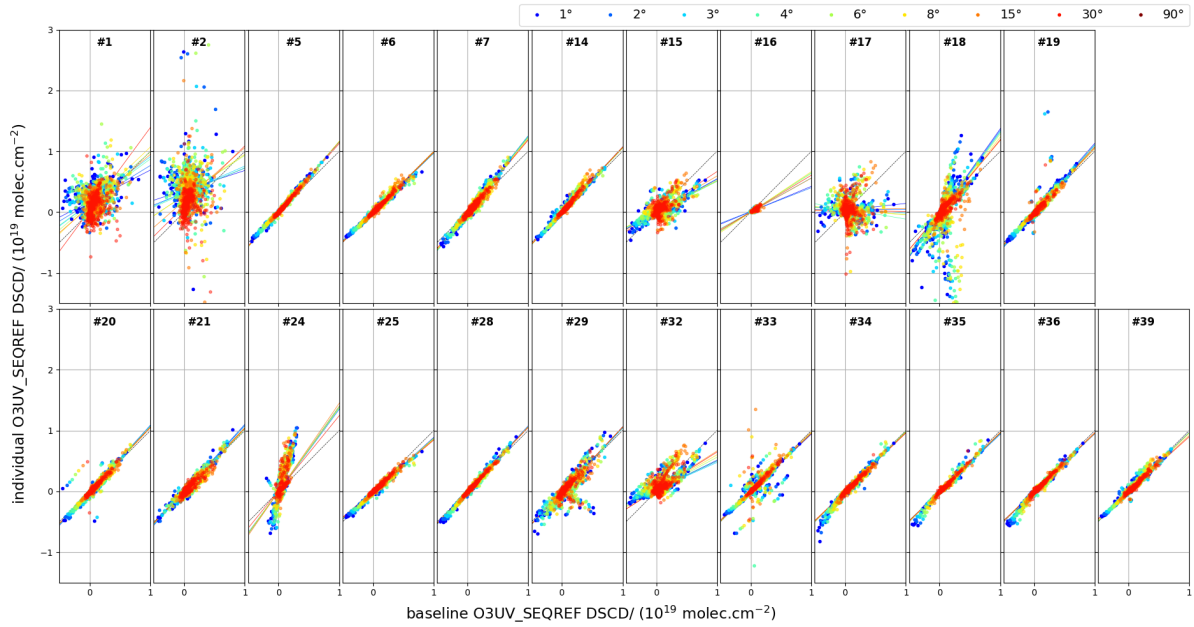

Figure 115: As Fig. 49 but for O3UV_SEQREF. Note that not all values for instrument #2 are displayed.

A.8 HCHO product

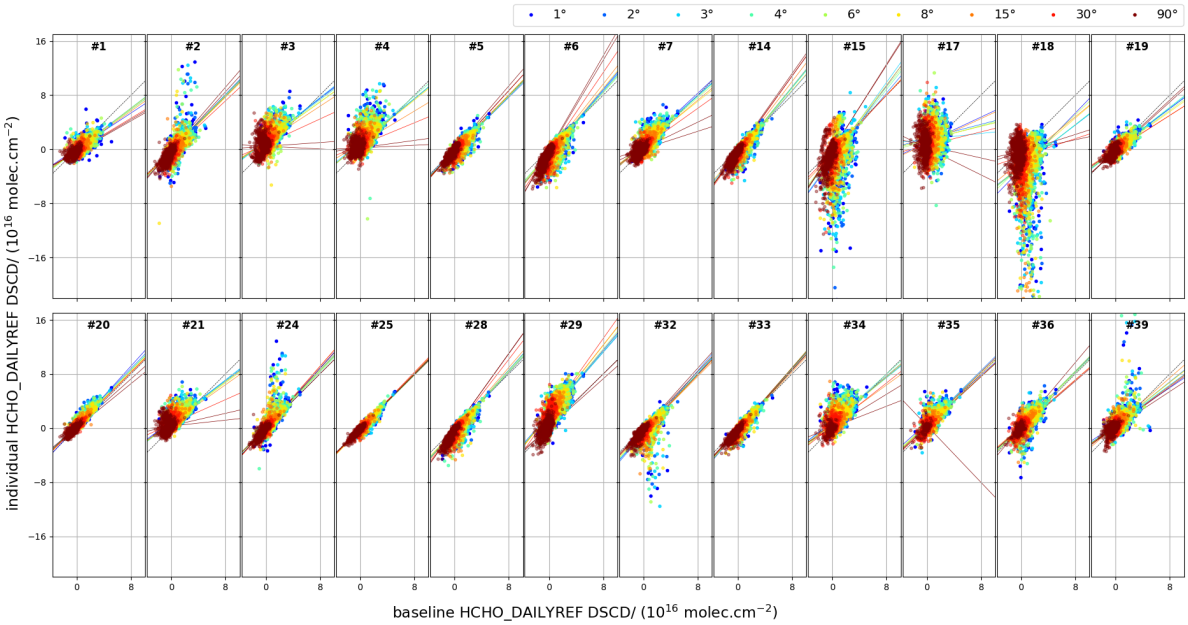

Figure 116: As Fig. 49 but for HCHO_DAILYREF. Note that not all values for instruments #15, #17 and #18 are displayed.

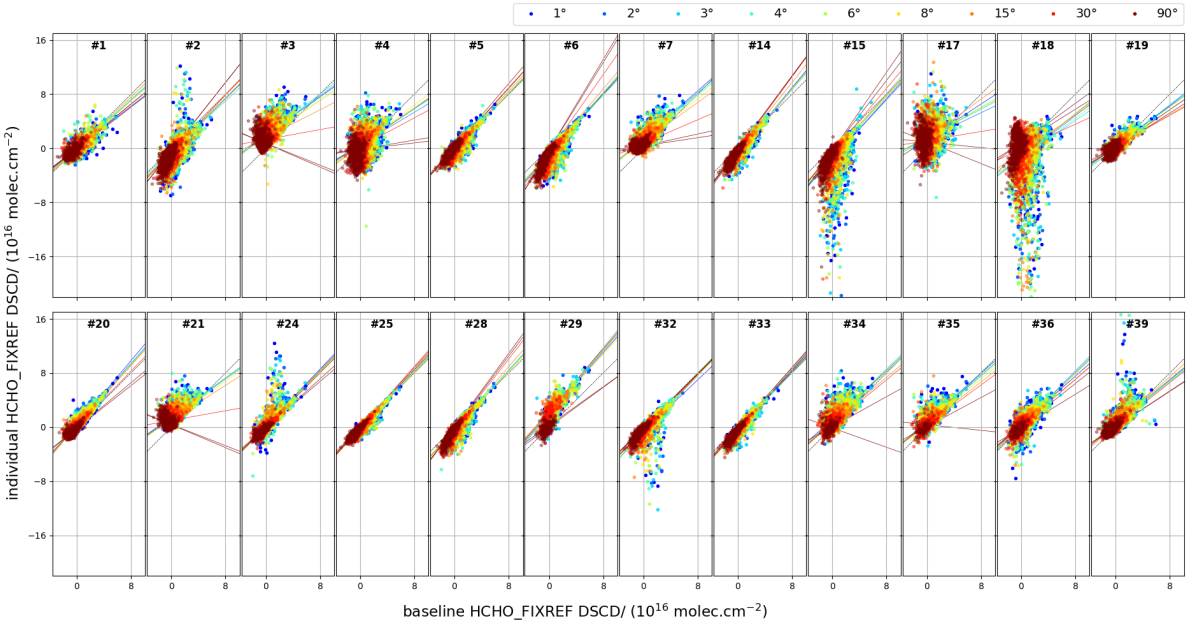

Figure 117: As Fig. 49 but for HCHO_FIXREF. Note that not all values for instruments #15 and #18 are displayed.

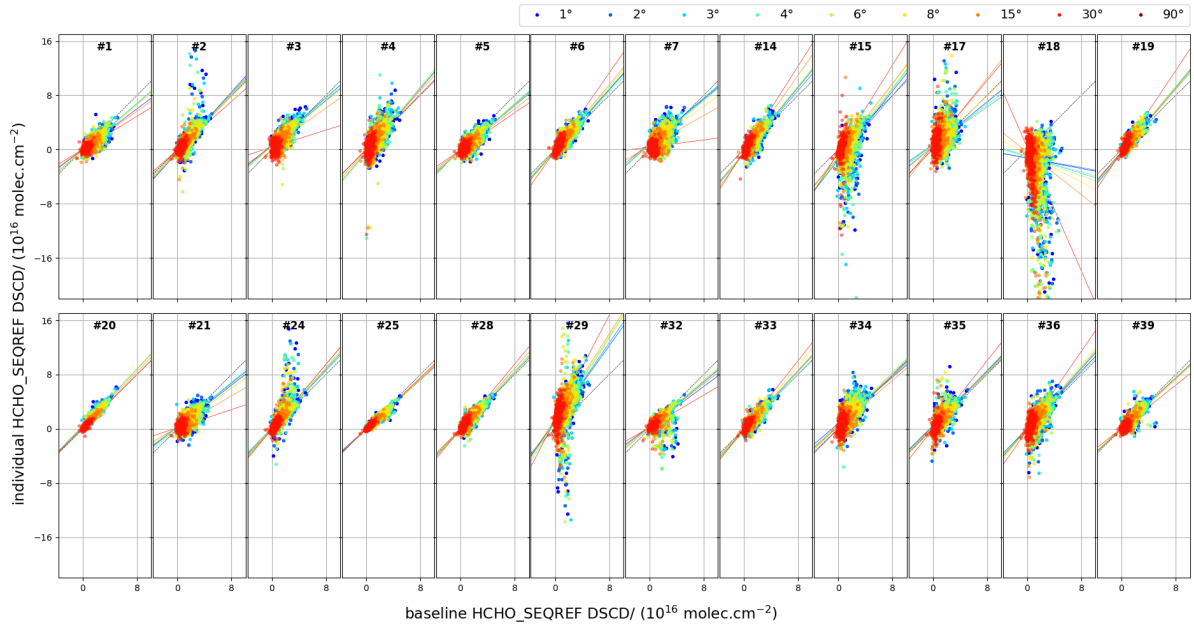

Figure 118: As Fig. 49 but for HCHO_SEQREF. Note that not all values for instruments #2, #15, #17, #17, #18, and #29 are displayed.

A.9 HCHO-WIDE product

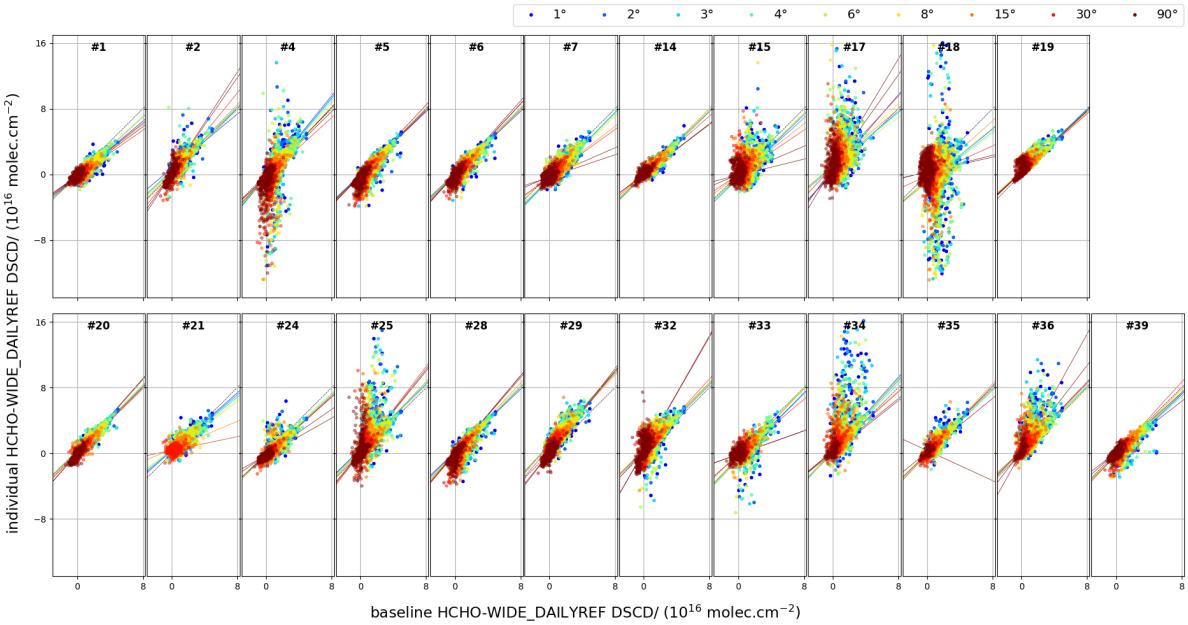

Figure 119: As Fig. 49 but for HCHO-WIDE_DAILYREF. Note that not all values for instruments #15, #17, #18 and #34 are displayed.

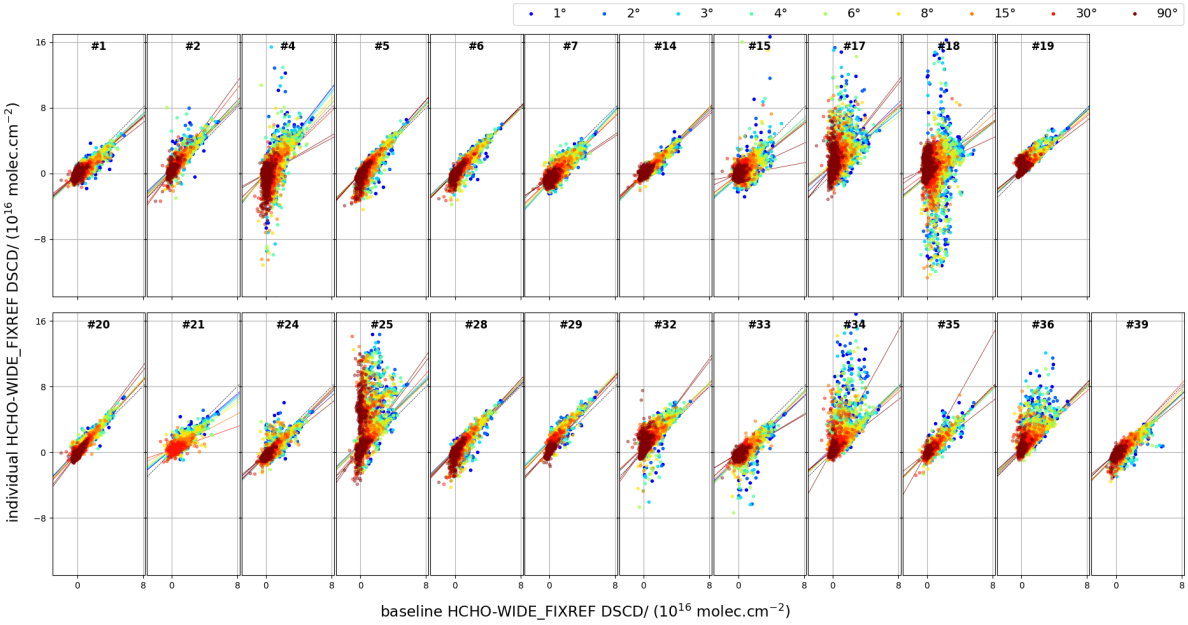

Figure 120: As Fig. 49 but for HCHO-WIDE_FIXREF. Note that not all values for instruments #4, #17, #18 and #34 are displayed.

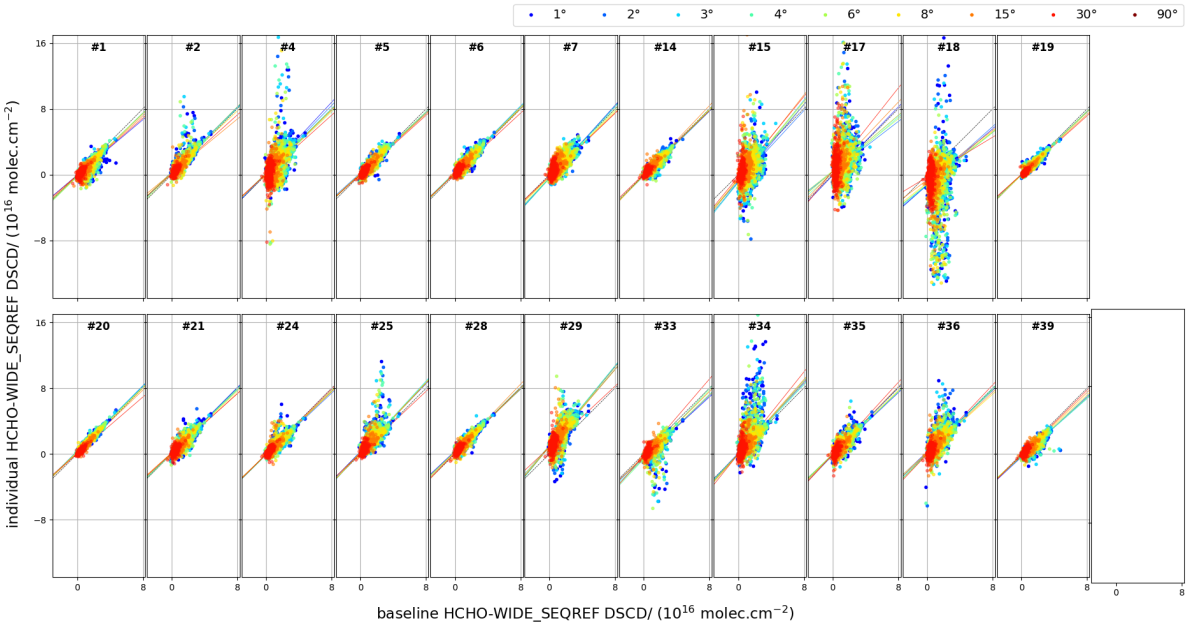

Figure 121: As Fig. 49 but for HCHO-WIDE_SEQREF. Note that not all values for instruments #4, #15, and #34 are displayed.

A.10 HONO product

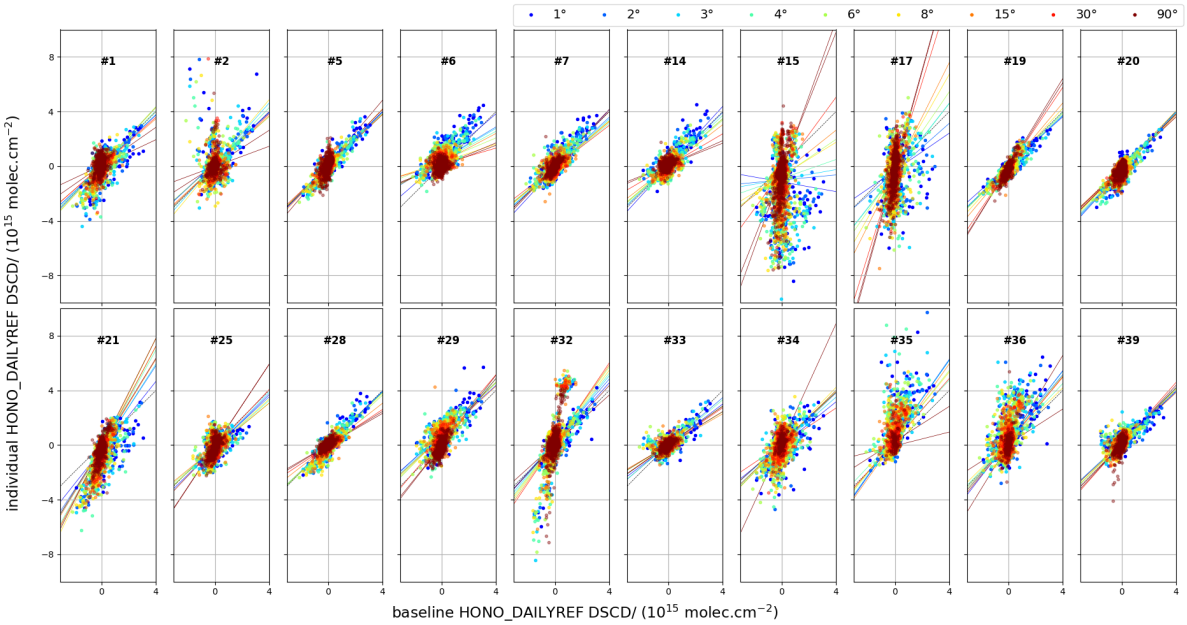

Figure 122: As Fig. 49 but for HONO_DAILYREF. Note that not all values for instruments #2, #15, #17, #32 and #35 are displayed.

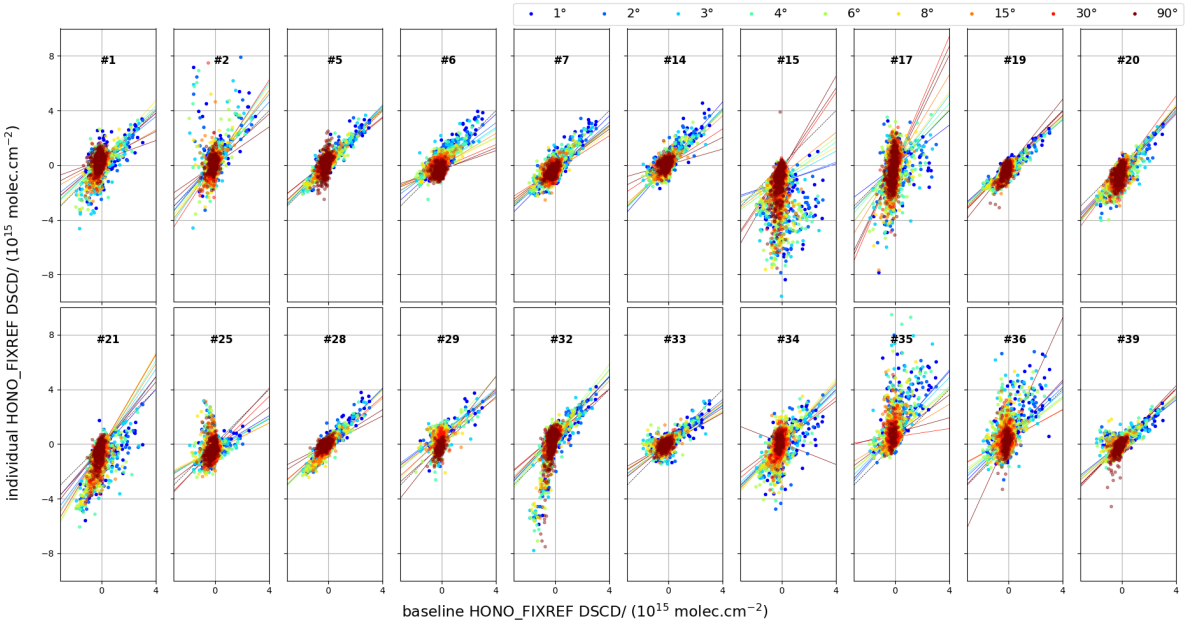

Figure 123: As Fig. 49 but for HONO_FIXREF. Note that not all values for instruments #2, #32 and #35 are displayed.

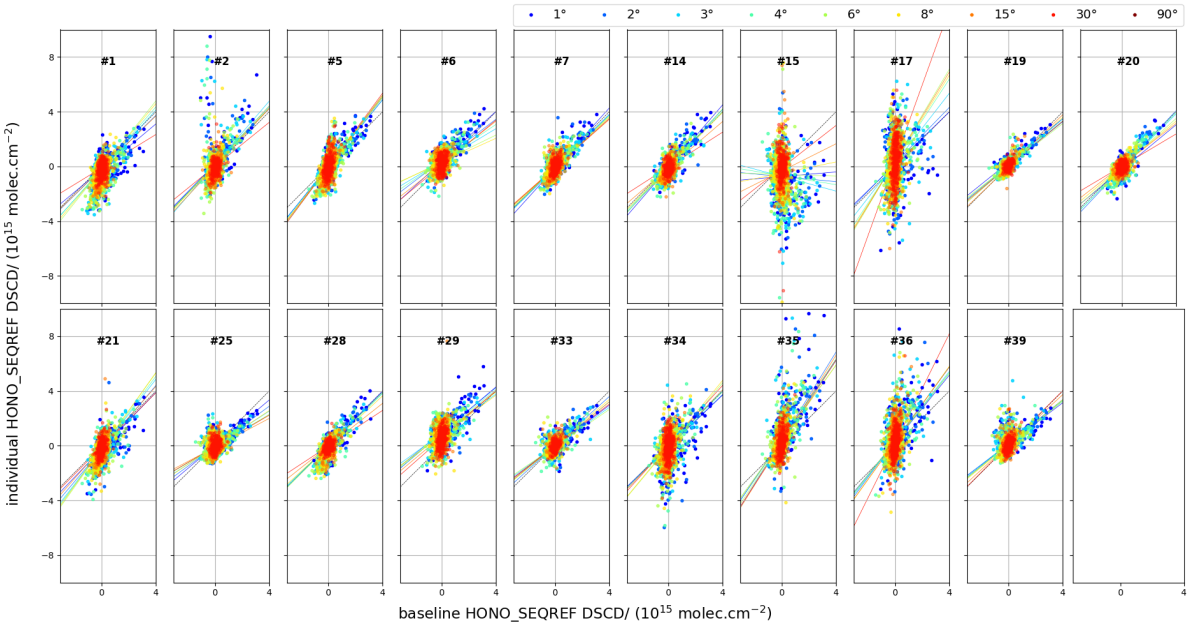

Figure 124: As Fig. 49 but for HONO_SEQREF. Note that not all values for instruments #2, #15, and #35 are displayed.

B Consistence between different reference types

Here we present the complete set of reference-type consistence checks where we compare the Δ $\mathrm{DSCD}_{\mathrm{DAILYREF}}$, Δ $\mathrm{DSCD}_{\mathrm{FIXREF}}$ and $\mathrm{DSCD}_{\mathrm{SEQREF}}$, where Δ is as defined in Eq. 4. NO2VIS was already shown in Fig. 51, Fig. 52 and Fig. 53.

