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

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

Deliverable D4.2:

Advanced urban tropospheric NO2 product ATBD

Date: 2026-05-12

Version: 2.1

Contributing authors:

A. Richter, K. Krause, T. Bösch (IUP Bremen)
S. Beirle, S. Ziegler and T. Wagner (MPIC)
M. M. Friedrich, M. van Roozendael, G. Pinardi, A. Merlaud, C. Fayt (BIRA)
A. Piters (KNMI)
D. Karagkiozidis, D. Nikolis, A. Bais (AUTH)
U. Frieß (IUP Heidelberg)

Content