B.1 NO2VIS-SMALL

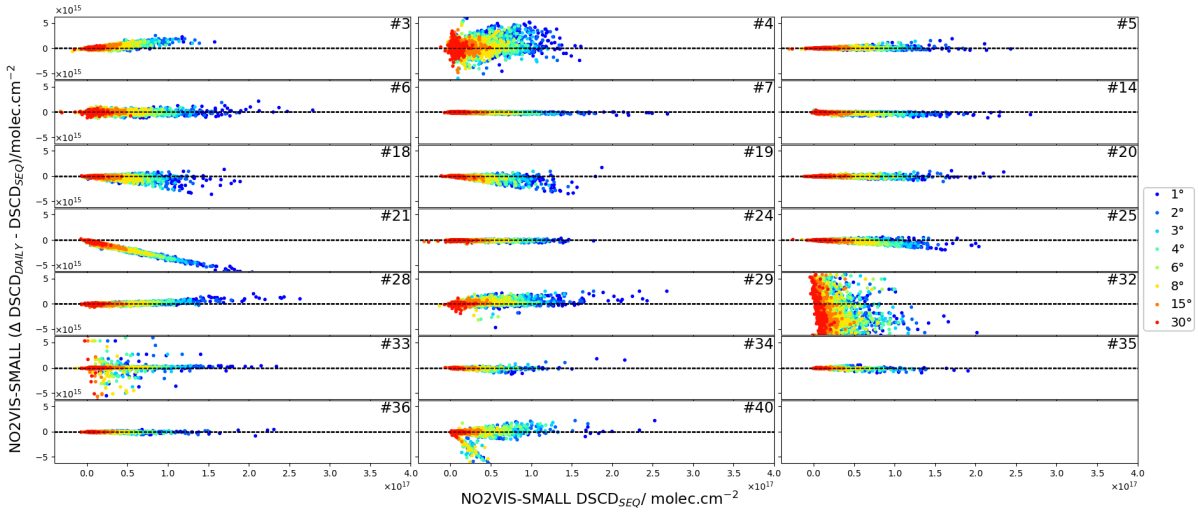

Figure 125: As Fig. 51 but for NO2VIS-SMALL

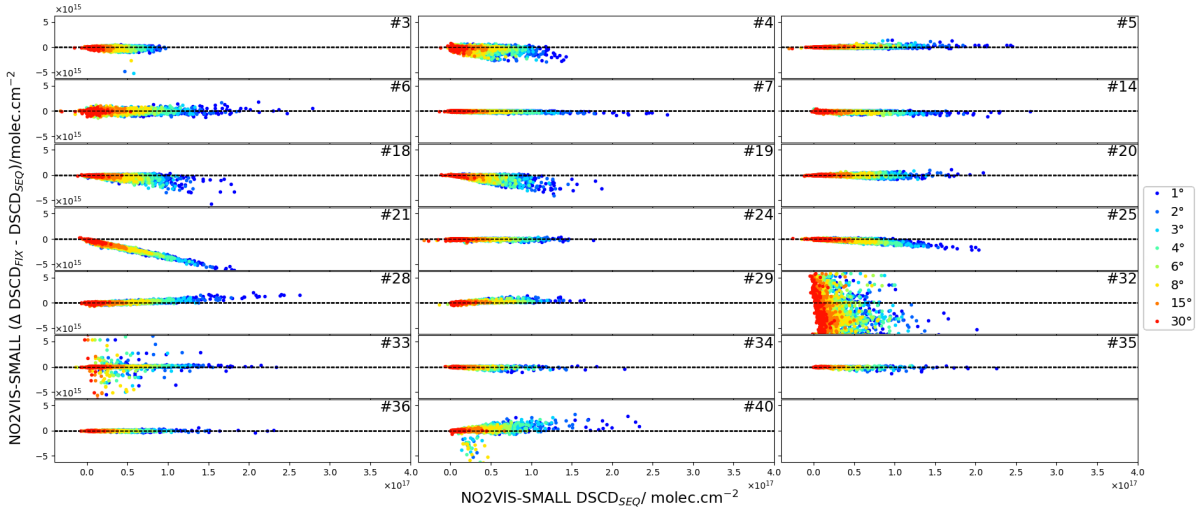

Figure 126: As Fig. 52 but for NO2VIS-SMALL

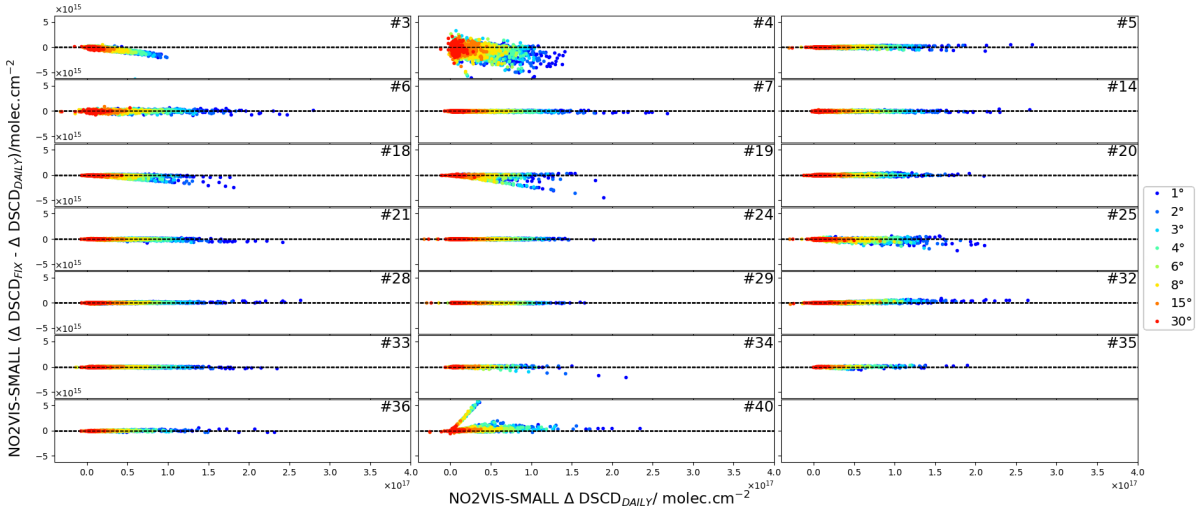

Figure 127: As Fig. 53 but for NO2VIS-SMALL

B.2 NO2UV

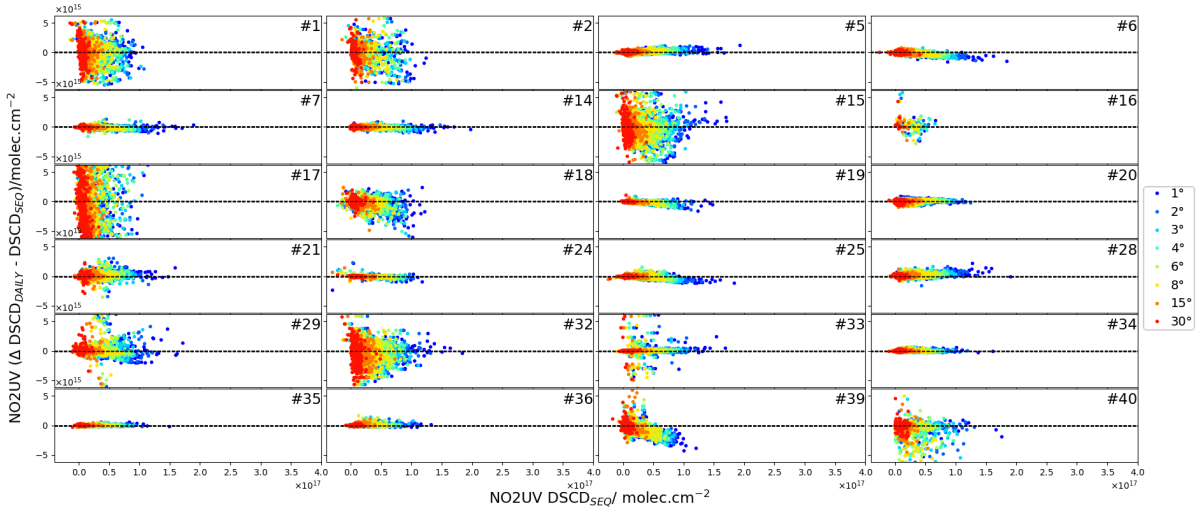

Figure 128: As Fig. 51 but for NO2UV

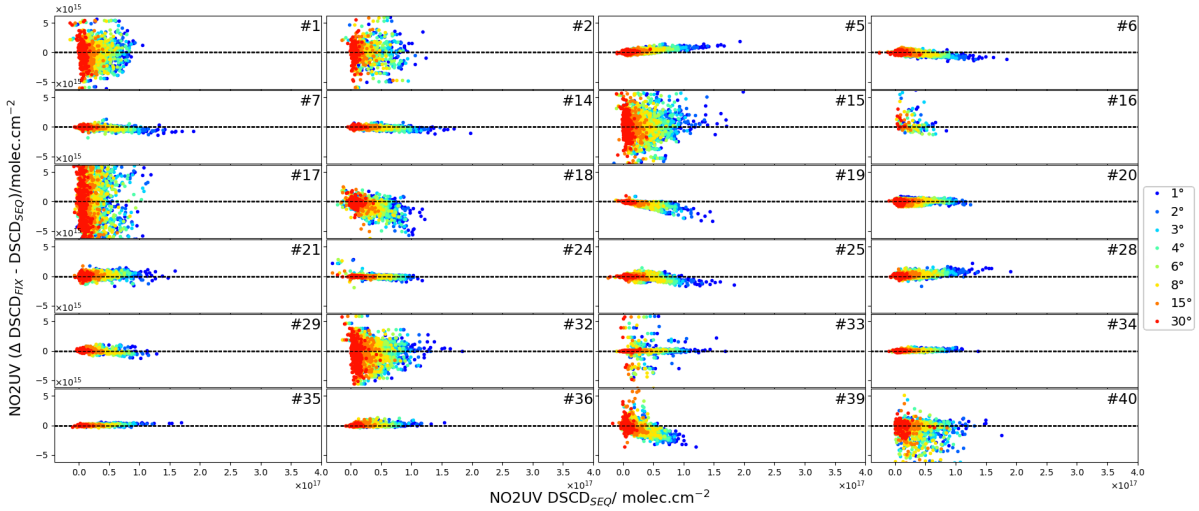

Figure 129: As Fig. 52 but for NO2UV

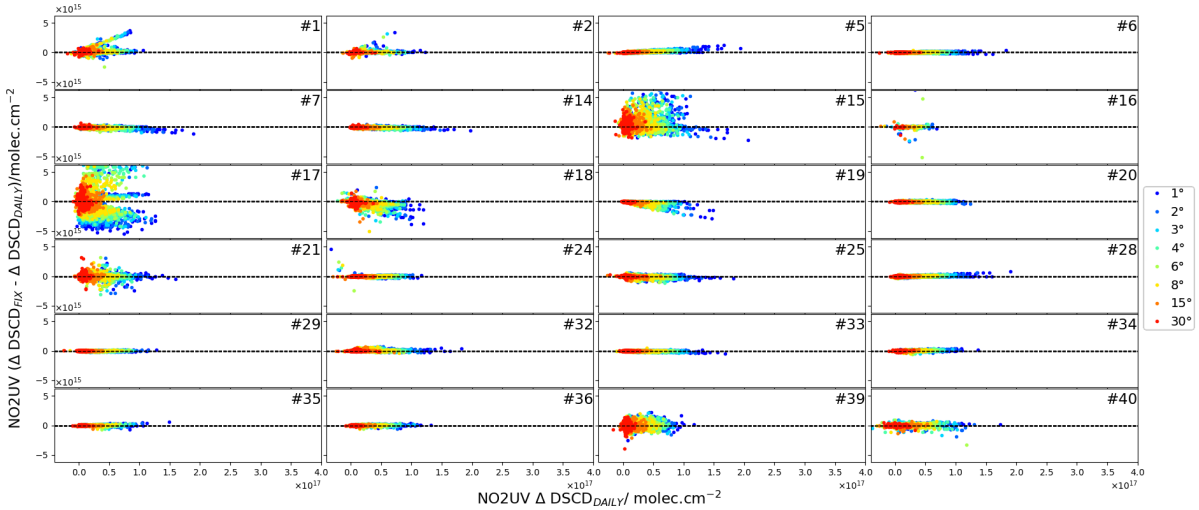

Figure 130: As Fig. 53 but for NO2UV

B.3 O3VIS

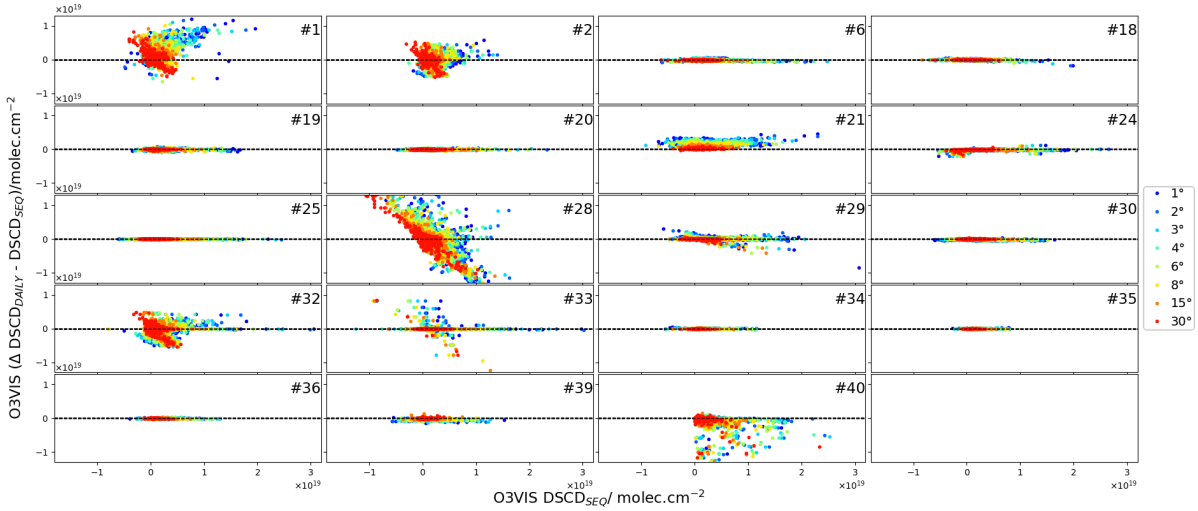

Figure 131: As Fig. 51 but for O3VIS

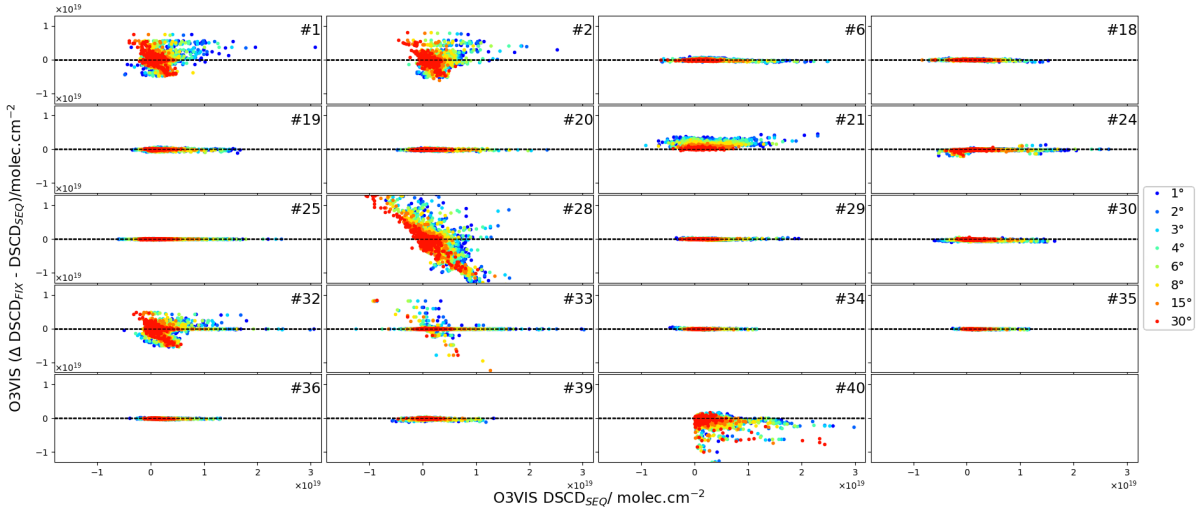

Figure 132: As Fig. 52 but for O3VIS

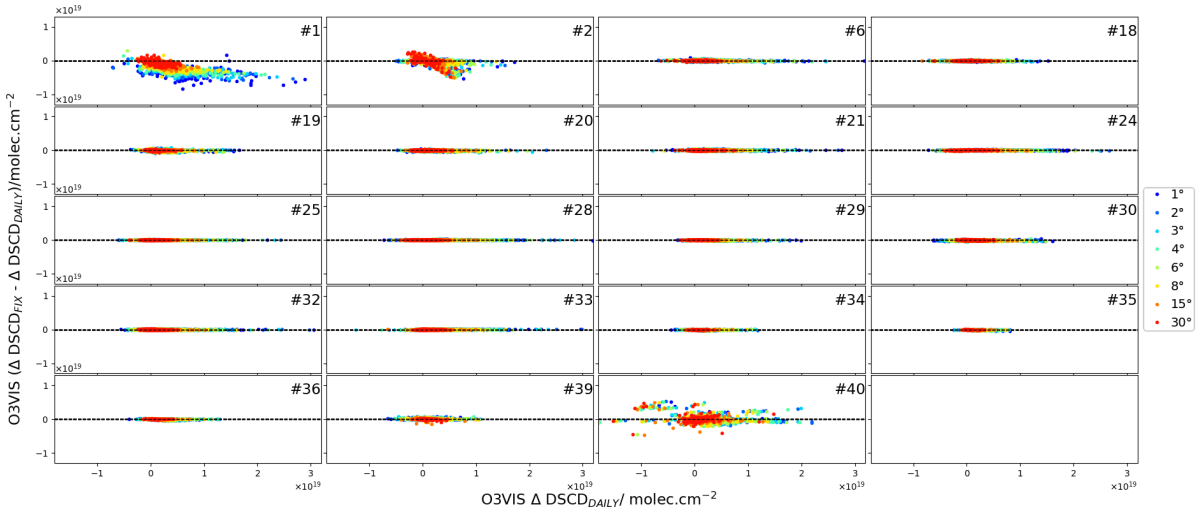

Figure 133: As Fig. 53 but for O3VIS

B.4 O3UV

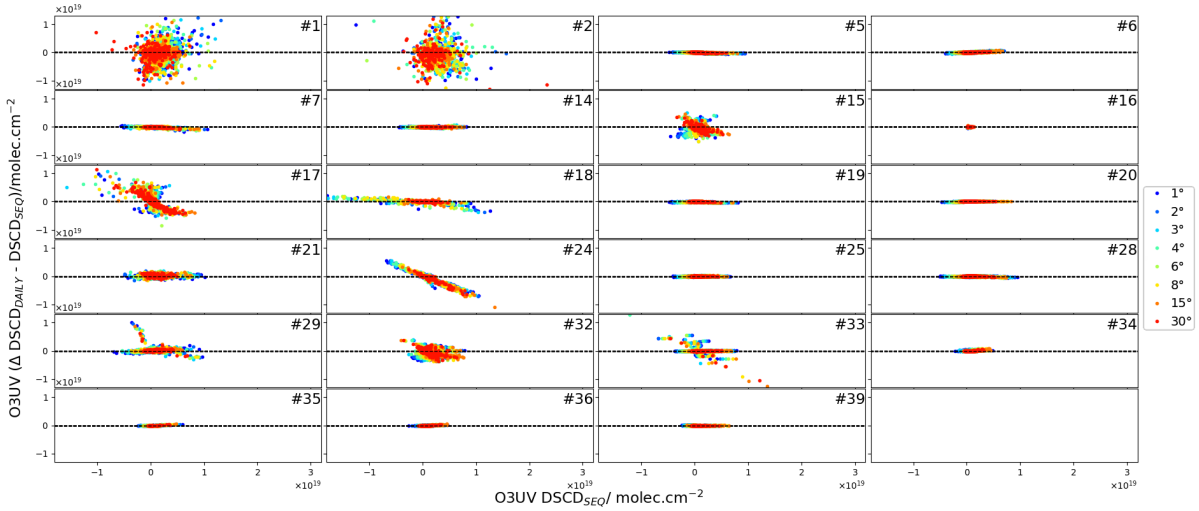

Figure 134: As Fig. 51 but for O3UV

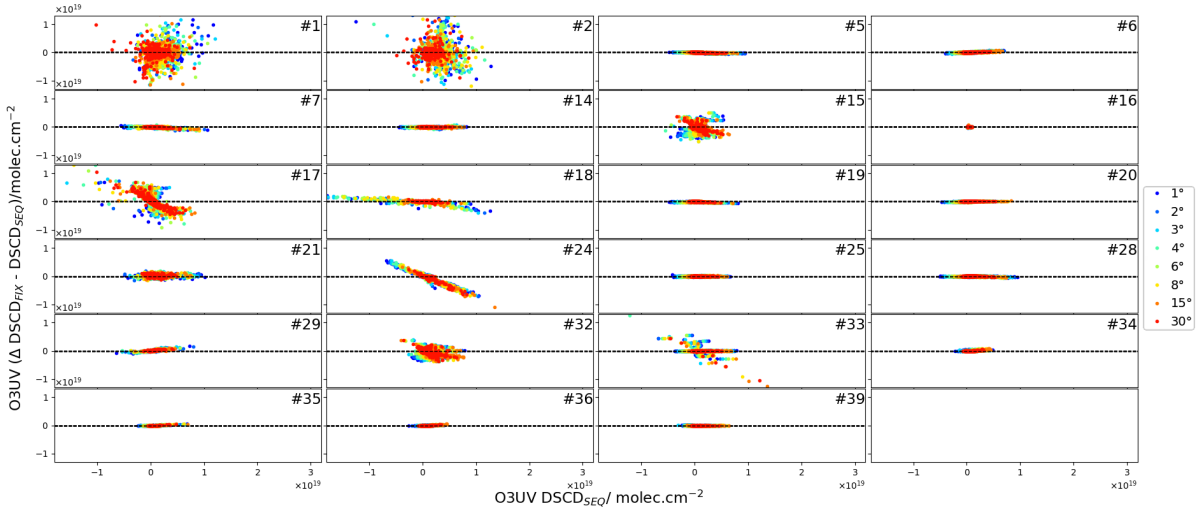

Figure 135: As Fig. 52 but for O3UV

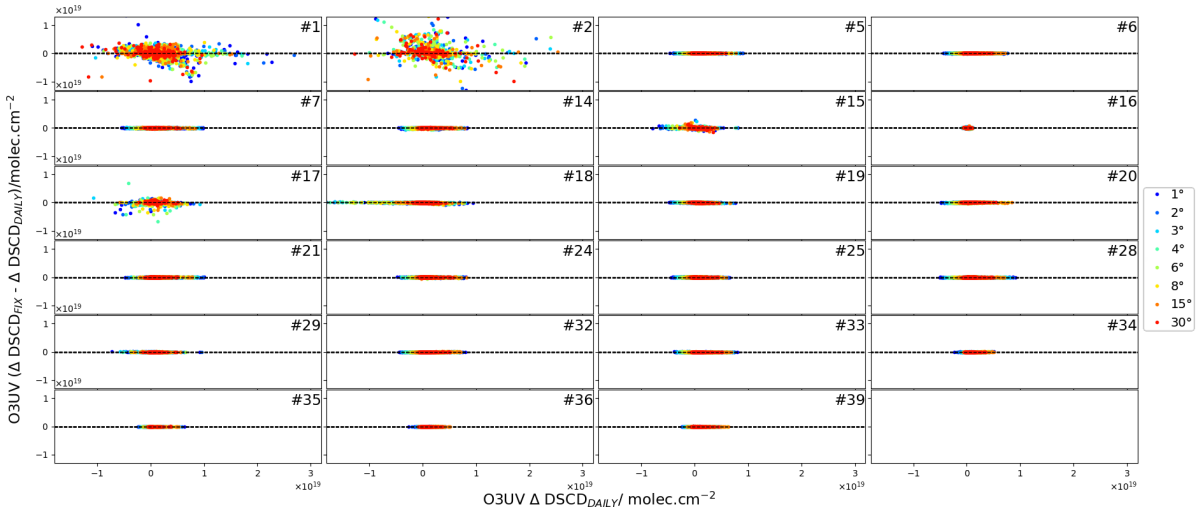

Figure 136: As Fig. 53 but for O3UV

B.5 O4VIS

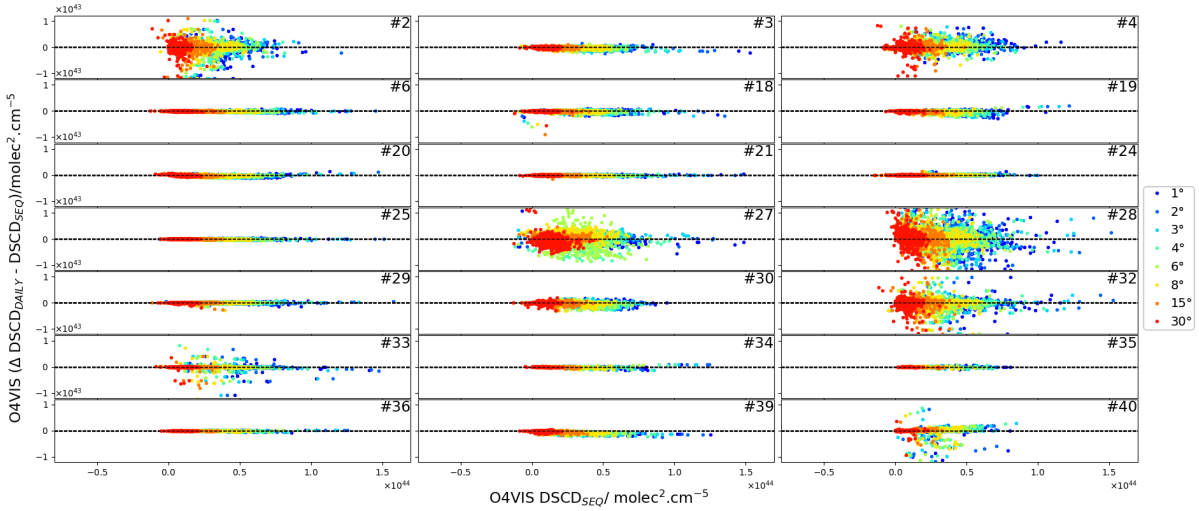

Figure 137: As Fig. 51 but for O4VIS

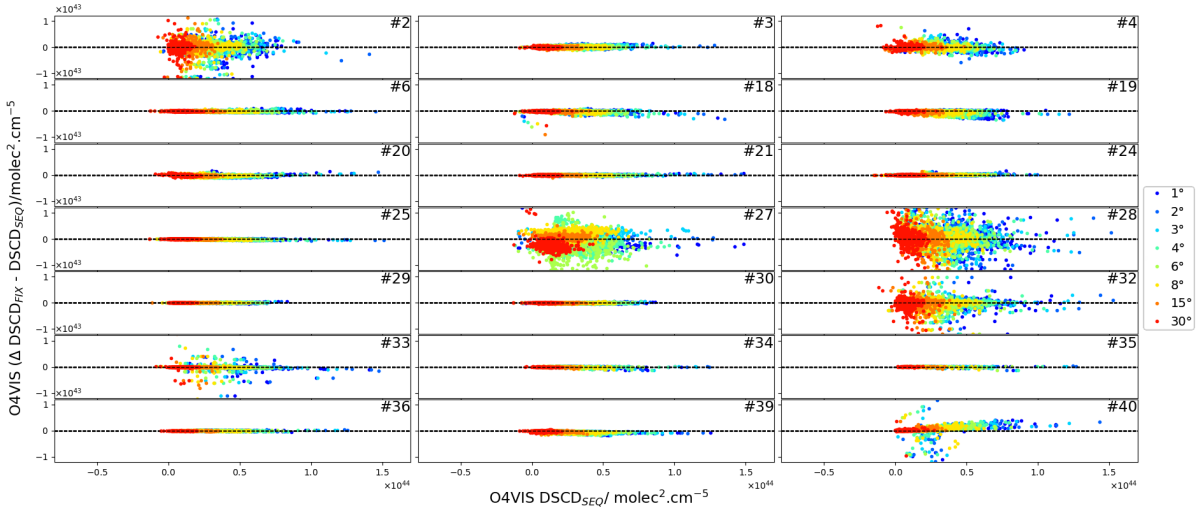

Figure 138: As Fig. 52 but for O4VIS

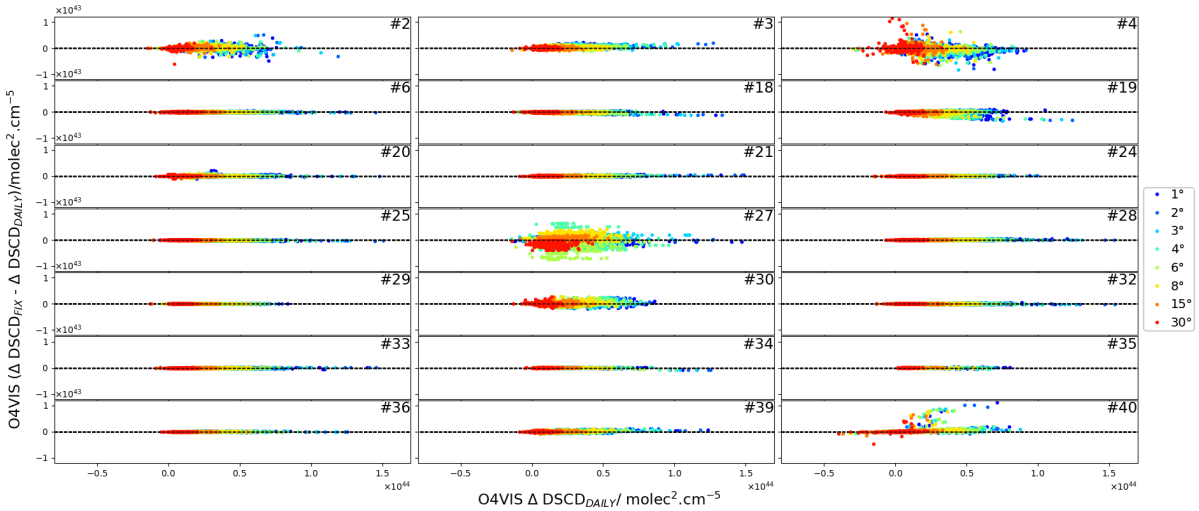

Figure 139: As Fig. 53 but for O4VIS

B.6 O4UV

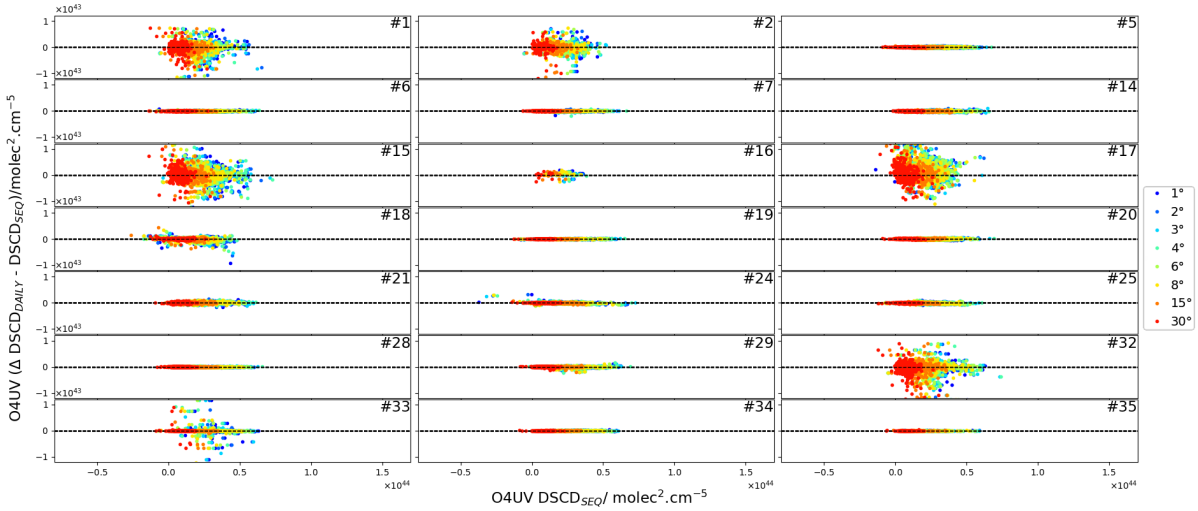

Figure 140: As Fig. 51 but for O4UV

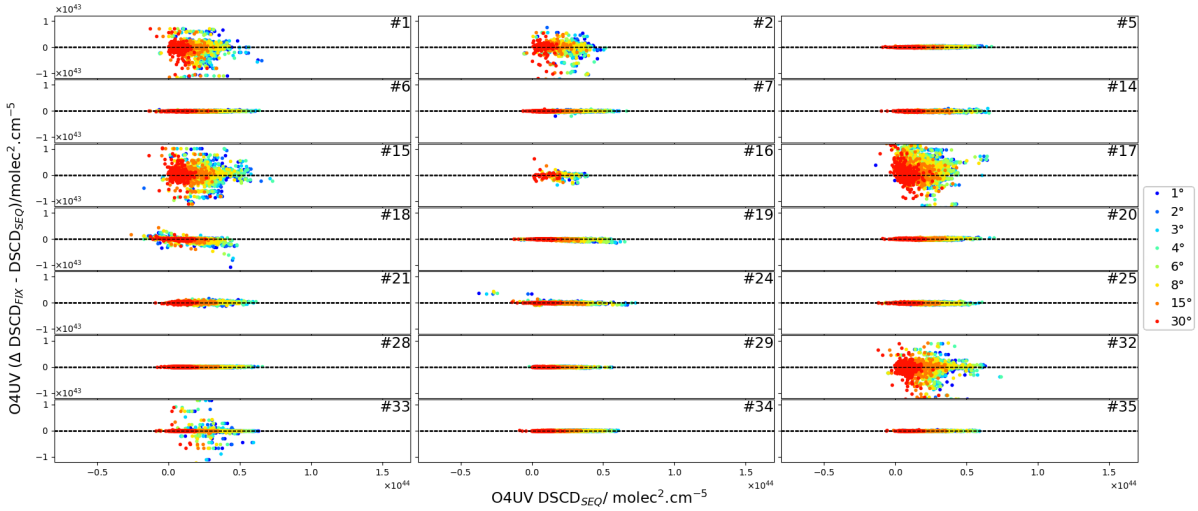

Figure 141: As Fig. 52 but for O4UV

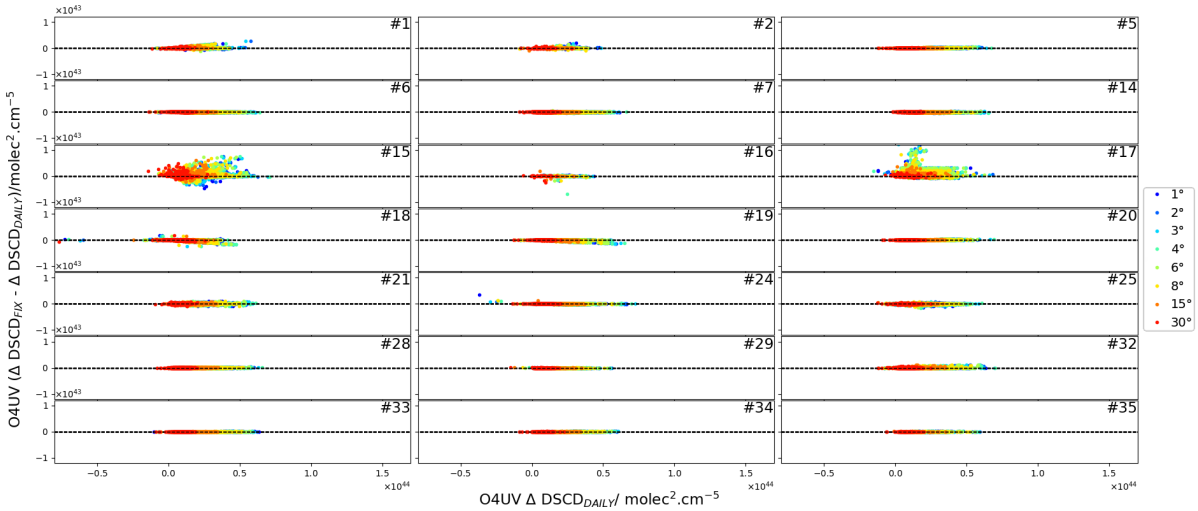

Figure 142: As Fig. 53 but for O4UV

B.7 HCHO

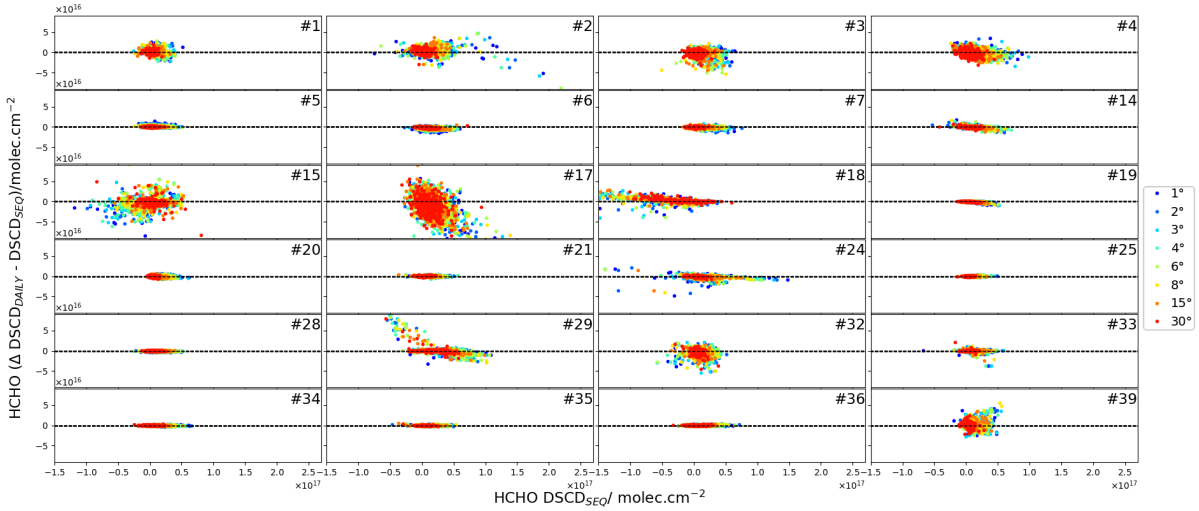

Figure 143: As Fig. 51 but for HCHO

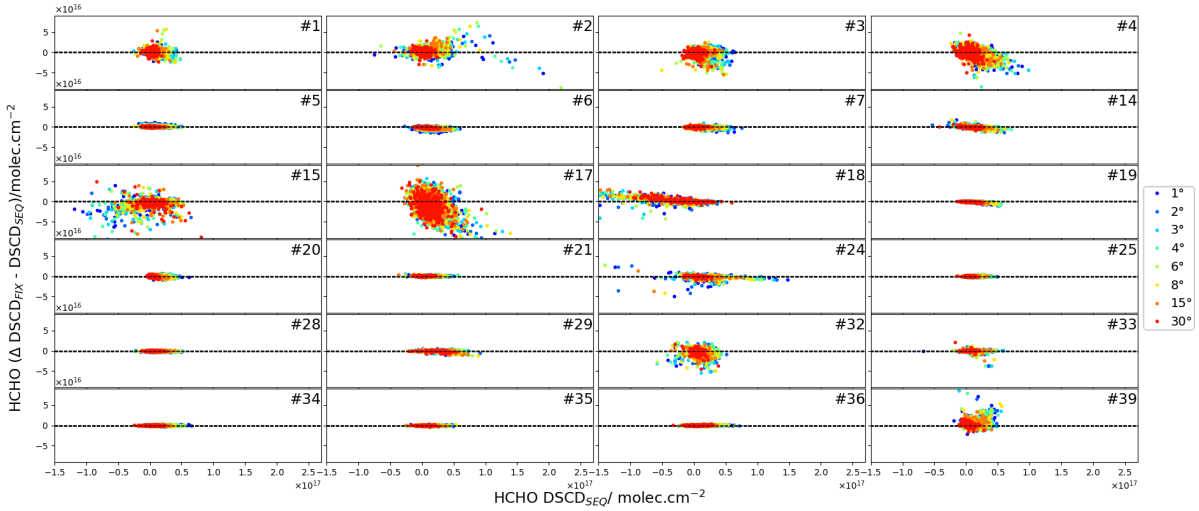

Figure 144: As Fig. 52 but for HCHO

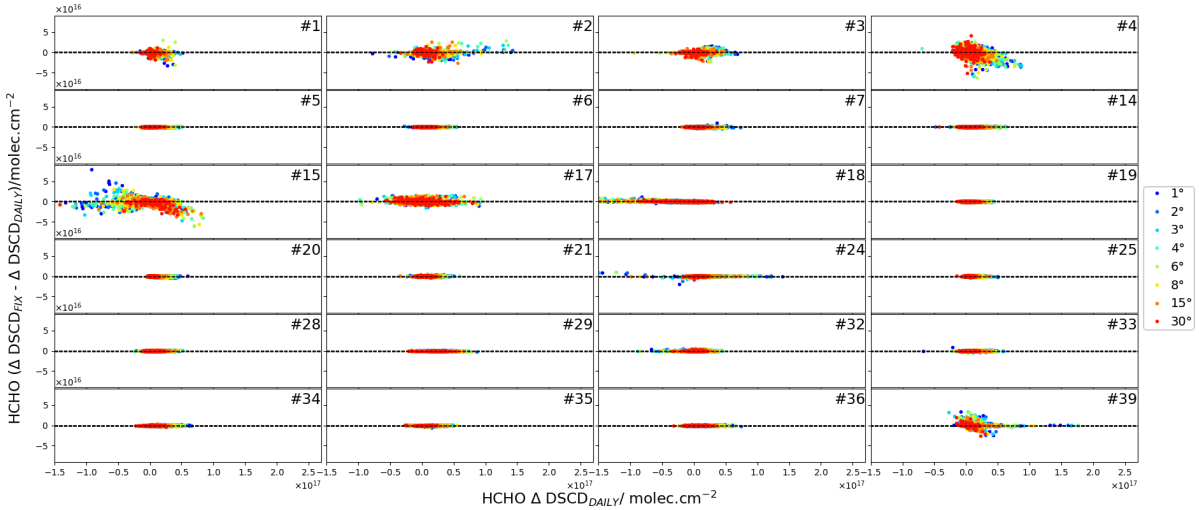

Figure 145: As Fig. 53 but for HCHO

B.8 HCHO-WIDE

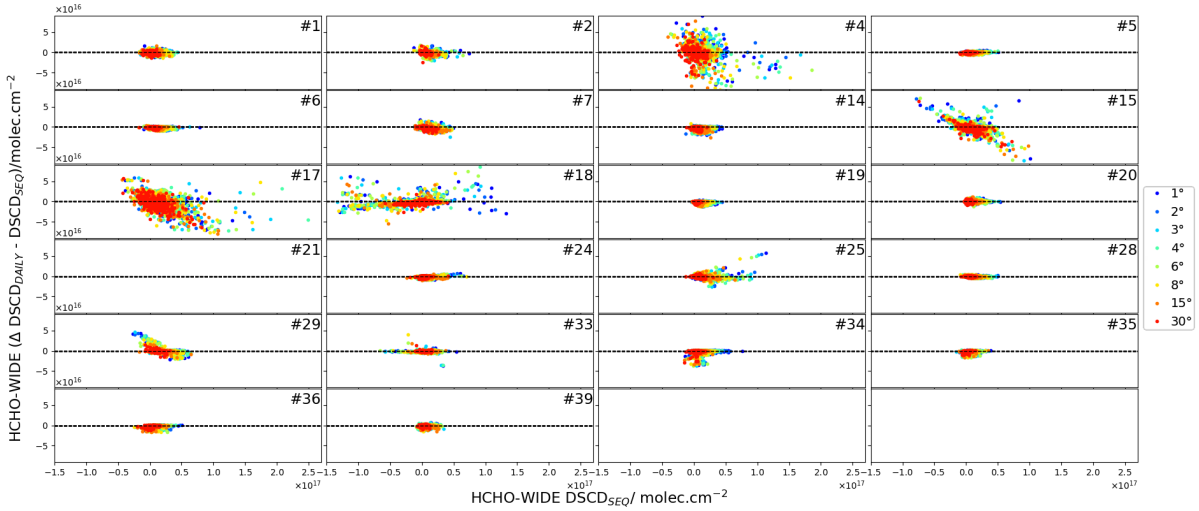

Figure 146: As Fig. 51 but for HCHO-WIDE

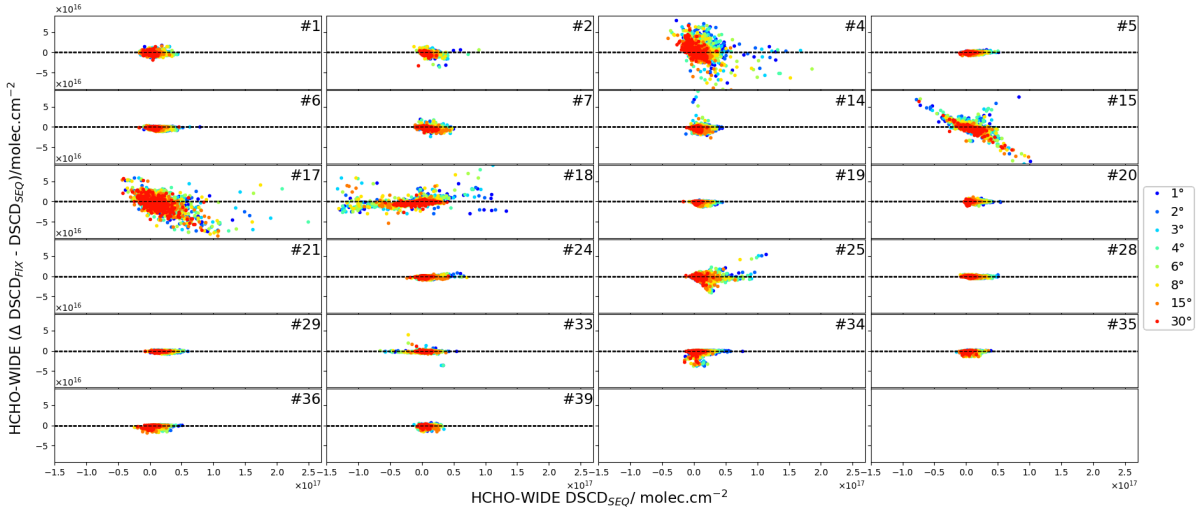

Figure 147: As Fig. 52 but for HCHO-WIDE

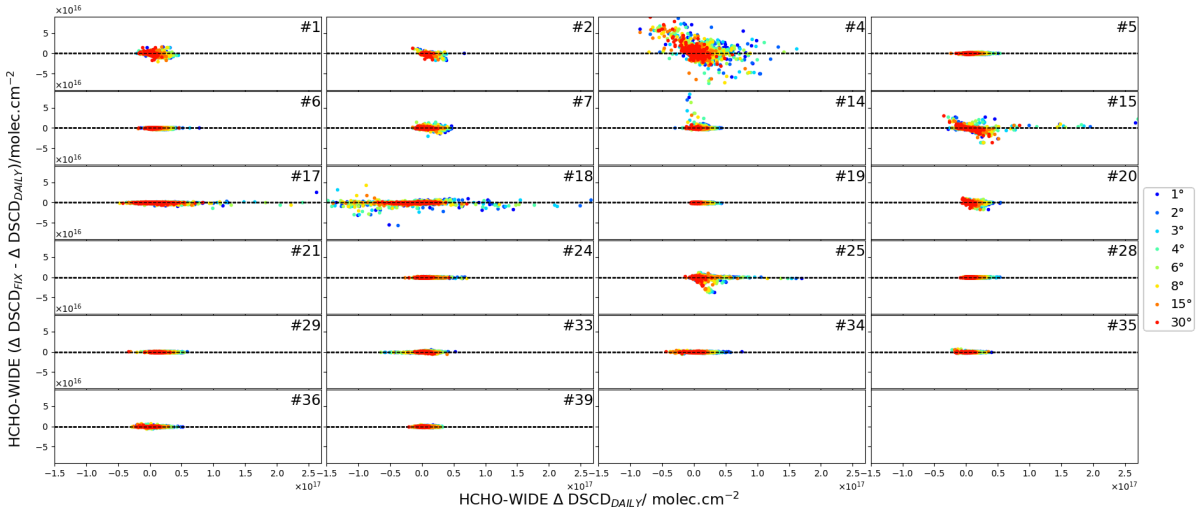

Figure 148: As Fig. 53 but for HCHO-WIDE

B.9 HONO

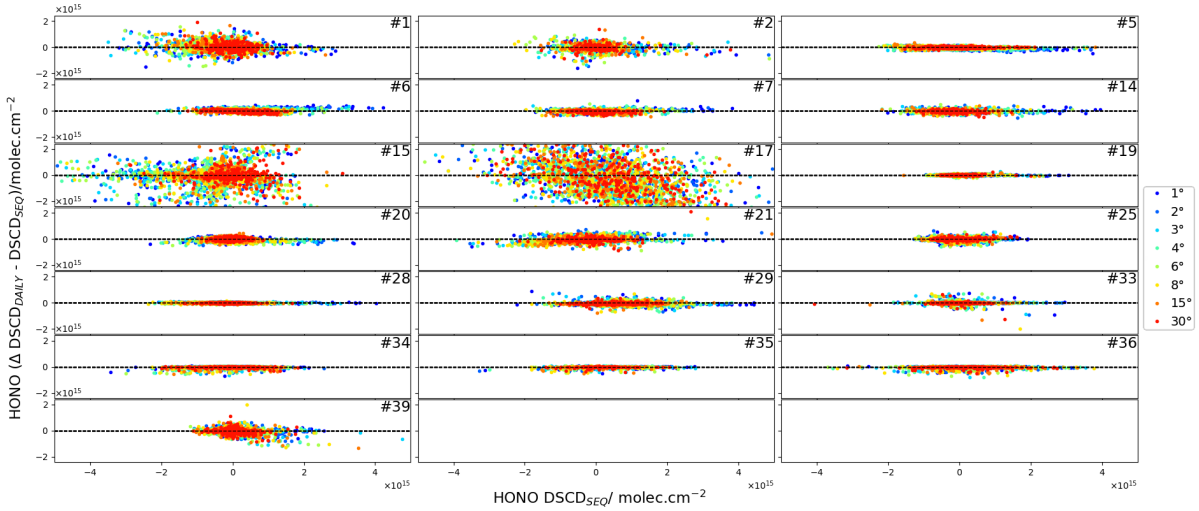

Figure 149: As Fig. 51 but for HONO

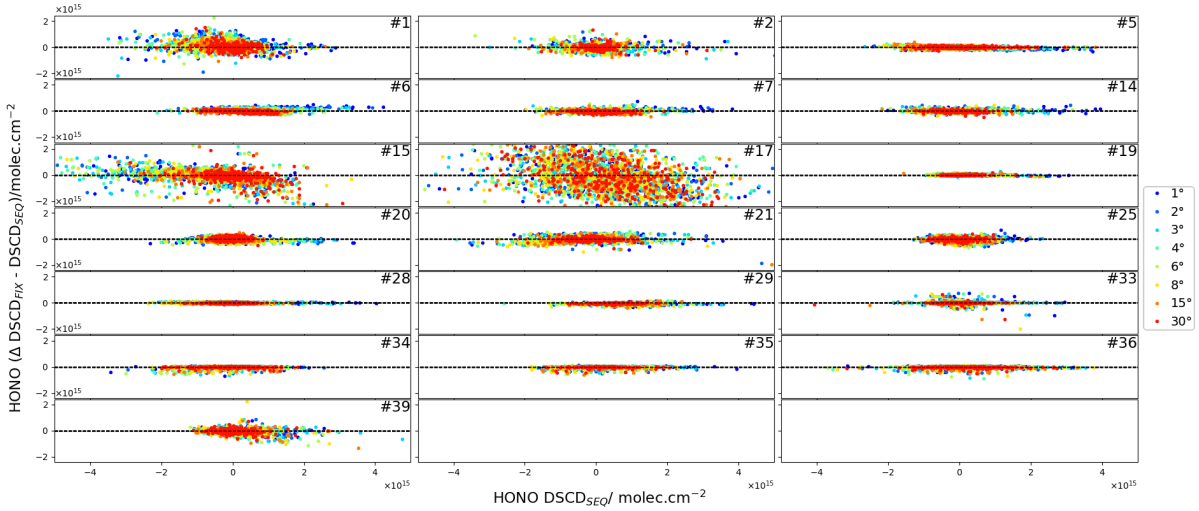

Figure 150: As Fig. 52 but for HONO

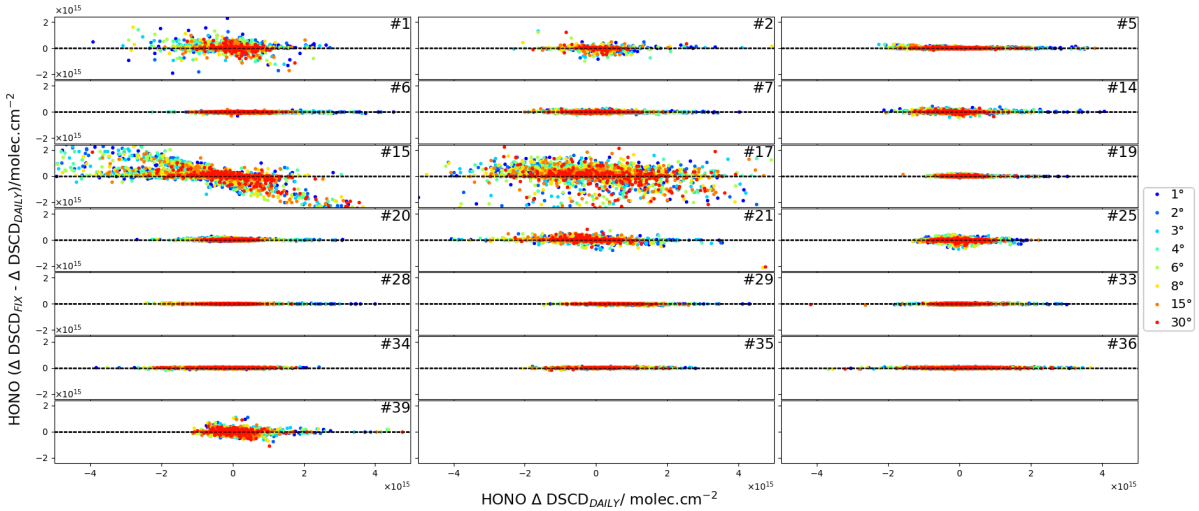

Figure 151: As Fig. 53 but for HONO

C FRM4DOAS comparison plots

C.1 NO2VIS-SMALL

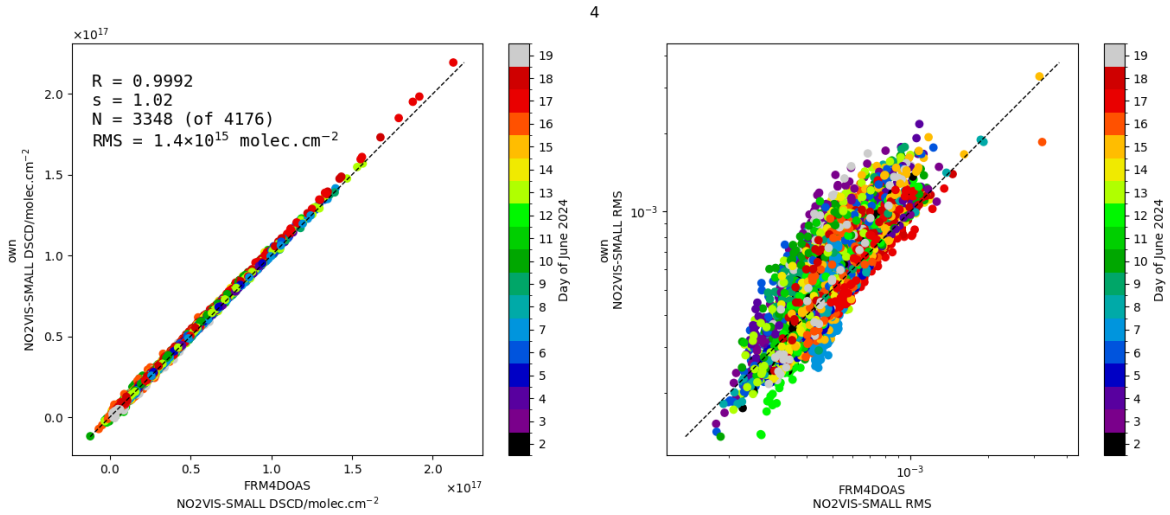

Figure 152: As Fig. 77 but for NO2VIS-SMALL and instrument #4

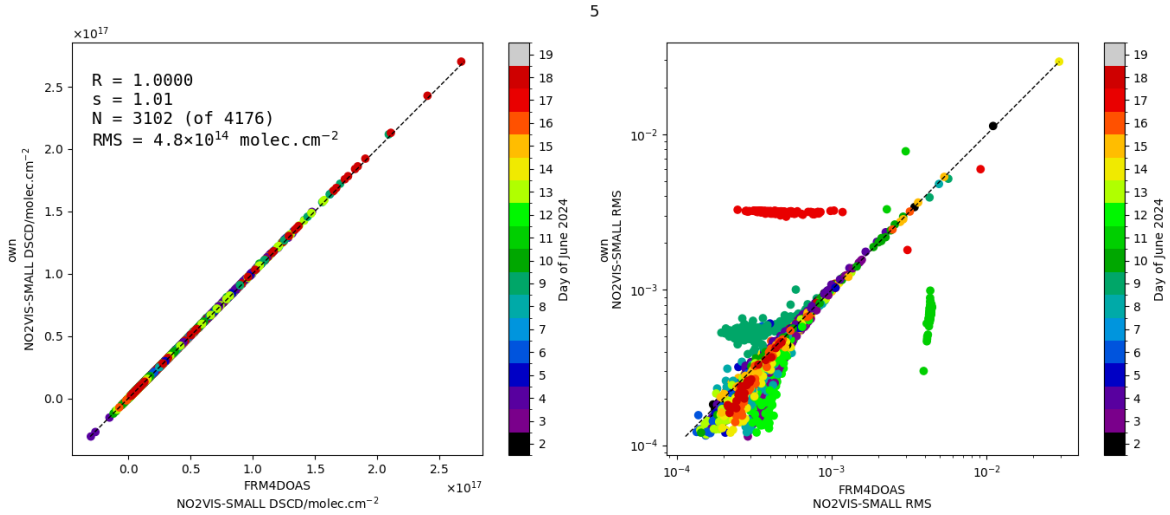

Figure 153: As Fig. 77 but for NO2VIS-SMALL and instrument #5

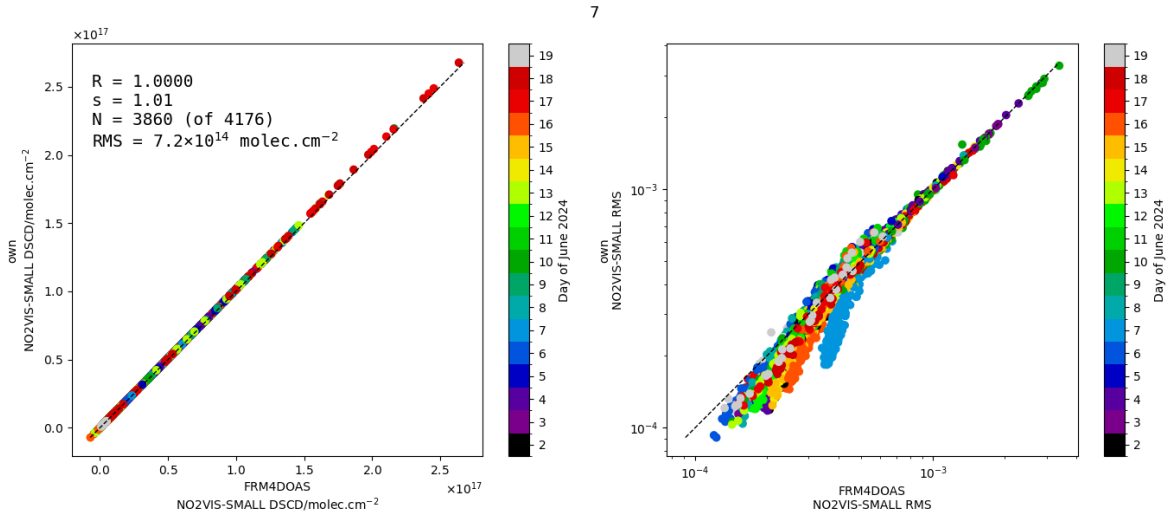

Figure 154: As Fig. 77 but for NO2VIS-SMALL and instrument #7

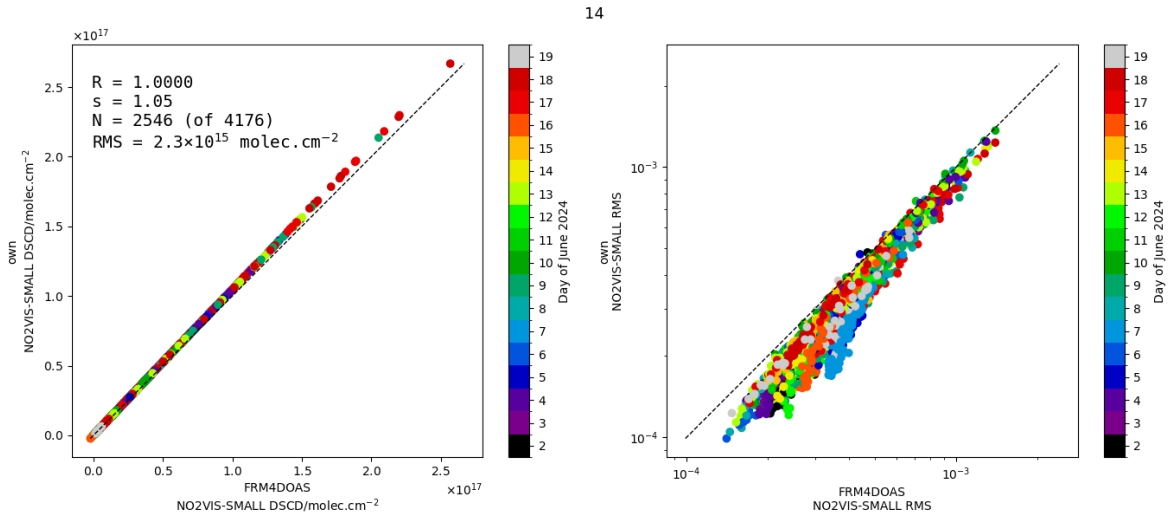

Figure 155: As Fig. 77 but for NO2VIS-SMALL and instrument #14

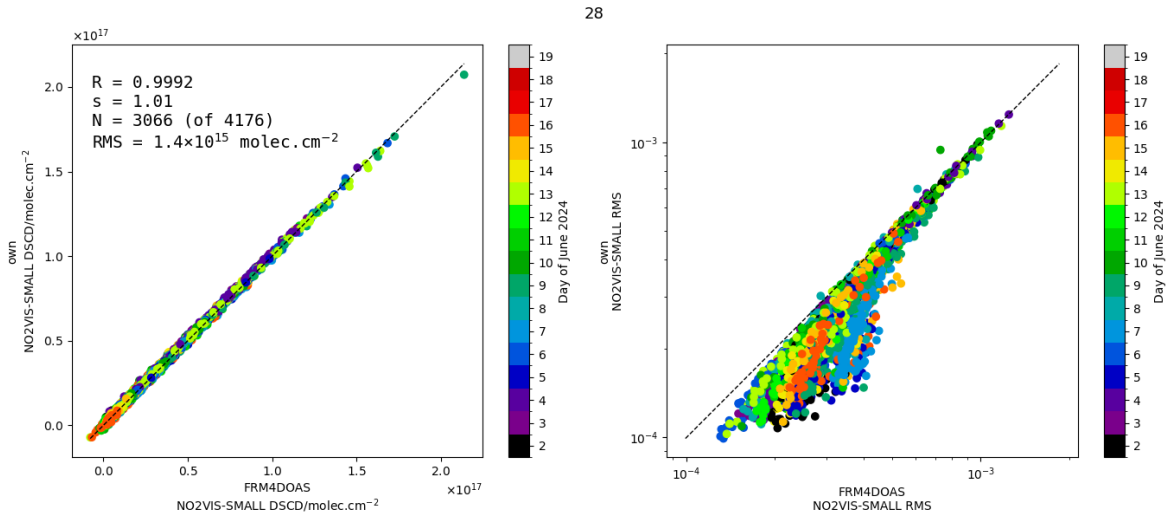

Figure 156: As Fig. 77 but for NO2VIS-SMALL and instrument #28

C.2 NO2UV

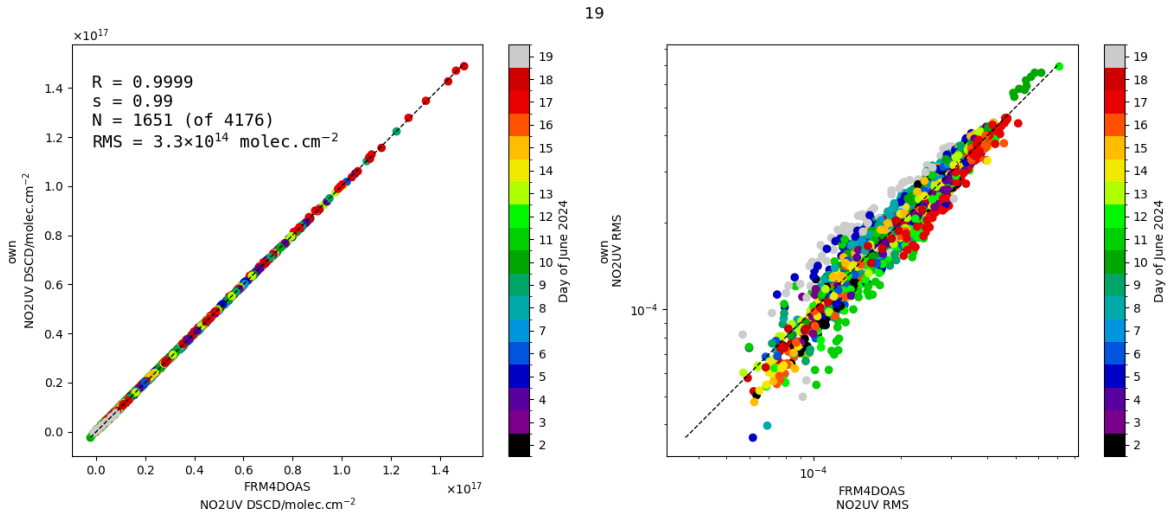

Figure 157: As Fig. 77 but for NO2UV and instrument #19

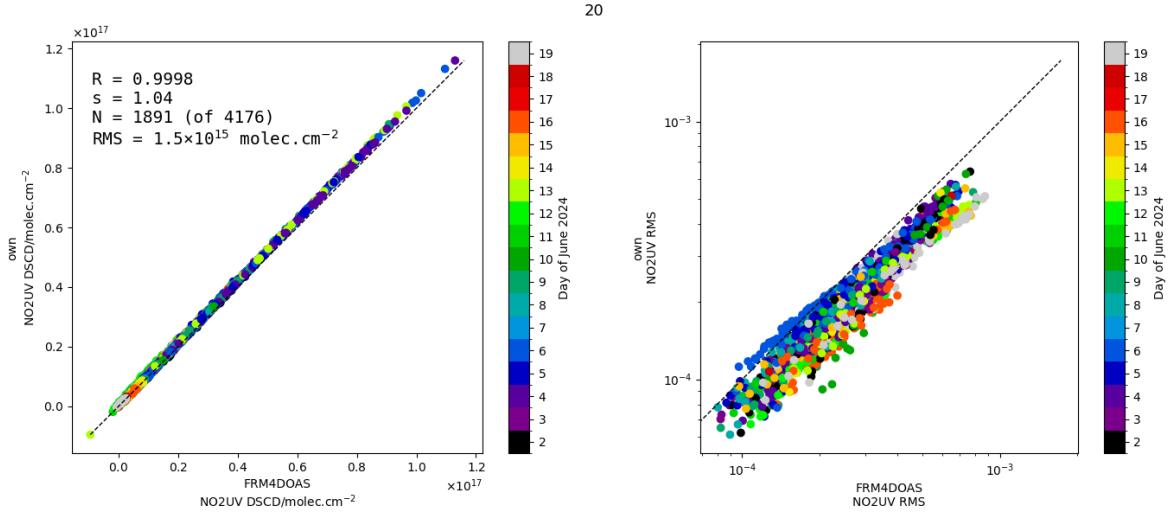

Figure 158: As Fig. 77 but for NO2UV and instrument #20

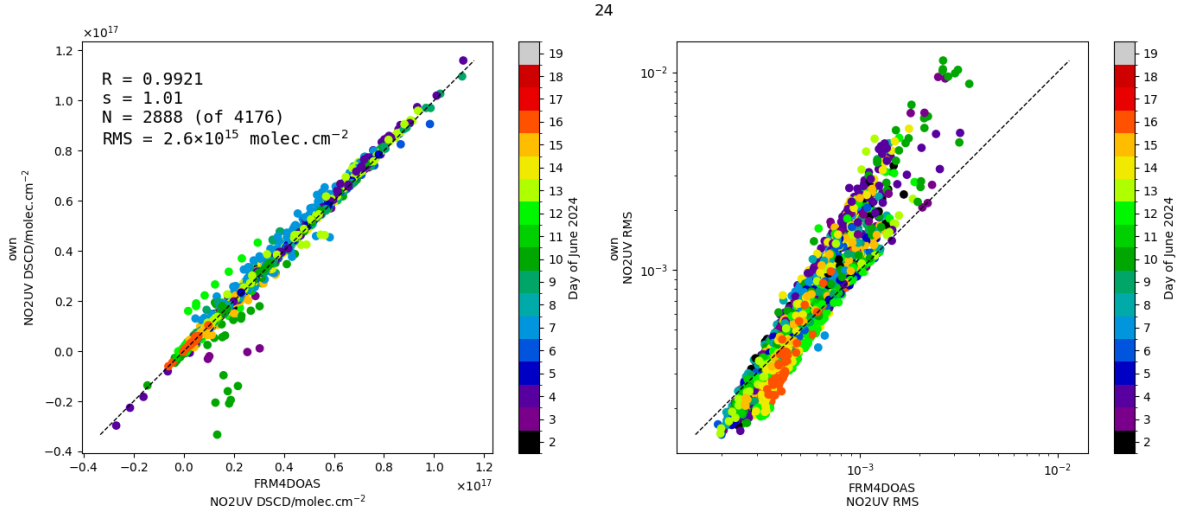

Figure 159: As Fig. 77 but for NO2UV and instrument #24

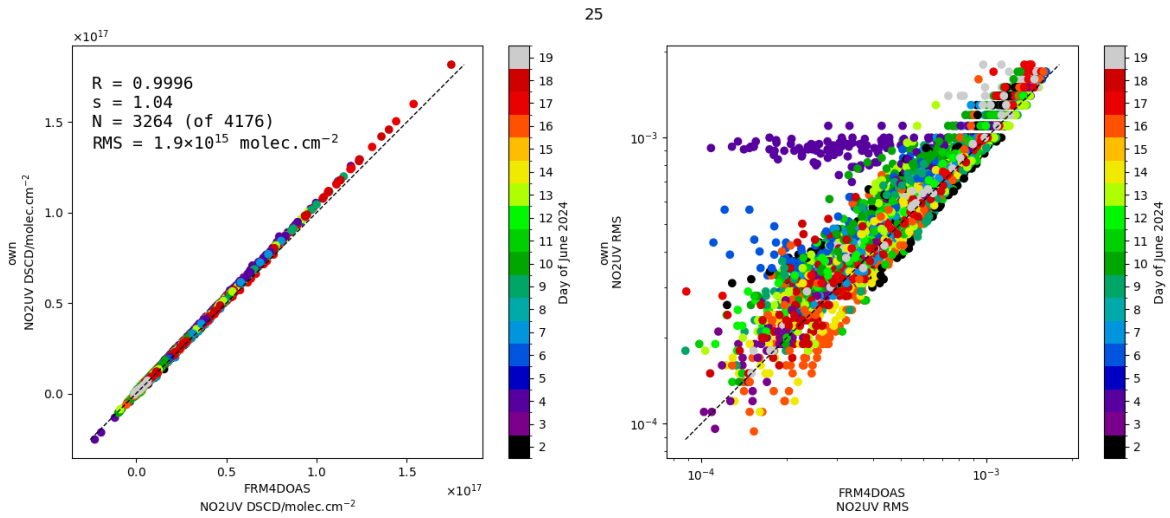

Figure 160: As Fig. 77 but for NO2UV and instrument #25

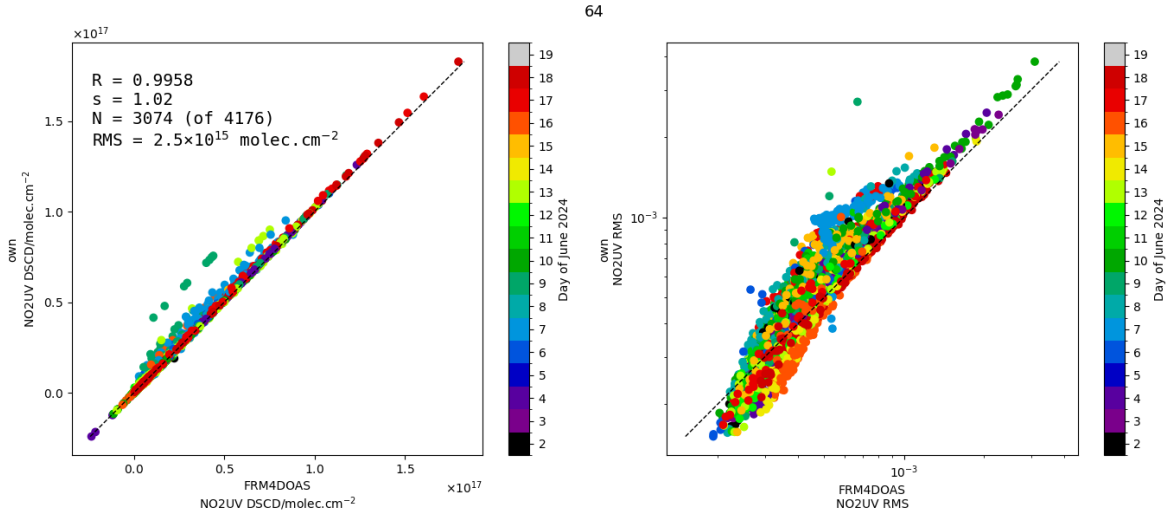

Figure 161: As Fig. 77 but for NO2UV and instrument #32

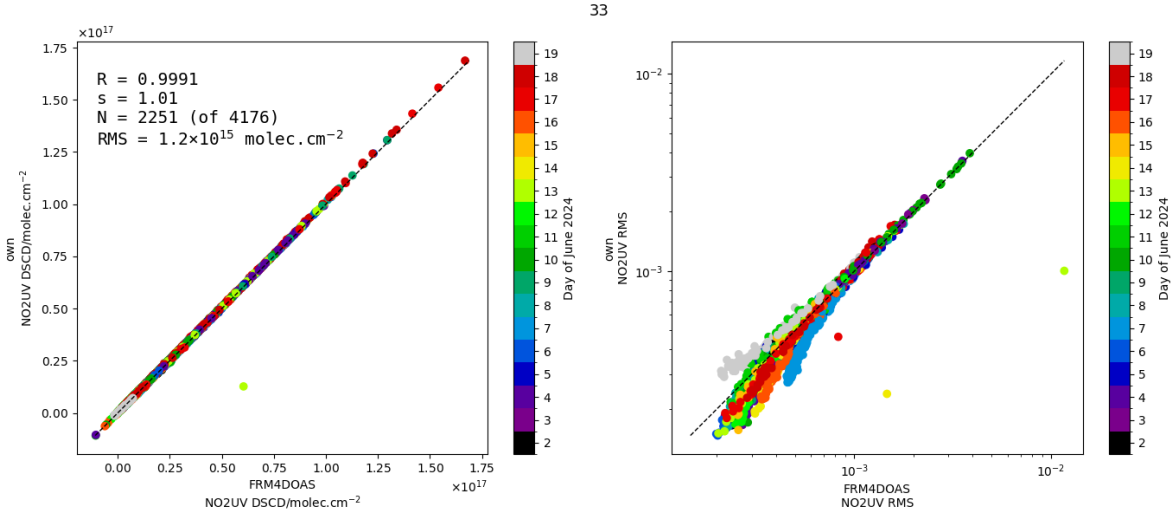

Figure 162: As Fig. 77 but for NO2UV and instrument #33

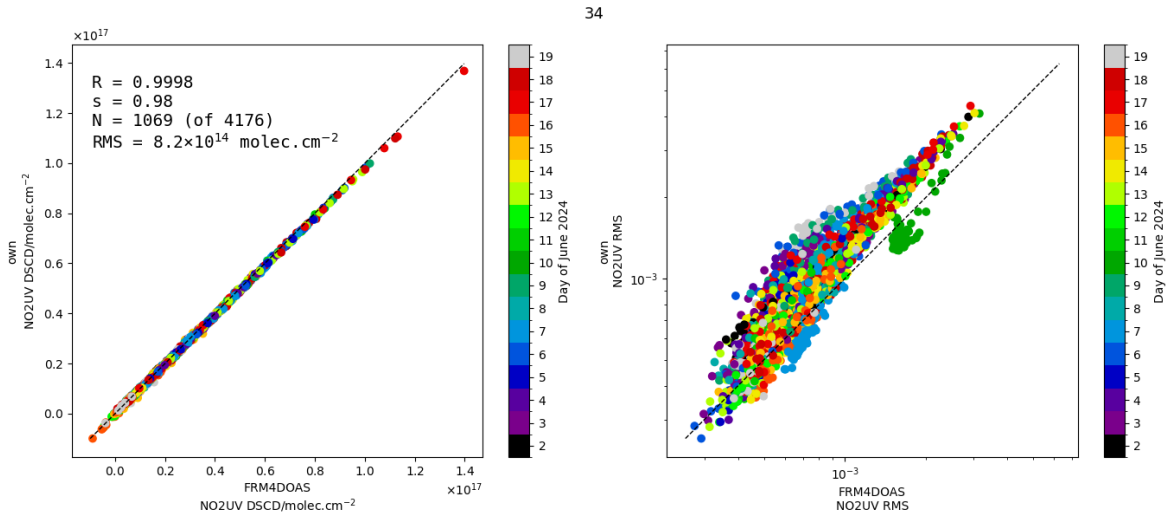

Figure 163: As Fig. 77 but for NO2UV and instrument #34

C.3 O4VIS

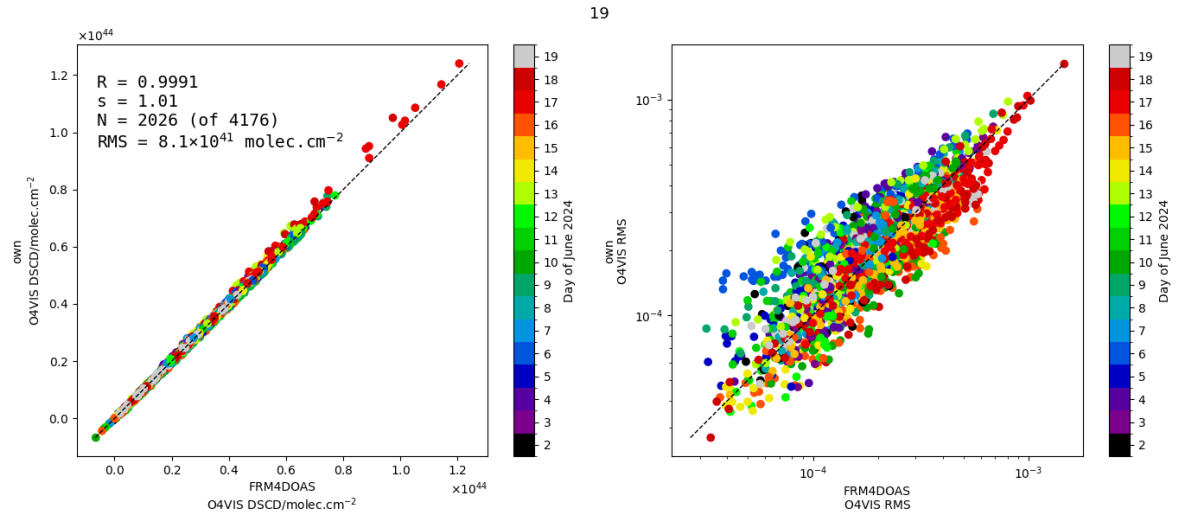

Figure 164: As Fig. 77 but for O4VIS and instrument #19

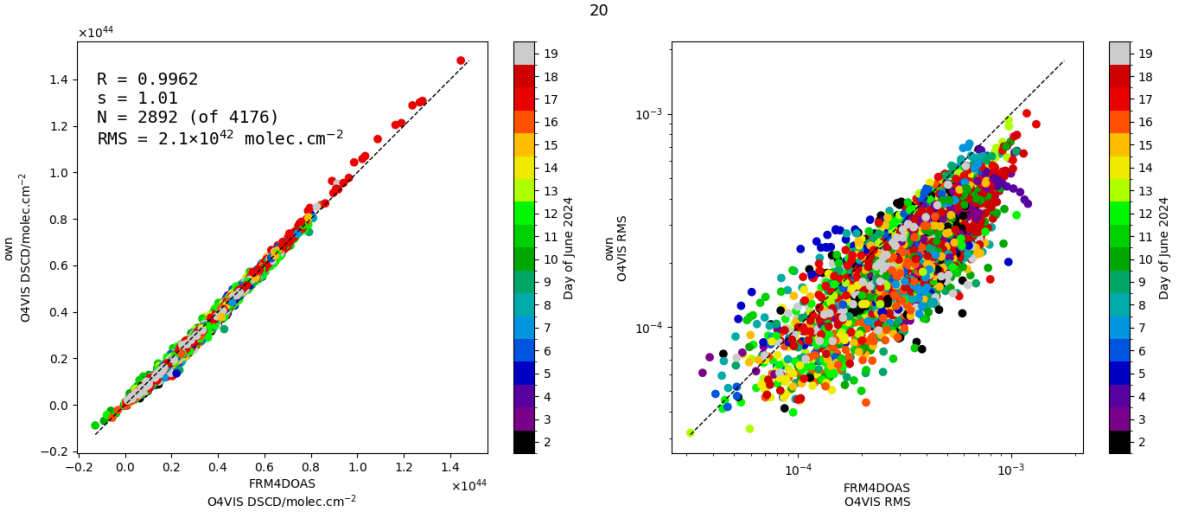

Figure 165: As Fig. 77 but for O4VIS and instrument #20

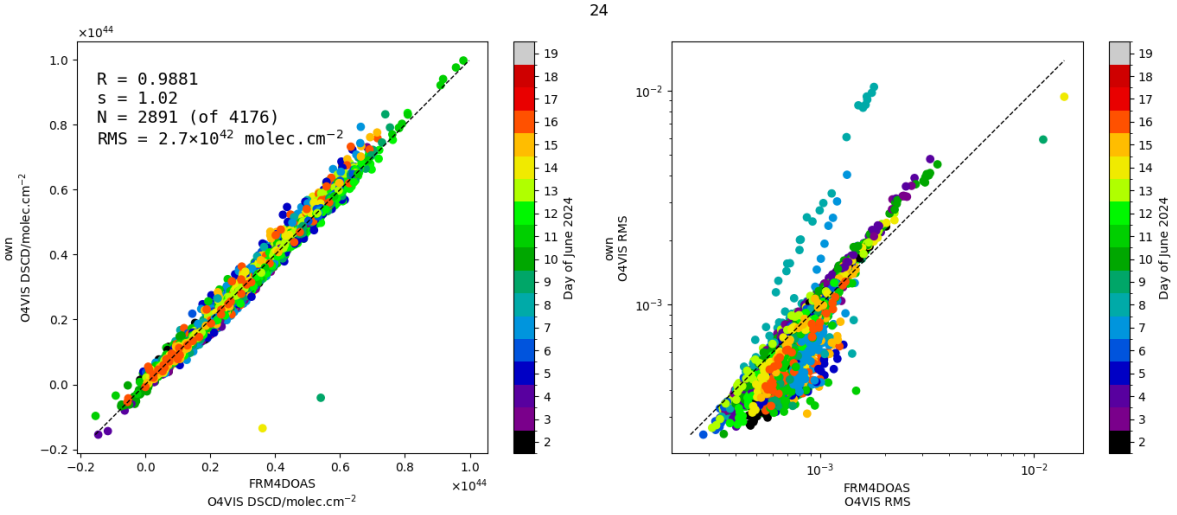

Figure 166: As Fig. 77 but for O4VIS and instrument #24

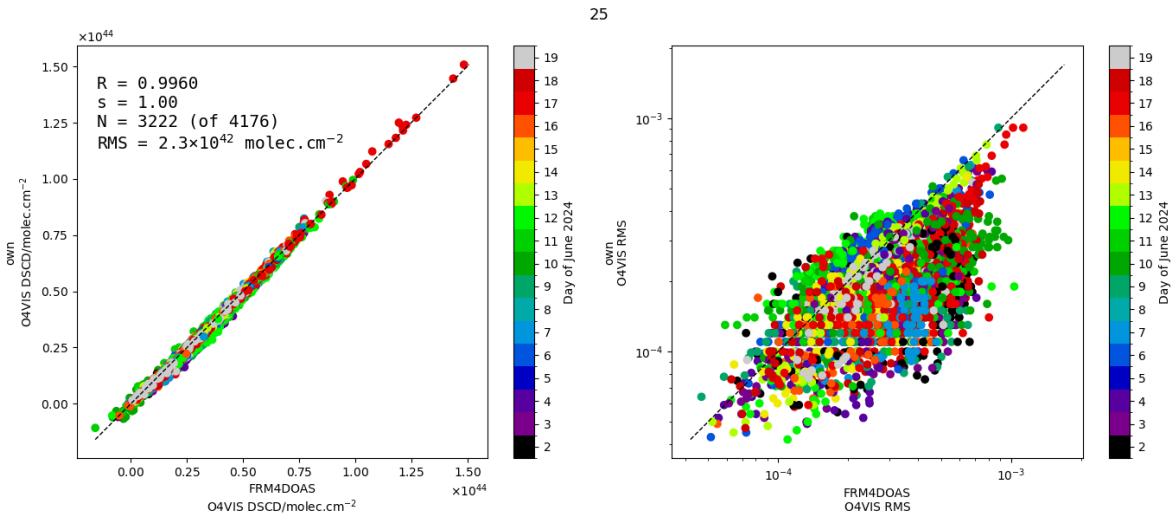

Figure 167: As Fig. 77 but for O4VIS and instrument #25

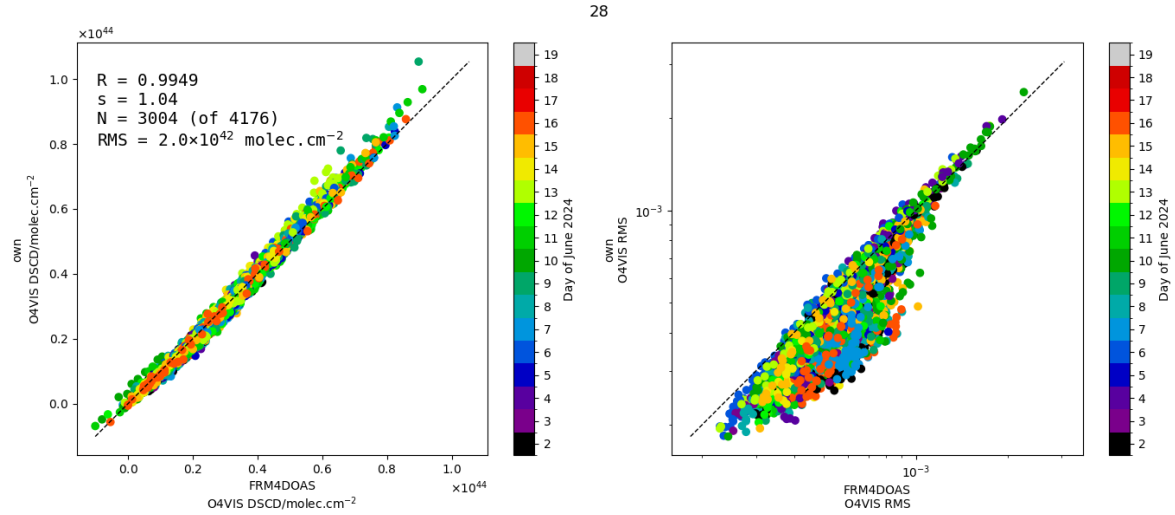

Figure 168: As Fig. 77 but for O4VIS and instrument #28

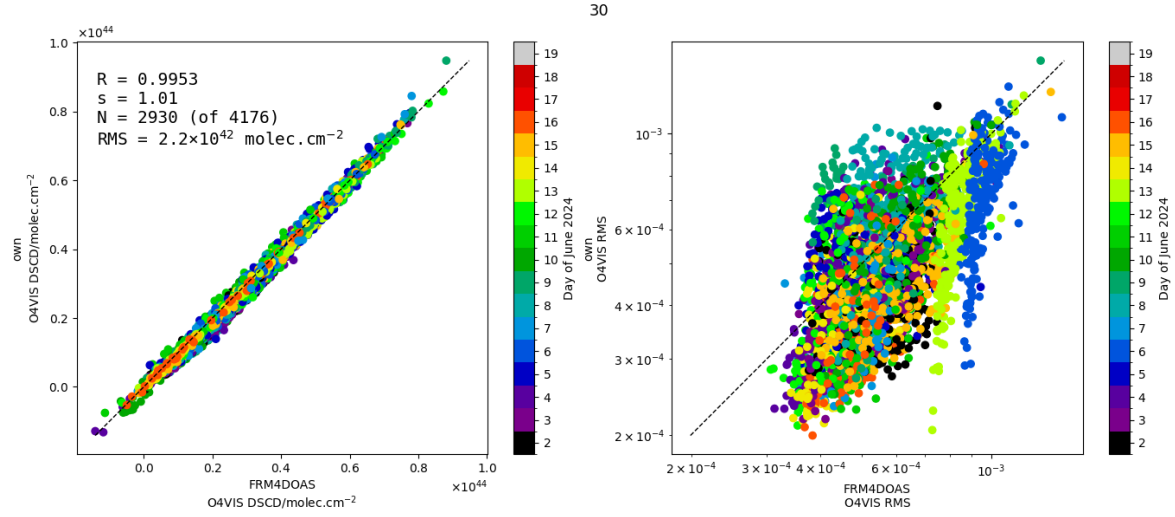

Figure 169: As Fig. 77 but for O4VIS and instrument #30

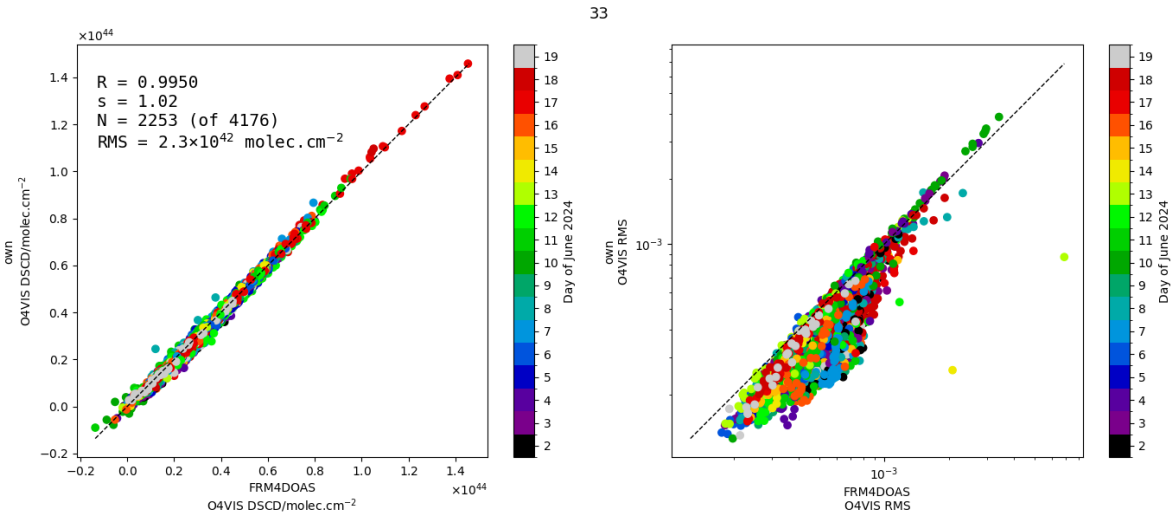

Figure 170: As Fig. 77 but for O4VIS and instrument #33

C.4 O4UV

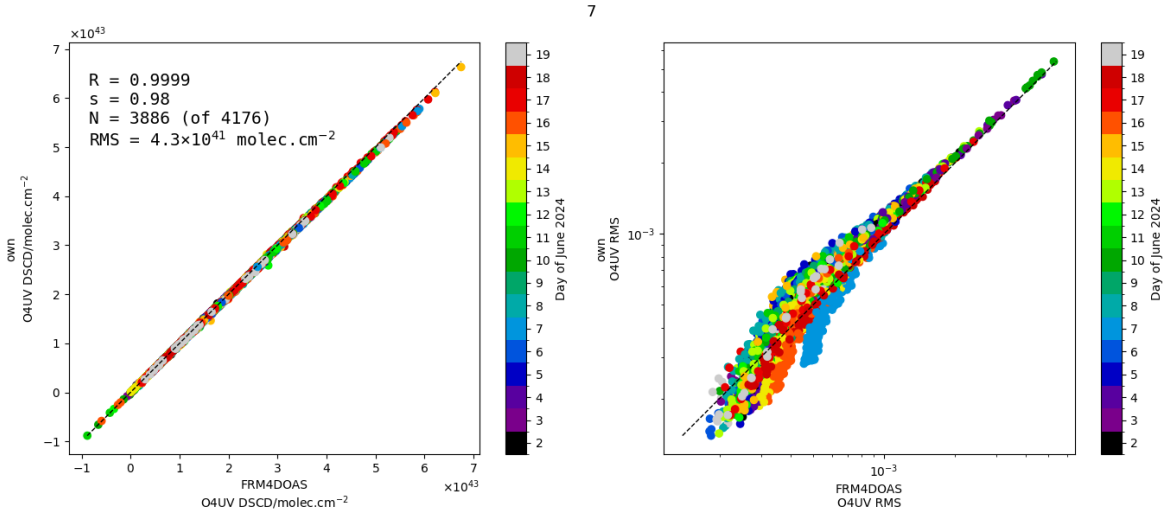

Figure 171: As Fig. 77 but for O4UV and instrument #7

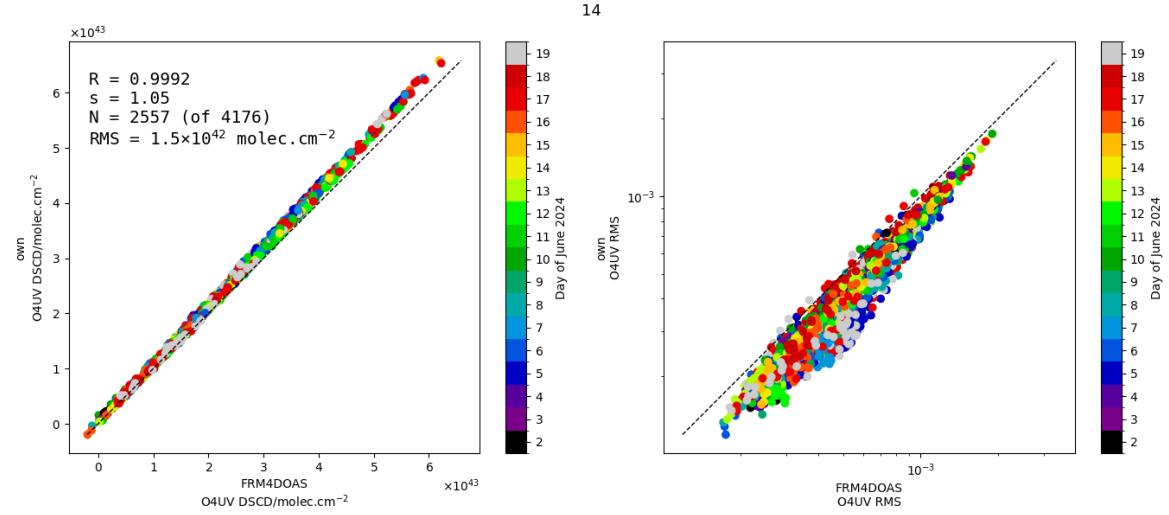

Figure 172: As Fig. 77 but for O4UV and instrument #14

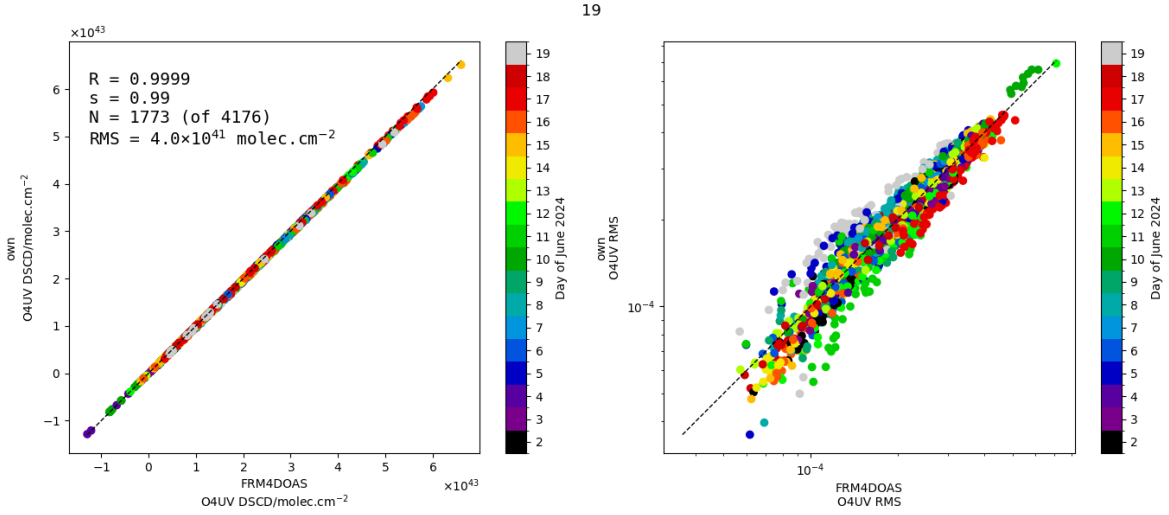

Figure 173: As Fig. 77 but for O4UV and instrument #19

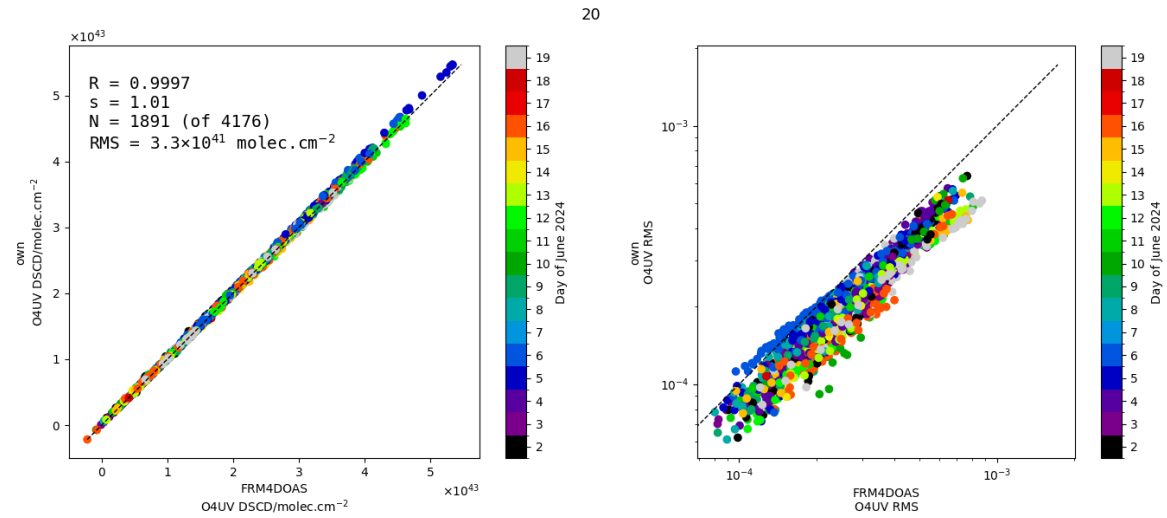

Figure 174: As Fig. 77 but for O4UV and instrument #20

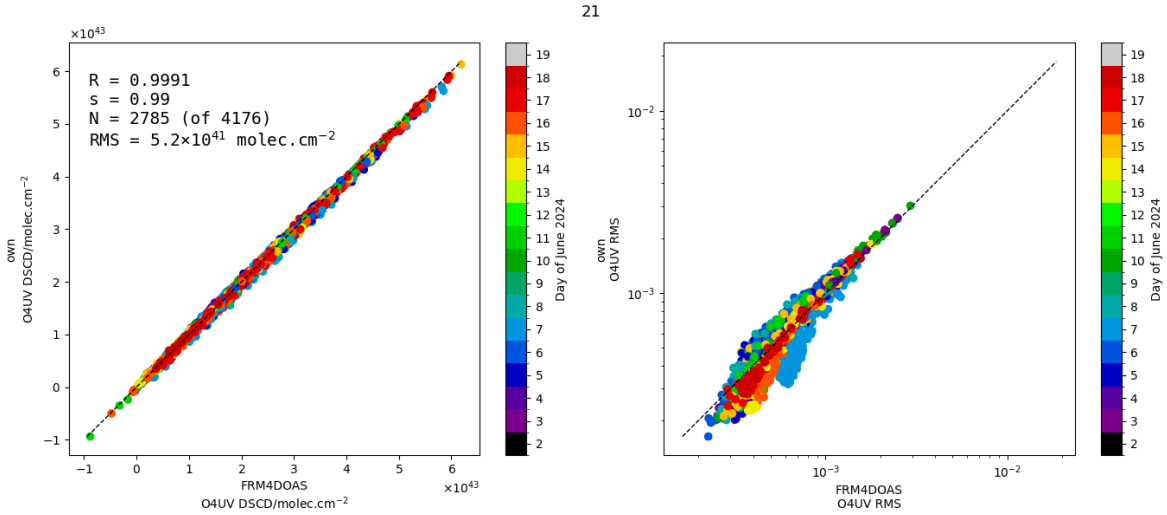

Figure 175: As Fig. 77 but for O4UV and instrument #21

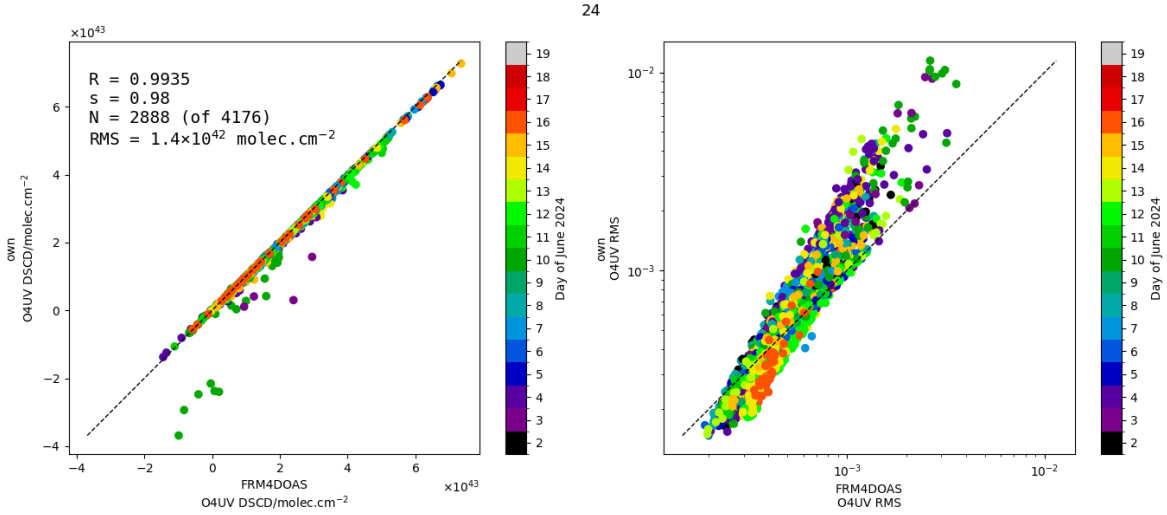

Figure 176: As Fig. 77 but for O4UV and instrument #24

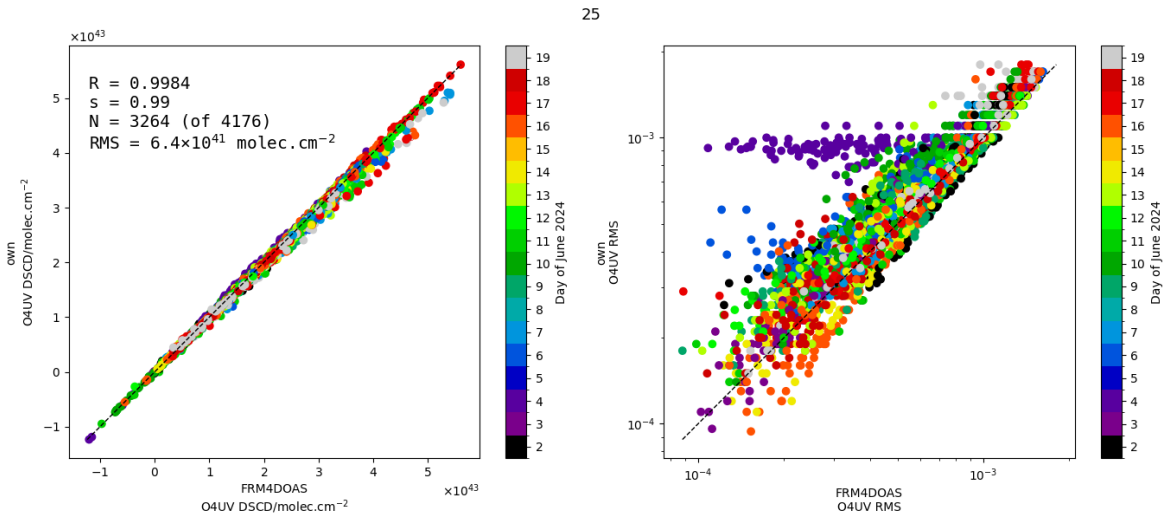

Figure 177: As Fig. 77 but for O4UV and instrument #25

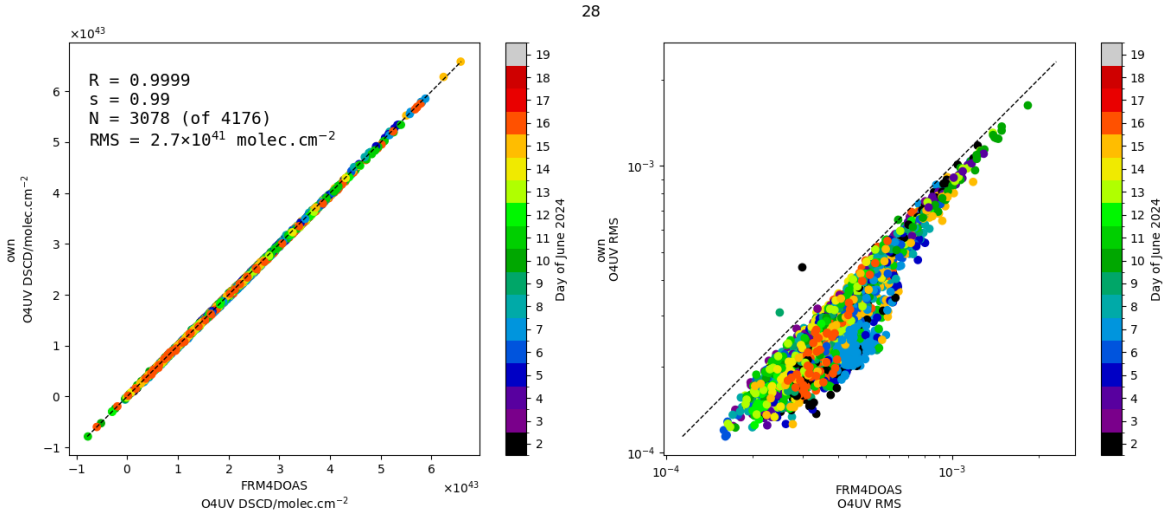

Figure 178: As Fig. 77 but for O4UV and instrument #28

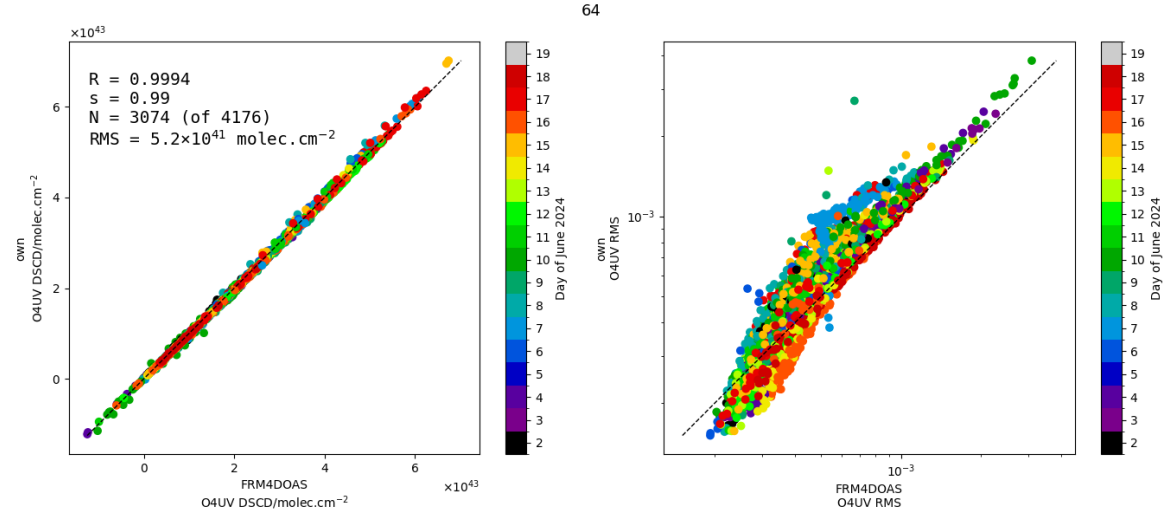

Figure 179: As Fig. 77 but for O4UV and instrument #32

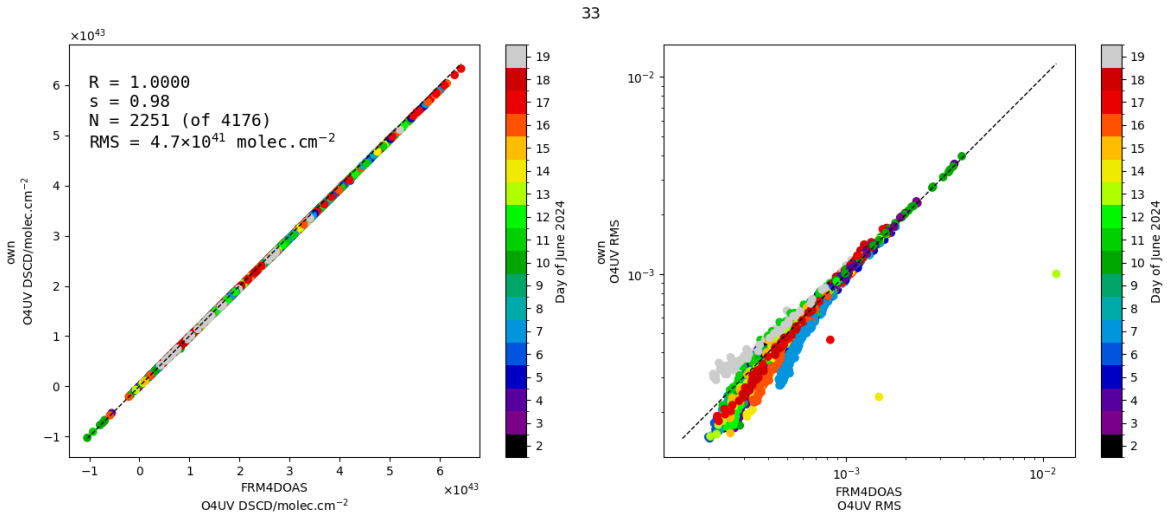

Figure 180: As Fig. 77 but for O4UV and instrument #33

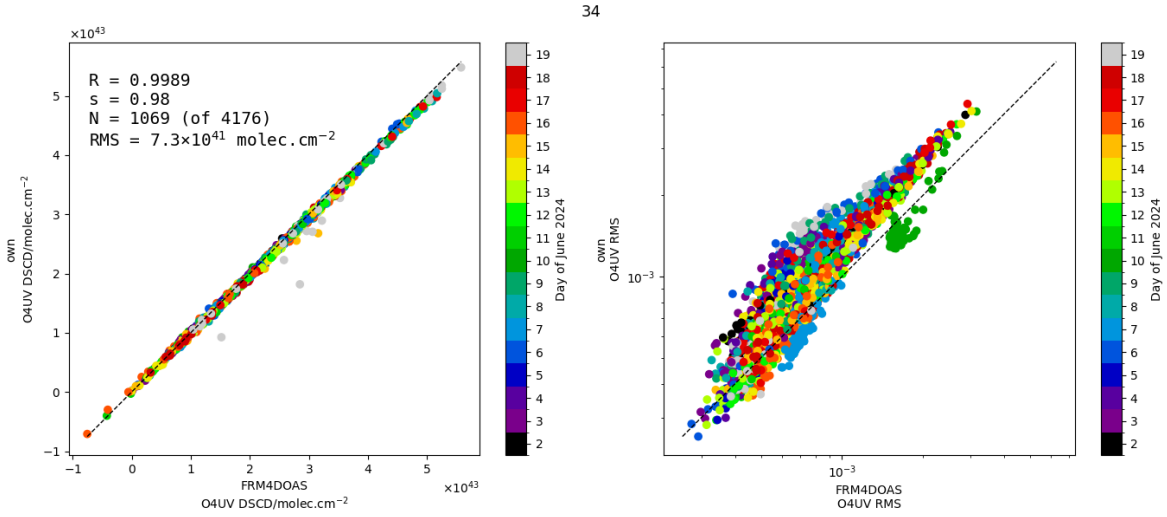

Figure 181: As Fig. 77 but for O4UV and instrument #34

C.5 HCHO-WIDE

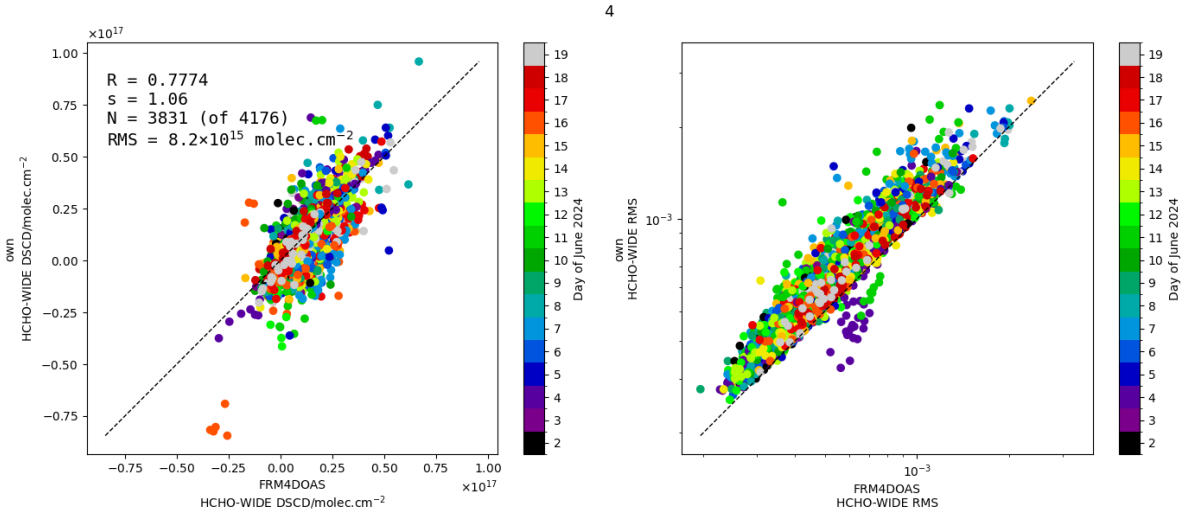

Figure 182: As Fig. 77 but for HCHO-WIDE and instrument #4

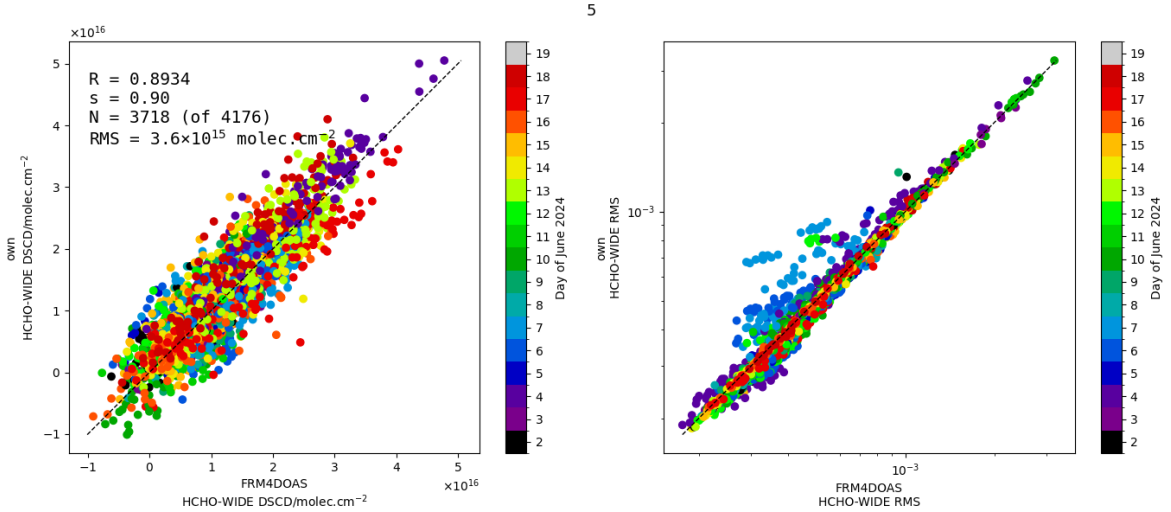

Figure 183: As Fig. 77 but for HCHO-WIDE and instrument #5

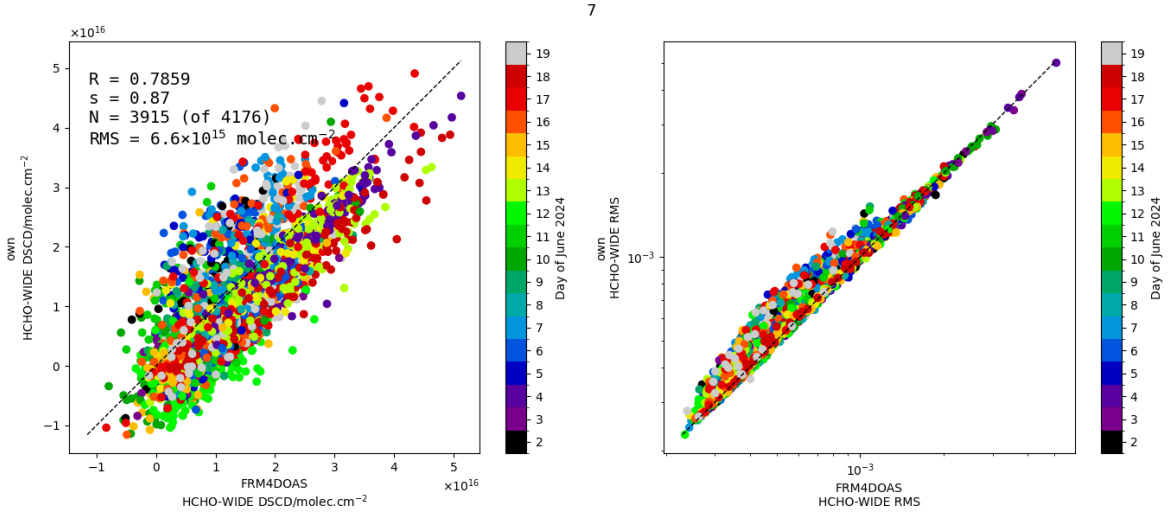

Figure 184: As Fig. 77 but for HCHO-WIDE and instrument #7

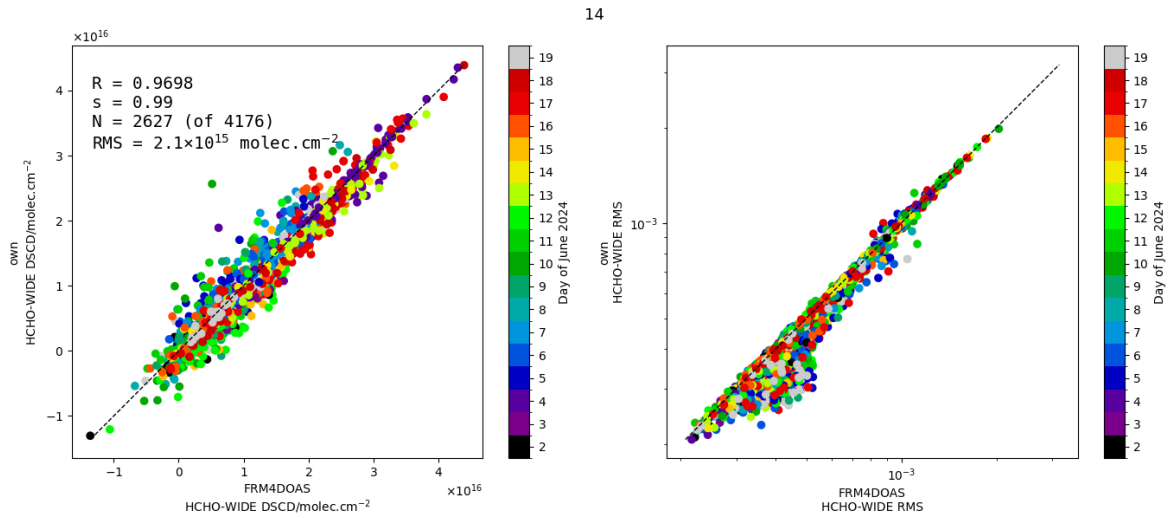

Figure 185: As Fig. 77 but for HCHO-WIDE and instrument #14

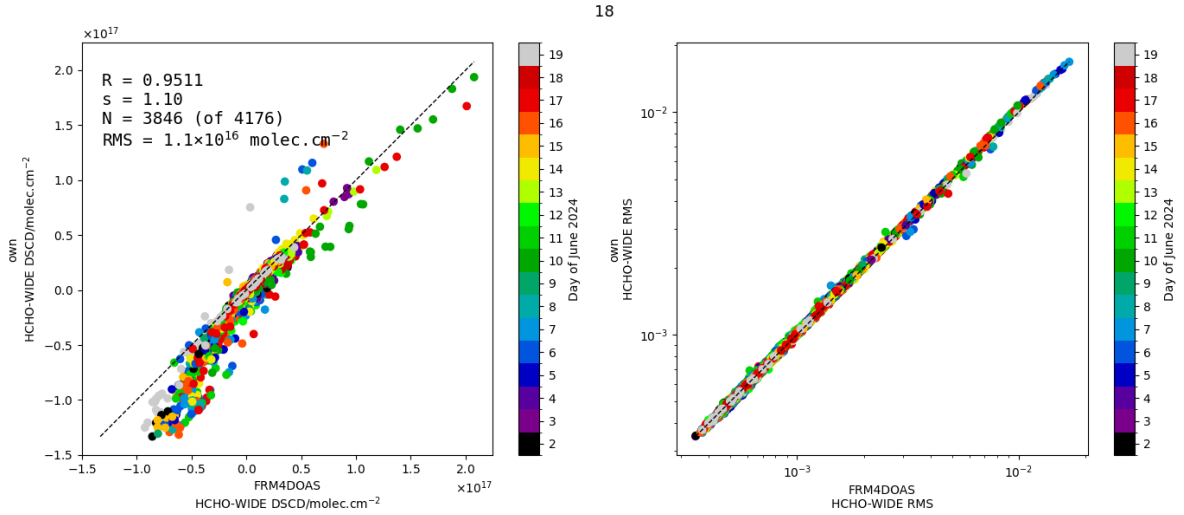

Figure 186: As Fig. 77 but for HCHO-WIDE and instrument #18

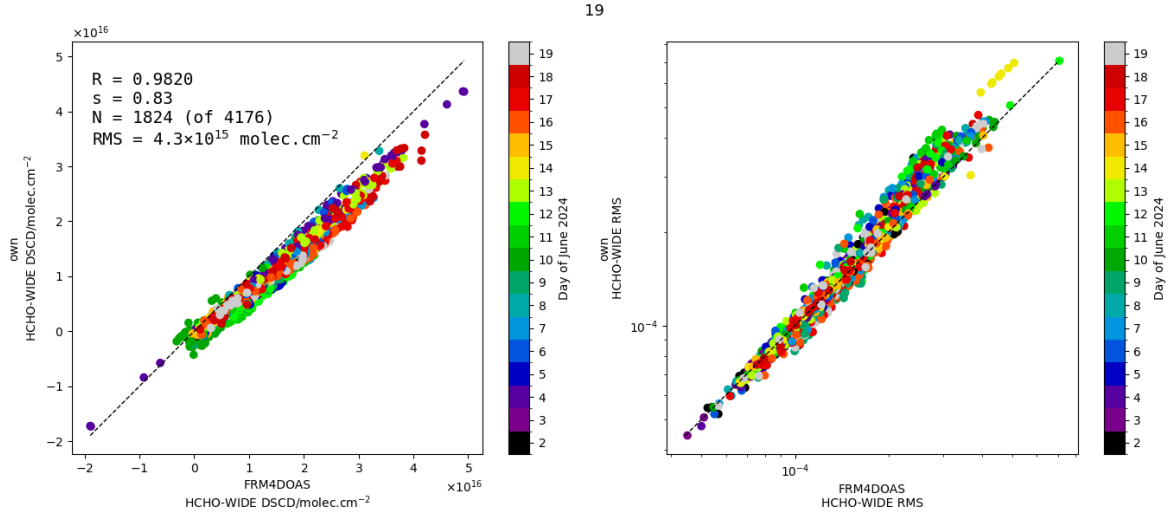

Figure 187: As Fig. 77 but for HCHO-WIDE and instrument #19

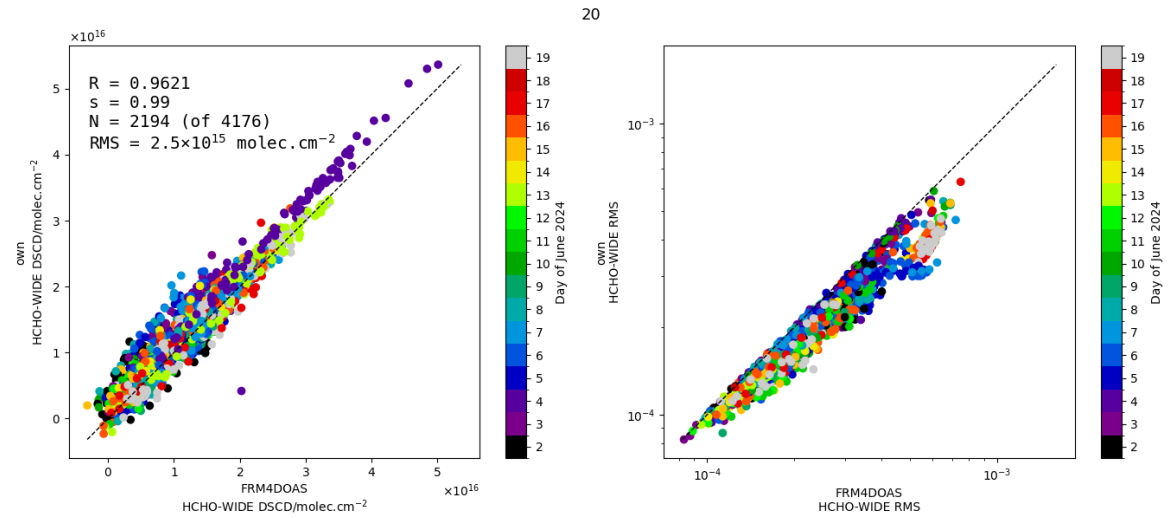

Figure 188: As Fig. 77 but for HCHO-WIDE and instrument #20

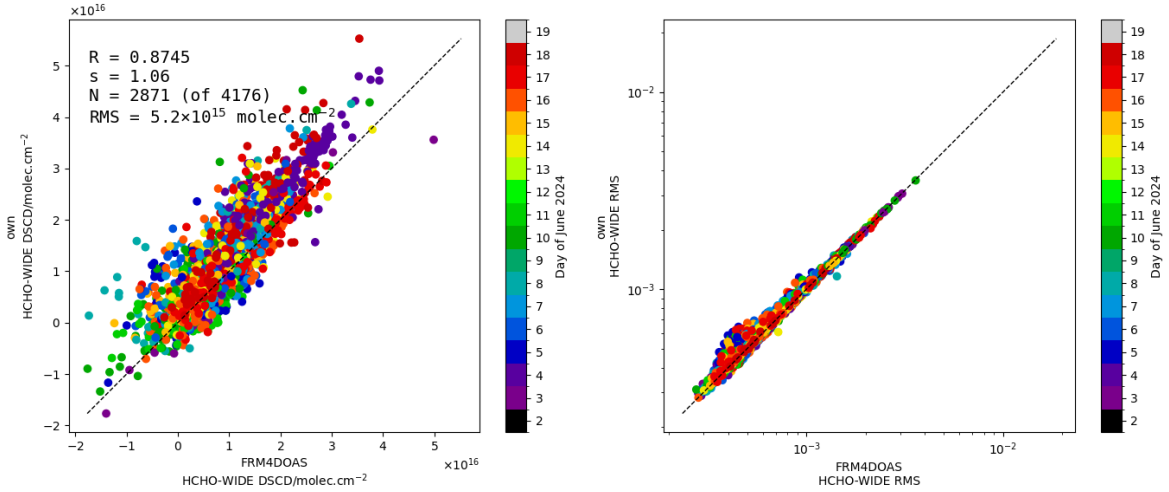

Figure 189: As Fig. 77 but for HCHO-WIDE and instrument #21

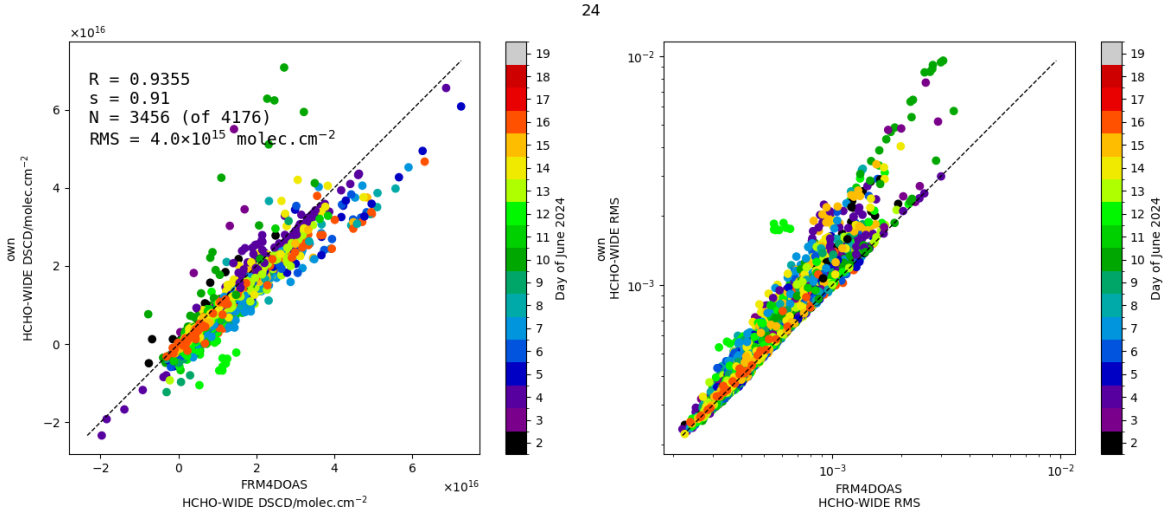

Figure 190: As Fig. 77 but for HCHO-WIDE and instrument #24

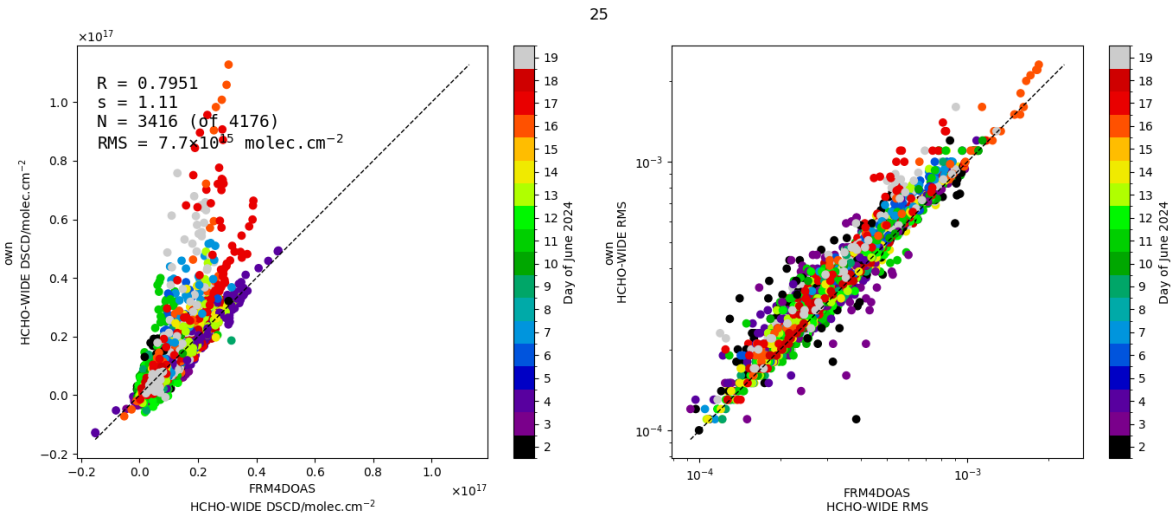

Figure 191: As Fig. 77 but for HCHO-WIDE and instrument #25

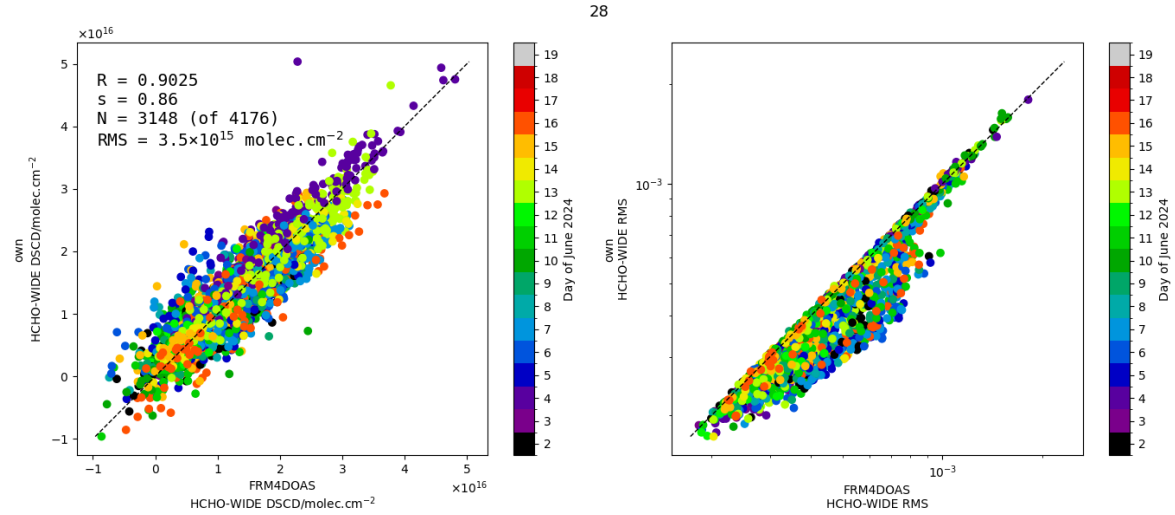

Figure 192: As Fig. 77 but for HCHO-WIDE and instrument #28

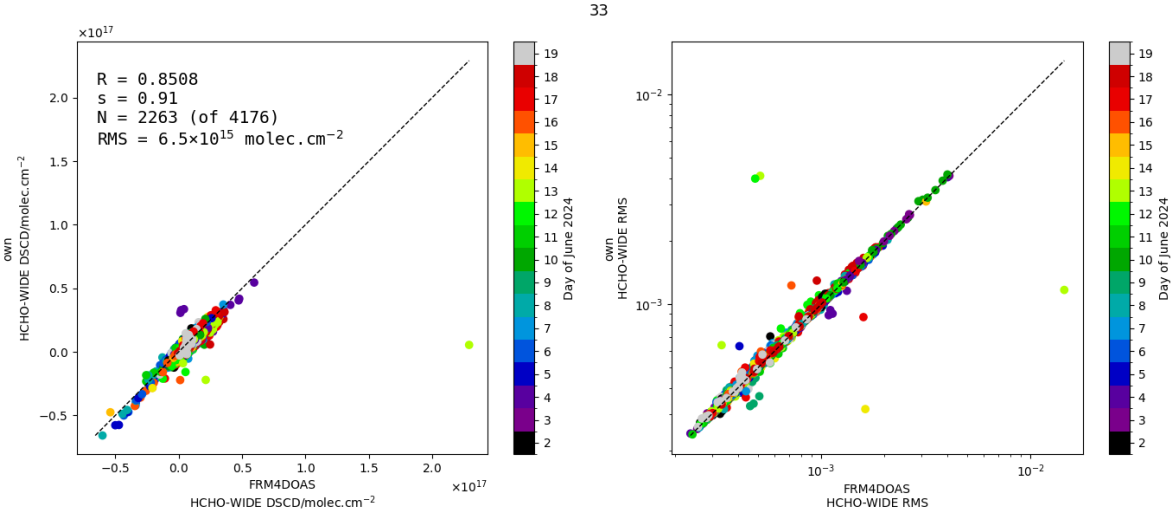

Figure 193: As Fig. 77 but for HCHO-WIDE and instrument #33

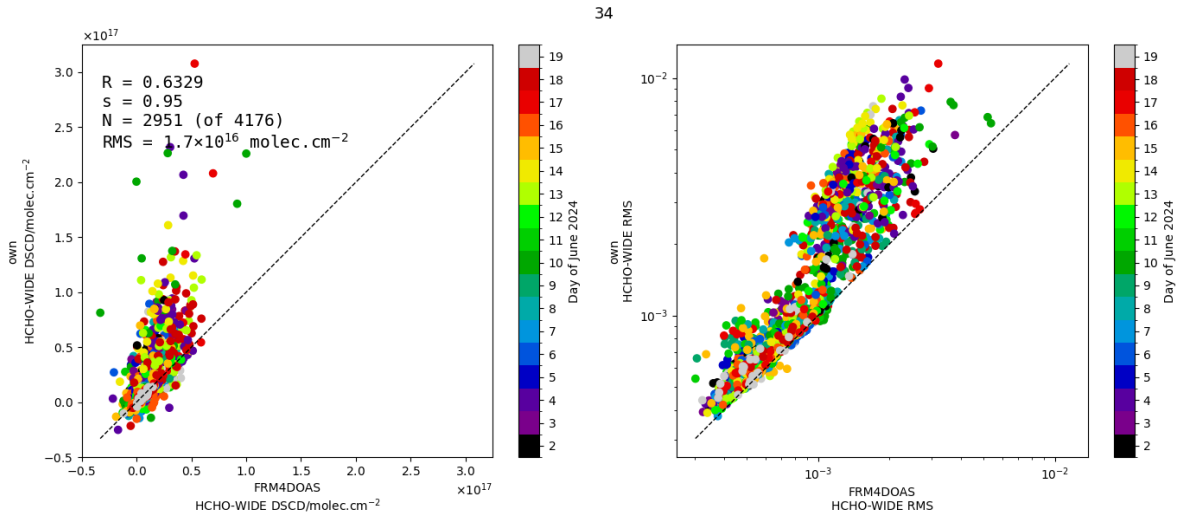

Figure 194: As Fig. 77 but for HCHO-WIDE and instrument #34

D summary of instrument information

All but two PIs (#22 and #40) filled out a questionnaire provided by BIRA-IASB regarding spectral corrections (such as dark, offset, stray-light), details about the wavelength calibration, details about applied filtering etc. This section is a compilation of the answers received.

Question	IUP-UB #25 #27
name	Andreas Richter
email	richter@iup..physik.uni-bremen.de
CINDI3 instrument numbers	25, 27
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual measurements, stop integration if individual measurement is in saturation
exposure time (fixed / calculated (how?))	calculated from previous measurement
integration time (fixed / calculated (how?))	integration of measurements until next measurement would have exceeded available time until next measurement
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	no
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start around midnight
Number of zenith measurements during noon window per day (min, max, median)	TBD
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	polynomial fitted through spectral lamp (HgCd) peaks in full detector coverage
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, shift and stretch of background spectrum on Fraunhofer atlas in the current fitting window, shift and stretch of current measurement on background measurement in DOAS fit
slit function shape (measured / gaussian / other)	measured
slit function width (fixed / fitted / fitted wl dependent)	fixed
Atmospheric absorption included (no / yes (which components, amount))	no
correction for Ring effect applied (no / yes (how?))	no
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	time interpolated
FIXREF (# measurements)	TBD
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum w.r.t reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	only on ln(spectrum)
additional cross sections used (yes / no)	no
offset correction implemented how?	1 / y[lam], normalised
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	SAOZ #8
name	Andrea Pazmino
email	andrea.pazmino@latmos.ipsl.fr
CINDI3 instrument numbers	8
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Saturation level =16000 counts calibration Level 13000 counts (Automatic to calculate exposure time and number of scans) no individual check
exposure time (fixed / calculated (how?))	Acquisition of two spectra at low exposure time and linear extrapolation of the exposure time to obtain 13000 counts at 450 nm. Selection of the final exposure time in a series of Renard of reason 7.
integration time (fixed / calculated (how?))	fixed (generally 60s) or it can reach 5 times de fixed value for important exposure time
measurement starts on the minute (exactly / approximately (max deviation))	It's not time that controls the start of measurements, but the zenith angle. Between 96 and 80 degrees from SZA, the measurement is in ASAP mode. Between 80 and 10 degrees, one measurement every 120 seconds (for the CINDI 3 campaign).
daily dark current measurements (no / yes (start at time / sza / other?))	The dark current spectrum is acquired after the spectrum acquisition with the same integration time (same exposure time and number of scans) for the following conditions: * Change in exposure time compared to the previous measurement * After the acquisition of 10 spectra with the same exposure time * After a 1-degree change in the spectrometer temperature.
daily electronic offset measurements (no / yes (start at time / sza / other?))	No
Number of zenith measurements during noon window per day (min, max, median)	The number of measurements depends on weather conditions. 5 acquisitions maximum between 11:35 a.m. and 11:45 a.m.
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	No
dark current / offset correction applied (no / yes (fixed, daily))	Yes, see item 5
straylight correction applied (no / yes (how?))	No
other corrections applied (no / yes (which, how?))	No
radiometric calibration applied (no / yes (how?))	No
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Not reported in L1 files any calibration with lampes. The wavelength law and the resolution of the reference spectrum are determined using a very high resolution solar spectrum.
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Optimization using Fraunhofer lines during processing (not reported in L1 files)
slit function shape (measured / gaussian / other)	Gaussian
slit function width (fixed / fitted / fitted wl dependent)	Fitted wavelenght dependent. Search for the best full-width-at-half-maximum Gaussians (minimizing the relative error between Kurucz and Mini-SAOZ spectra) by convolving the Kurucz spectrum for a 10 nm wide band with a 2 nm shift between 325 and 630 nm (sliding measurement), corresponding to 145 values.
Atmospheric absorption included (no / yes (which components, amount))	Reference spectrum is corrected only for the linearity if the detector and the hot pixels
correction for Ring effect applied (no / yes (how?))	Reference spectrum is corrected only for the linearity if the detector and the hot pixels
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	measured spectrum between 11:35 a.m. and 11:45 a.m.
SEQREF (before / after / time interpolated)	NaN
FIXREF (# measurements)	only one measured during the campaign (2024-06-06)
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	Yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	The spectrum is shift / stretch on the reference spectrum using a polynomial function
additional cross sections used (yes / no)	No
offset correction implemented how?	No
filtering of spectral outliers or spikes in residuals (no / yes (how?))	No

Question	bira #20
name	Michel Van Roozendael
email	michel.vanroozendael@aeronomie.be
CINDI3 instrument numbers	20
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated (exposure time extrapolated to fill 60% of the max. count rate)
integration time (fixed / calculated (how?))	calculated as the product of the exposure time and the number of acquisitions within one minute
measurement starts on the minute (exactly / approximately (max deviation))	measurement typically starts 7-10 sec after the minute (max deviation 50 sec)
daily dark current measurements (no / yes (start at time / sza / other?))	yes, starts after 95° SZA
daily electronic offset measurements (no / yes (start at time / sza / other?))	mean electronic offset measured after each acquisition
Number of zenith measurements during noon window per day (min, max, median)	10, 10, 10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	yes (based on offset measured after each acquisition, and daily dark current)
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	initial calibration based on HgCd lamp measurement
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, based on QDOAS calibration tool using SAO2010 solar atlas
slit function shape (measured / gaussian / other)	measured (using HgCd lamp)
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependent, starting from measured line shape
Atmospheric absorption included (no / yes (which components, amount))	yes (ozone in the UV range)
correction for Ring effect applied (no / yes (how?))	yes, using a standard Ring effect spectrum (calculated for a standard atmosphere at 50° SZA)
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum only
additional cross sections used (yes / no)	yes (resolution change)
offset correction implemented how?	yes, non linear fit
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes, based oo QDOAS spike removal function (tolerance factor=5)

Question	INTA #30
name	Mónica Navarro-Comas
email	navarrocm@inta.es
CINDI3 instrument numbers	30
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Check individual exposures
exposure time (fixed / calculated (how?))	calculated to reach maximum signal avoiding saturation
integration time (fixed / calculated (how?))	number of scans x exposure time
measurement starts on the minute (exactly / approximately (max deviation))	Approximately. Min = 8s and max = 25s depending on the VAA and IEA for offaxis observations. Min =2s max=4s for zenith twilight observations.
daily dark current measurements (no / yes (start at time / sza / other?))	yes, a few days. Starting at 21:00
daily electronic offset measurements (no / yes (start at time / sza / other?))	no
Number of zenith measurements during noon window per day (min, max, median)	8,10,9
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	yes fixed
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	yes sprectal lamp
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	no
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	wavelength dependent
Atmospheric absorption included (no / yes (which components, amount))	according to settings (NO2, O3)
correction for Ring effect applied (no / yes (how?))	no
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	yes filtered if saturated
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	8
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	yes, according to settings
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	Mini SAOZ #9 & #10
name	Andrea Pazmino
email	andrea.pazmino@latmos.ipsl.fr
CINDI3 instrument numbers	9 and 10
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Saturation level =16000 counts calibration Level 12000 counts (Automatic to calculate exposure time and number of scans) no individual check
exposure time (fixed / calculated (how?))	Acquisition of two spectra at low exposure time and linear extrapolation of the exposure time to obtain 12000 counts at 450 nm. Selection of the final exposure time in a series of Renard of reason 17.
integration time (fixed / calculated (how?))	fixed (generally 60s) or it can reach 5 times de fixed value for important exposure time
measurement starts on the minute (exactly / approximately (max deviation))	It's not time that controls the start of measurements, but the zenith angle. Between 96 and 80 degrees from SZA, the measurement is in ASAP mode. Between 80 and 10 degrees, one measurement every 120 seconds (for the CINDI 3 campaign).
daily dark current measurements (no / yes (start at time / sza / other?))	The dark current spectrum is acquired after the spectrum acquisition with the same integration time (same exposure time and number of scans) for the following conditions: * Change in exposure time compared to the previous measurement * After the acquisition of 10 spectra with the same exposure time * After a 1-degree change in the spectrometer temperature.
daily electronic offset measurements (no / yes (start at time / sza / other?))	No
Number of zenith measurements during noon window per day (min, max, median)	The number of measurements depends on weather conditions. 5 acquisitions maximum between 11:35 a.m. and 11:45 a.m.
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	Yes, polynomial function.
dark current / offset correction applied (no / yes (fixed, daily))	Yes, see item 5
straylight correction applied (no / yes (how?))	No
other corrections applied (no / yes (which, how?))	Hot pixels correction. The hot pixels detected during calibration.The values of the hot pixel correspond to the mean value using neighbors pixels.
radiometric calibration applied (no / yes (how?))	No
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Not reported in L1 files any calibration with lampes. The wavelength law and the resolution of the reference spectrum are determined using a very high resolution solar spectrum.
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Optimization using Fraunhofer lines during processing (not reported in L1 files)
slit function shape (measured / gaussian / other)	Gaussian
slit function width (fixed / fitted / fitted wl dependent)	Fitted wavelenght dependent. Search for the best full-width-at-half-maximum Gaussians (minimizing the relative error between Kurucz and Mini-SAOZ spectra) by convolving the Kurucz spectrum for a 10 nm wide band with a 2 nm shift between 340 and 800 nm (sliding measurement), corresponding to 225 values.
Atmospheric absorption included (no / yes (which components, amount))	Reference spectrum is corrected only for the linearity if the dectecteur and the hot pixels
correction for Ring effect applied (no / yes (how?))	Reference spectrum is corrected only for the linearity if the dectecteur and the hot pixels
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	measured spectrum between 11:35 a.m. and 11:45 a.m.
SEQREF (before / after / time interpolated)	NaN
FIXREF (# measurements)	only one measured during the campaign (2024-06-06)
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	Yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	The spectrum is shift / stretch on the reference spectrum using a polynomial function
additional cross sections used (yes / no)	No
offset correction implemented how?	No
filtering of spectral outliers or spikes in residuals (no / yes (how?))	No

Question	AIOFM #15
name	Ang Li, Zhaokun Hu, Chaonan Lv
email	angli@aiofm.ac.cn, zkhu@aiofom.ac.cn, cnlv@aiofm.ac.cn
CINDI3 instrument numbers	AIOFM#15
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated to reach max signal avoiding saturation
integration time (fixed / calculated (how?))	exposure time * scan times
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes, start at 2.30
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start at 2.30
Number of zenith measurements during noon window per day (min, max, median)	10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	yes, daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Calibration based on Hg lamp instrument
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	YES baesd on QDOAS calibration tool using SAO2010 solar atlas
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	fixed
Atmospheric absorption included (no / yes (which components, amount))	yes. O3_223K
correction for Ring effect applied (no / yes (how?))	yes.The correction is done by using the specified ring effect cross-section file.Ring_QDOAScalc_HighResSAO2010_Norm.xs
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	yes.For spectrum shfit="true" stfit="1st" For ring shfit="true" stfit="1st"
additional cross sections used (yes / no)	no
offset correction implemented how?	yes (via DOAS fit offset term)
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	BIRA #21
name	Alexis Merlaud
email	alexis.merlaud@aeronomie.be
CINDI3 instrument numbers	21
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Check individual exposures
exposure time (fixed / calculated (how?))	Calculated from intensity of previous scan to yield constant saturation level
integration time (fixed / calculated (how?))	Fixed according to CINDI3 schedule
measurement starts on the minute (exactly / approximately (max deviation))	Exactly
daily dark current measurements (no / yes (start at time / sza / other?))	Continuously at night for SZA > 105°
daily electronic offset measurements (no / yes (start at time / sza / other?))	Continuously at night for SZA > 105°
Number of zenith measurements during noon window per day (min, max, median)	10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, only for the visible channel
dark current / offset correction applied (no / yes (fixed, daily))	yes, daily
straylight correction applied (no / yes (how?))	yes for the visible channel, zong method with the luftblick tunable laser
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Optical density fitting, 8 windows in 330-455 nm (UV), 12 in 420-540 nm (VIS)
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	No
slit function shape (measured / gaussian / other)	UV : Gaussian, Vis: Assymetric Gaussian
slit function width (fixed / fitted / fitted wl dependent)	Wl dependent, fitted
Atmospheric absorption included (no / yes (which components, amount))	ring, o3 (UV), ring (Vis)
correction for Ring effect applied (no / yes (how?))	yes, with a fixed Raman scaled cross section, without fitting
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	9
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum only
additional cross sections used (yes / no)	no
offset correction implemented how?	yes, non linear fit, but no offset applied for o3 visible
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	FING #32
name	Lampel, Johannes; Erna Frins; Roberto Barragán
email	johannes.lampel@airyx.de; efrins@fing.edu.uy; roberto.barragan.c@academicos.udg.mx
CINDI3 instrument numbers	32
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Rejection by MSDOAS above 95% sat
exposure time (fixed / calculated (how?))	calculated from previous spectra by MSDOAS
integration time (fixed / calculated (how?))	calculated from previous spectra by MSDOAS
measurement starts on the minute (exactly / approximately (max deviation))	exactly, as in MSDOAS script defined by Udo
daily dark current measurements (no / yes (start at time / sza / other?))	yes, but as stability of instrument was sufficient, one fixed was used
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, but as stability of instrument was sufficient, one fixed was used
Number of zenith measurements during noon window per day (min, max, median)	NaN
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, according to polynomial measured in lab
dark current / offset correction applied (no / yes (fixed, daily))	yes,
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	NaN
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	based on hg spectrum recorded on-site at the beginning of the campaign
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	no, but recorded also nightly
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependent
Atmospheric absorption included (no / yes (which components, amount))	no
correction for Ring effect applied (no / yes (how?))	yes,
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	mean of 10 noon spectra
SEQREF (before / after / time interpolated)	Averages references after and before
FIXREF (# measurements)	mean of 9 noon spectra; 2024-06-06
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	gaussian, from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum & ref (in the UV only ref)
additional cross sections used (yes / no)	no
offset correction implemented how?	we follow CINDI-3 protocol
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	AIOFM #17
name	Yuhan Luo
email	yhluo@aiofm.ac.cn
CINDI3 instrument numbers	17
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated to reach maximum signal avoiding saturation
integration time (fixed / calculated (how?))	calculated as the product of the exposure time and the number of acquisitions within one minute
measurement starts on the minute (exactly / approximately (max deviation))	Exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes
Number of zenith measurements during noon window per day (min, max, median)	10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	calibration based on Hg lamp measurement
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, based on QDOAS calibration tool using SAO2010 solar atlas
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	fixed
Atmospheric absorption included (no / yes (which components, amount))	yes (O3 in the UV range)
correction for Ring effect applied (no / yes (how?))	Yes, using the CINDI-3 reference spectra (Ring_QDOAScalc_HighResSAO2010_Norm.xs)
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	yes(constant and order one)
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	USTC #39
name	Cheng Liu
email	chliu81@ustc.edu.cn, xgji2017@mail.ustc.edu.cn, xingcz@aiofm.ac.cn, sytang2023@mail.ustc.edu.cn
CINDI3 instrument numbers	#39
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	fixed
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes
Number of zenith measurements during noon window per day (min, max, median)	min
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes
dark current / offset correction applied (no / yes (fixed, daily))	yes,fixed
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	HG-calibration,
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes,sun light calibration in QDOAS
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	fixed
Atmospheric absorption included (no / yes (which components, amount))	yes, ozone, ring
correction for Ring effect applied (no / yes (how?))	yes,included Ring XS
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	NaN
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	no
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum and Cross-sections
additional cross sections used (yes / no)	no
offset correction implemented how?	constant and order one
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes, included gaps

Question	PANDORA #34, #35, #36
name	Alexander Cede
email	alexander.cede@luftblick.at
CINDI3 instrument numbers	34,35,36
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated: exp. time = exp. time(intensity check) / (max. counts(intensity check) / (0.8 * saturation limit))
integration time (fixed / calculated (how?))	fixed20s for Twilight Zenith-Sky reference (including dark)52s for Twilight Zenith-Sky scans (including dark)24s/position for elevation scan (including dark)20s/position Almucantar scan (including dark)10s for Direct Sun (including dark)4s/position for horizon scan (no dark)Note: we switch between two functional filters within the same routine, hence the real integration times are lower than the integration times specified by the CINDI-3 protocol
measurement starts on the minute (exactly / approximately (max deviation))	approximately: after the previous measurement finishes (max deviation < 5 minutes)
daily dark current measurements (no / yes (start at time / sza / other?))	other: after each routine measurement (for all routines except horizon scan)
daily electronic offset measurements (no / yes (start at time / sza / other?))	other: included in dark current measurements
Number of zenith measurements during noon window per day (min, max, median)	min: 4, max: 7, median: 5
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, polynomial fit characterized in lab calibration
dark current / offset correction applied (no / yes (fixed, daily))	yes, from measurement, as it is included in the dark counts
straylight correction applied (no / yes (how?))	no, only "simple stray light correction" (constant correction by mean intensity between 288-290nm, which is assumed to be attributed to instrument stray light only)
other corrections applied (no / yes (which, how?))	yes, flat field, dispersion & resolution corrected by field data
radiometric calibration applied (no / yes (how?))	yes, from lab calibration and corrected by field data
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	lab calibration with gas emission lamps
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, using noon spectrafrom 20240605 (#34)from 20240619 (#35 & #36)
slit function shape (measured / gaussian / other)	Super-Gauss
slit function width (fixed / fitted / fitted wl dependent)	fitted wavelength dependent
Atmospheric absorption included (no / yes (which components, amount))	yes: O3 fitted to extraterrestrial reference
correction for Ring effect applied (no / yes (how?))	no
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	filtered: no saturation, number of cycles > 1
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	9
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	no, but the other way around: wavelength alignment of reference spectrum wrt spectrum using interpolation
slit function shape&width (fixed / from wavelength calibration)	width fitted, width and shape determined from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	yes, on spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	offset in intensity space with polynomial scaled to mean intensity
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	CNR-ISAC #24
name	Andrè Achilli
email	a.achilli@isac.cnr.it
CINDI3 instrument numbers	24
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	set max detectors saturation at 40%, then the exposition times are calculated based on this information
exposure time (fixed / calculated (how?))	calculated by the acquisition program MSDOAS
integration time (fixed / calculated (how?))	fixed as per CINDI3 protocol
measurement starts on the minute (exactly / approximately (max deviation))	NaN
daily dark current measurements (no / yes (start at time / sza / other?))	yes, start at time 21:00 UTC after offset measurements
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start at time 21:00 UTC
Number of zenith measurements during noon window per day (min, max, median)	10, 10, 10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, the non-linearity coefficients are given by Airyx, then for each wavelength they are used to calculate a polynomial of the spectrum and the orignal one is divided by it.
dark current / offset correction applied (no / yes (fixed, daily))	yes, daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	HG lamp
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	QDOAS contiguous division, for each molecule the number of windows to bew used has been selected manually doing multiple tests.
slit function shape (measured / gaussian / other)	gaussian
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependant
Atmospheric absorption included (no / yes (which components, amount))	no
correction for Ring effect applied (no / yes (how?))	no
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	8
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	QDOAS linear offset fit
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	SUWON #7
name	Hyeong-Ahn Kwon
email	hakwon@suwon.ac.kr
CINDI3 instrument numbers	7
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated. Total integration time per elevation angle is fixed, exposure time is adjusted according to saturation
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
Number of zenith measurements during noon window per day (min, max, median)	8,10,10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes (from the DOASIS software)
dark current / offset correction applied (no / yes (fixed, daily))	yes (fixed)
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Using HG lamp
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Yes (optical density fitting using solar reference spectrum with several calibration windows (the number of sub-windows depends on fitting windows for target species); shift: non-linear, stretch: 1st order)
slit function shape (measured / gaussian / other)	Gaussian
slit function width (fixed / fitted / fitted wl dependent)	Fitted
Atmospheric absorption included (no / yes (which components, amount))	no
correction for Ring effect applied (no / yes (how?))	Yes (using theprovided ring spectrum in QDOAS)
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	8 (11:39 UTC and 11:44 UTC measurements are excluded on June 6)
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	From wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum and ref spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	follow suggestions from the protocol
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes (spike tolerance factor 4) and a sza dependent rms criterion

Question	IUPHD #28, #29
name	Udo Frieß
email	udo.friess@iup.uni-heidelberg.de
CINDI3 instrument numbers	28, 29
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Check individual exposures
exposure time (fixed / calculated (how?))	Calculated from intensity of previous scan to yield constant saturation level
integration time (fixed / calculated (how?))	Fixed according to CINDI-3 schedule
measurement starts on the minute (exactly / approximately (max deviation))	Exactly
daily dark current measurements (no / yes (start at time / sza / other?))	Continuously at night for SZA > 105°
daily electronic offset measurements (no / yes (start at time / sza / other?))	Continuously at night for SZA > 105°
Number of zenith measurements during noon window per day (min, max, median)	10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	28: yes using the polynomial coefficiens supplied by Airyx; 29: no
dark current / offset correction applied (no / yes (fixed, daily))	daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Supergaussian fit to Hg lines at 365.02, 404.66, 407.78, 435.84, 491.6, 507.3, 546.07, 576.96, 579.07, 593.46, 604.9, 625.14, 626.34 nm, polynomial fit using peak locations
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Non-linear fit to high-resolution Kurucz spectrum
slit function shape (measured / gaussian / other)	Measured Hg lines
slit function width (fixed / fitted / fitted wl dependent)	Wl-dependent using several Hg lines
Atmospheric absorption included (no / yes (which components, amount))	O3, NO2, O4, Raman, fitted during Kurucz calibration
correction for Ring effect applied (no / yes (how?))	Yes, using high-resolution Raman spectrum as part of the CINDI-3 reference spectra (Ring_QDOAScalc_HighResSAO2010_Norm.xs)
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	filtered for oversatured spectra
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	on measured spectrum and cross sections
additional cross sections used (yes / no)	no
offset correction implemented how?	subtracted according to number of scans
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	BIRA #23
name	Cedric Busschots
email	cedric.busschots@aeronomie.be
CINDI3 instrument numbers	23
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Check individual exposures
exposure time (fixed / calculated (how?))	1 second, fixed
integration time (fixed / calculated (how?))	69 seconds (1 second per wavelength)
measurement starts on the minute (exactly / approximately (max deviation))	approximately (5 minutes)
daily dark current measurements (no / yes (start at time / sza / other?))	Yes, at each acquisition
daily electronic offset measurements (no / yes (start at time / sza / other?))	No
Number of zenith measurements during noon window per day (min, max, median)	2, but never used due to clouds
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	No
dark current / offset correction applied (no / yes (fixed, daily))	Yes, each acquisition
straylight correction applied (no / yes (how?))	Yes, measured at each acquisition and subtracted
other corrections applied (no / yes (which, how?))	No
radiometric calibration applied (no / yes (how?))	No
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Non-linear minimization of the high-pass-filtered absolute difference between spectrum and zenith, on window 434-455 nm (scipy minimize with Powell's method)
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Yes, linearized using 2 pseudo-absorbers for shift & stretch
slit function shape (measured / gaussian / other)	Gaussian
slit function width (fixed / fitted / fitted wl dependent)	Fixed, wl-dependent (1.2-1.5 nm)
Atmospheric absorption included (no / yes (which components, amount))	No
correction for Ring effect applied (no / yes (how?))	Only for calibration of the zenith spectrum vs Fraunhofer ref spectrum; fixed at 0.05
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	/
SEQREF (before / after / time interpolated)	/
FIXREF (# measurements)	One measurement using Spectralon flat field diffuser + additive term for conversion to official FIXREF (computed using instrument #20); we did this because we had no perfectly cloud-free zenith.
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	pseudo absorbers for shift & stretch and offset correction
offset correction implemented how?	2 pseudo absorbers, allowing linear wavelength dependence
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	KNMI #03, #04
name	Ankie Piters
email	piters@knmi.nl
CINDI3 instrument numbers	#03, #04
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	#03: max rel intensity = 50%; #04: check every 10sec
exposure time (fixed / calculated (how?))	#03: calculated for 0.5*saturation; #04: calculated for 0.9*saturation
integration time (fixed / calculated (how?))	#03: fixed 52 sec; #04: check every 10 sec until IT>51 sec
measurement starts on the minute (exactly / approximately (max deviation))	approximately, #03: max 20 sec, #04: max 52sec
daily dark current measurements (no / yes (start at time / sza / other?))	yes, SZA>100
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, SZA>100
Number of zenith measurements during noon window per day (min, max, median)	#03: Min:9, max:11, med:9; #04: min:8, max:10, med:9.5
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	yes, daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	fit to solar spectrum, for window slightly larger than doas fitting window, to include strong solar lines
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	no
slit function shape (measured / gaussian / other)	gauss
slit function width (fixed / fitted / fitted wl dependent)	fitted
Atmospheric absorption included (no / yes (which components, amount))	yes, O3 1e20, no2 1e17
correction for Ring effect applied (no / yes (how?))	no
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	8
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	cross sections
additional cross sections used (yes / no)	no
offset correction implemented how?	Polynomial in wavelength divided by normalised intensity
filtering of spectral outliers or spikes in residuals (no / yes (how?))	Yes, saturated and zero pixels are removed

Question	LMU #14
name	Zeqing Chen
email	Zeqing.Chen@physik.uni-muenchen.de
CINDI3 instrument numbers	14
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated. Total integration time per elevation angle is fixed, exposure time is adjusted according to saturation
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
Number of zenith measurements during noon window per day (min, max, median)	(7, 9, 8)
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, the spectral measurements are corrected using a fifth-degree polynomial to accurately model and compensate for the detector's pixel-wise variation.
dark current / offset correction applied (no / yes (fixed, daily))	yes, fixed
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Calibration using Hg lamp emission lines
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, based on the functions inside QDOAS software, the original wavelength-pixels relation from step 13 is corrected using a procedure based on the alignment of the Fraunhofer structures of the reference spectrum with those of an accurately calibrated high-resolution solar reference atlas, degraded at the resolution of the instrument, i.e. convolved with the instrumental slit function.
slit function shape (measured / gaussian / other)	measured
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependent
Atmospheric absorption included (no / yes (which components, amount))	yes, O3(1.4E19)
correction for Ring effect applied (no / yes (how?))	yes, convolve the ring effect pseudo cross-section with the slit function, and set a constant amount (0.07) for it
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	8
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	yes, non-linear fit
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes, based on QDOAS spike removal function with tolerance factor of 5

Question	RAL #33
name	Ka Lok Chan
email	ka.chan@stfc.ac.uk
CINDI3 instrument numbers	NaN
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	calculated. Total integration time per elevation angle is fixed, exposure time is adjusted according to saturation
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, start at sza 100 degree
Number of zenith measurements during noon window per day (min, max, median)	(8, 10, 10)
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes, the spectral measurements are corrected using a fifth-degree polynomial to accurately model and compensate for the detector's pixel-wise variation.
dark current / offset correction applied (no / yes (fixed, daily))	yes, fixed
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Calibration using Hg lamp emission lines
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, based on the functions inside QDOAS software, the original wavelength-pixels relation from step 13 is corrected using a procedure based on the alignment of the Fraunhofer structures of the reference spectrum with those of an accurately calibrated high-resolution solar reference atlas, degraded at the resolution of the instrument, i.e. convolved with the instrumental slit function.
slit function shape (measured / gaussian / other)	measured
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependent
Atmospheric absorption included (no / yes (which components, amount))	yes, O3(1.4E19)
correction for Ring effect applied (no / yes (how?))	yes, convolve the ring effect pseudo cross-section with the slit function, and set a constant amount (0.07) for it
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	yes, non-linear fit
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes, based on QDOAS spike removal function with tolerance factor of 5

Question	AUTH #18, #19
name	Alkis Bais; Dimitris Karagkiozidis; Dimitris Nikolis
email	abais@auth.gr; dkaragki@auth.gr; dnikolis@auth.gr
CINDI3 instrument numbers	#18 and #19
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Check individual exposures. Saturation of the detector is prevented by appropriately adjusting its exposure time
exposure time (fixed / calculated (how?))	Calculated. Before each actual measurement, two test spectra are recorded with very short exposure times. The optimal exposure time is calculated by interpolation so that the maximum intensity of the measured spectrum is of about 50,000 counts. During twilight Zenith-Sky and Direct-Sun measurements, the optimal exposure time is recalculated every few seconds to ensure saturation prevention due to potential atmospheric/weather changes (e.g., cloud passing in the FOV). For all other observations, the optimal exposure time remains constant.
integration time (fixed / calculated (how?))	Calculated based on the CINDI-3 Data Acquisition Protocol. See also question #4
measurement starts on the minute (exactly / approximately (max deviation))	Approximately. A few seconds are needed before starting a new measurement to move the optics to the appropriate viewing position, set the filterwheel and determine the exposure time. The time required for these actions depends mainly on the previous/next position of the optics. The higher the "distance" the instrument is moved in azimuth and elevation, the higher the deviation from starting exactly on the minute. It is noted that instrument #19 is significantly faster than #18. The maximum deviation during CINDI-3 was 20 sec.If the suggested integration time for a measurement is e.g., 60 sec and it takes 5 sec for the instrument to get in position, the rest of the time (55 sec) is set as the integration time of the measurement, so that it stays in line with the Data Acquisition Protocol.
daily dark current measurements (no / yes (start at time / sza / other?))	Yes on a daily basis (starting at 105 deg. SZA with an opaque position on the filter-wheel). Dark current measurements are performed for each exposure time used during the day for a total acquisition time of 30 min.
daily electronic offset measurements (no / yes (start at time / sza / other?))	No
Number of zenith measurements during noon window per day (min, max, median)	All non-saturated spectra (#18, min: 5, max: 12, median: 10; #19 UV channel, min: 4, max: 6, median: 5; #19 VIS channel, min: 3, max: 6, median: 5)
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	No
dark current / offset correction applied (no / yes (fixed, daily))	Yes (daily)
straylight correction applied (no / yes (how?))	No
other corrections applied (no / yes (which, how?))	No
radiometric calibration applied (no / yes (how?))	No
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Using Hg lamp measurements in the lab
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Yes, using the standard "calibration" of QDOAS algorithm
slit function shape (measured / gaussian / other)	Measured using Hg lamp spectra. Approximately gaussian of 0.85 nm FWHM for #19 and 0.55 nm FWHM for #18
slit function width (fixed / fitted / fitted wl dependent)	Fitted and wavelength dependent (using QDOAS)
Atmospheric absorption included (no / yes (which components, amount))	Yes, O3 (determined by QDOAS)
correction for Ring effect applied (no / yes (how?))	Yes
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	All non-saturated spectra
SEQREF (before / after / time interpolated)	After
FIXREF (# measurements)	All non-saturated spectra (#18: 9, #19 UV channel: 5, #19 VIS channel: 3)
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	No
slit function shape&width (fixed / from wavelength calibration)	From wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	Allowed on spectrum (shift fit: non-linear; stretch fit: 1st order)
additional cross sections used (yes / no)	No
offset correction implemented how?	Standard QDOAS algorithm
filtering of spectral outliers or spikes in residuals (no / yes (how?))	No

Question	UOM & ABOM #1, #2
name	Robert Ryan
email	rryan1@unimelb.edu.au
CINDI3 instrument numbers	1 + 2
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	Max. relative intensity
exposure time (fixed / calculated (how?))	Calculated within MS-DOAS software using fixed integration time (60 s) and saturation
integration time (fixed / calculated (how?))	Fixed (60 s)
measurement starts on the minute (exactly / approximately (max deviation))	Exactly
daily dark current measurements (no / yes (start at time / sza / other?))	Yes, nightly, after SZA 110
daily electronic offset measurements (no / yes (start at time / sza / other?))	Yes, nightly, after SZA 110
Number of zenith measurements during noon window per day (min, max, median)	Fixed at 12
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	No
dark current / offset correction applied (no / yes (fixed, daily))	Yes, daily
straylight correction applied (no / yes (how?))	No
other corrections applied (no / yes (which, how?))	No
radiometric calibration applied (no / yes (how?))	No
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Initial wavelength cal calculated as a linear function from Hg lamp peaks using DOASIS
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	Yes, using QDOAS. 20 contiguous windows fitted within QDOAS calib tab including fitting Ring and O3_223K cross sections. 5th order polynomials fitted for both wavelength shift and slit function parameter (gaussian).
slit function shape (measured / gaussian / other)	Initially measured from Hg lamp, then fitted with gaussian function in QDOAS
slit function width (fixed / fitted / fitted wl dependent)	From gaussian fit in QDOAS calib procedure
Atmospheric absorption included (no / yes (which components, amount))	No
correction for Ring effect applied (no / yes (how?))	Yes, Ring spectrum calculated in QDOAS tool with intitial wavelength cal function and initial measured slit function
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	average of 12 noon spectra on 6 June
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calib
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	ref spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	linear offset order 1
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

Question	UOT #6
name	Kimberly Strong; Kevin Joshy; Ramina Alwarda; Darby Bates
email	strong@atmosp.physics.utoronto.ca; kevin.joshy@mail.utoronto.ca; ramina.alwarda@mail.utoronto.ca; dbates@physics.utoronto.ca
CINDI3 instrument numbers	6
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	max rel intensity
exposure time (fixed / calculated (how?))	calculated (calculated in MS-DOAS software)
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	approximately (max deviation: 1 sec)
daily dark current measurements (no / yes (start at time / sza / other?))	yes (start at 19:10:00 UTC)
daily electronic offset measurements (no / yes (start at time / sza / other?))	no
Number of zenith measurements during noon window per day (min, max, median)	min: 9; max: 11; median: 10
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes (applied in DOASIS software)
dark current / offset correction applied (no / yes (fixed, daily))	yes (fixed)
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Kurucz calibration fits in QDOAS
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes (calibration using reference spectrum only in QDOAS)
slit function shape (measured / gaussian / other)	measured
slit function width (fixed / fitted / fitted wl dependent)	fixed
Atmospheric absorption included (no / yes (which components, amount))	yes (O3 absorption, Vis SCD initial: 9e+19, UV SCD initial: 1e+19)
correction for Ring effect applied (no / yes (how?))	yes (ring effect cross section is included in initial calibration procedure)
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	9 out of 10
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	fixed
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	shift & stretch allowed on spectruml; shift allowed on ref spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	Implemented in QDOAS software
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes (manual filtering via identification of spikes in residuals, and setting spike tolerance factor in QDOAS)

Question	MPIC #5
name	Steffen Ziegler
email	Steffen.Ziegler@mpic.de
CINDI3 instrument numbers	5, 52, 53, 65, 66
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	max rel intensity, individual spectra can be saturated. This can be usually identified by significantly increased fit residuals
exposure time (fixed / calculated (how?))	calculated from test measurement just before the "real" measurement
integration time (fixed / calculated (how?))	calculated from the remaining time within the 60s time window of the CINDI3-schedule
measurement starts on the minute (exactly / approximately (max deviation))	approximately (typically below 5s, but exceptions might occur when light conditions change dramatically and/or the motor needs to move a lot, e.g. from 90° to -2°
daily dark current measurements (no / yes (start at time / sza / other?))	yes, outside of the CINDI3-daytime schedule. However, as temperature is stabilized we typically use the same darkcurrent for all measurements
daily electronic offset measurements (no / yes (start at time / sza / other?))	yes, outside of the CINDI3-daytime schedule. However, as temperature is stabilized we typically use the same offset for all measurements
Number of zenith measurements during noon window per day (min, max, median)	23, 23, 23
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	no
dark current / offset correction applied (no / yes (fixed, daily))	yes, fixed, offset spectrum: 20240523 1:30:15 1000scans x 3ms; darkcurrent spectrum: 20240523 2:00:41 2scans x 10s
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	yes, QDOAS: asymmetric Gaussian, 8 windows between 316.7 and 460.0nm using a constant reference spectrum for the whole campaign (same as fixref spectrum, see 21: FIXREF)
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	only for SEQREF
slit function shape (measured / gaussian / other)	asymmetric gaussian
slit function width (fixed / fitted / fitted wl dependent)	fitted wl dependent
Atmospheric absorption included (no / yes (which components, amount))	yes, Ozone, SCD is being fitted
correction for Ring effect applied (no / yes (how?))	yes, as prescribed in the campaign document
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	filtered: Only spectra with a saturation (max counts/65536) between 10% and 70% are considered
SEQREF (before / after / time interpolated)	starting from version 9: after
FIXREF (# measurements)	9 measurements between 11:35:02 and 11:43:57 on 06.06.2024
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes, in QDOAS "refonly"
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	yes, spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	fitted offset (constant + 1 st order according to campaign description)
filtering of spectral outliers or spikes in residuals (no / yes (how?))	yes, hot pixel filtering from characteristic spectra at the beginning of the campaign

Question	AIOFM#16
name	2D-MAX-DOAS
email	phxie@aiofm.ac.cn
CINDI3 instrument numbers	#16
operations	NaN
saturation prevention (max rel intensity / check individual exposures / none / other)	check individual exposures
exposure time (fixed / calculated (how?))	fixed
integration time (fixed / calculated (how?))	fixed
measurement starts on the minute (exactly / approximately (max deviation))	exactly
daily dark current measurements (no / yes (start at time / sza / other?))	yes,other
daily electronic offset measurements (no / yes (start at time / sza / other?))	no
Number of zenith measurements during noon window per day (min, max, median)	median
NaN	NaN
Spectral corrections	NaN
Non-linearity correction applied (no / yes (how?))	yes,Polynomial fit
dark current / offset correction applied (no / yes (fixed, daily))	yes, daily
straylight correction applied (no / yes (how?))	no
other corrections applied (no / yes (which, how?))	no
radiometric calibration applied (no / yes (how?))	no
NaN	NaN
wavelength calibration and slit function	NaN
Basic wavelength calibration (method and settings, e.g. window-edges)	Hg Calibration
Additional wavelength calibration during DOAS retrieval (no / yes (how?))	yes, Frounhofer calibration by QDOAS
slit function shape (measured / gaussian / other)	measured
slit function width (fixed / fitted / fitted wl dependent)	fitted
Atmospheric absorption included (no / yes (which components, amount))	yes, Frounhofer calibration by QDOAS
correction for Ring effect applied (no / yes (how?))	yes convolution ring by QDOAS
NaN	NaN
reference spectrum	NaN
DAILYREF (all / filtered (how?))	all
SEQREF (before / after / time interpolated)	after
FIXREF (# measurements)	NaN
NaN	NaN
doas settings	NaN
wavelength alignment of spectrum wrt reference spectrum using interpolation (yes / no)	yes
slit function shape&width (fixed / from wavelength calibration)	from wavelength calibration
shift&stretch allowed on (spectrum? ref spectrum? Ratio? Cross-sections?)	spectrum
additional cross sections used (yes / no)	no
offset correction implemented how?	no
filtering of spectral outliers or spikes in residuals (no / yes (how?))	no

References

Kreher, K., M. Van Roozendael, F. Hendrick, A. Apituley, E. Dimitropoulou, U. Frieß, A. Richter, T. Wagner, J. Lampel, N. Abuhassan, L. Ang, M. Anguas, A. Bais, N. Benavent, T. Bösch, K. Bognar, A. Borovski, I. Bruchkouski, A. Cede, K. L. Chan, S. Donner, T. Drosoglou, C. Fayt, H. Finkenzeller, D. Garcia-Nieto, C. Gielen, L. Gómez-Martín, N. Hao, B. Henzing, J. R. Herman, C. Hermans, S. Hoque, H. Irie, J. Jin, P. Johnston, J. Khayyam Butt, F. Khokhar, T. K. Koenig, J. Kuhn, V. Kumar, C. Liu, J. Ma, A. Merlaud, A. K. Mishra, M. Müller, M. Navarro-Comas, M. Ostendorf, A. Pazmino, E. Peters, G. Pinardi, M. Pinharanda, A. Piters, U. Platt, O. Postylyakov, C. Prados-Roman, O. Puentedura, R. Querel, A. Saiz-Lopez, A. Schönhardt, S. F. Schreier, A. Seyler, V. Sinha, E. Spinei, K. Strong, F. Tack, X. Tian, M. Tiefengraber, J.-L. Tirpitz, J. van Gent, R. Volkamer, M. Vrekoussis, S. Wang, Z. Wang, M. Wenig, F. Wittrock, P. H. Xie, J. Xu, M. Yela, C. Zhang, and X. Zhao (2020). "Intercomparison of NO, O, O and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2". In: *Atmospheric Measurement Techniques* 13, pp. 2169–2208. DOI: 10.5194/amt-13-2169-2020. URL: https://amt.copernicus.org/articles/13/2169/2020/.

Piters, A. J. M., K. F. Boersma, M. Kroon, J. C. Hains, M. Van Roozendael, F. Wittrock, N. Abuhassan, C. Adams, M. Akrami, M. A. F. Allaart, A. Apituley, J. B. Bergwerff, A. J. C. Berkhout, D. Brunner, A. Cede, J. Chong, K. Clémer, C. Fayt, U. Frieß, L. F. L. Gast, M. Gil-Ojeda, F. Goutail, R. Graves, A. Griesfeller, K. Großmann, G. Hemerijckx, F. Hendrick, B. Henzing, J. Herman, C. Hermans, M. Hoexum, G. R. van der Hoff, H. Irie, P. V. Johnston, Y. Kanaya,

Y. J. Kim, H. Klein Baltink, K. Kreher, G. de Leeuw, R. Leigh, A. Merlaud, M. M. Moerman, P. S. Monks, G. H. Mount, M. Navarro-Comas, H. Oetjen, A. Pazmino, M. Perez-Camacho, E. Peters, A. du Piesanie, G. Pinardi, O. Puentadura, A. Richter, H. K. Roscoe, A. Schönhardt, B. Schwarzenbach, R. Shaiganfar, W. Sluis, E. Spinei, A. P. Stolk, K. Strong, D. P. J. Swart, H. Takashima, T. Vlemmix, M. Vrekoussis, T. Wagner, C. Whyte, K. M. Wilson, M. Yela, S. Yilmaz, P. Zieger, and Y. Zhou (2012). "The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results". In: *Atmospheric Measurement Techniques* 5, pp. 457–485.

Roscoe, H. K., M. Van Roozendael, C. Fayt, A. du Piesanie, N. Abuhassan, C. Adams, M. Akrami, A. Cede, J. Chong, K. Clémer, U. Frieß, M. Gil Ojeda, F. Goutail, R. Graves, A. Griesfeller, K. Grossmann, G. Hemerijckx, F. Hendrick, J. Herman, C. Hermans, H. Irie, P. V. Johnston, Y. Kanaya, K. Kreher, R. Leigh, A. Merlaud, G. H. Mount, M. Navarro, H. Oetjen, A. Pazmino, M. Perez-Camacho, E. Peters, G. Pinardi, O. Puentedura, A. Richter, A. Schönhardt, R. Shaiganfar, E. Spinei, K. Strong, H. Takashima, T. Vlemmix, M. Vrekoussis, T. Wagner, F. Wittrock, M. Yela, S. Yilmaz, F. Boersma, J. Hains, M. Kroon, A. Piters, and Y. J. Kim (2010). "Intercomparison of slant column measurements of NO and O by MAXDOAS and zenith-sky UV and visible spectrometers". In: *Atmospheric Measurement Techniques* 3, pp. 1629–1646. DOI: 10.5194/amt-3-1629-2010.

Sen, Pranab Kumar (1968). "Estimates of the Regression Coefficient Based on Kendall's Tau". In: *Journal of the American Statistical Association* 63.324, pp. 1379–1389. DOI: 10.1080/01621459.1968.10480934.

Theil, Henri (1950). "A rank-invariant method of linear and polynomial regression analysis". In: *Nederl. Akad. Wetensch., Proc.* 53. Also published in Indagationes Mathematicae, 12:386–392, pp. 386–392.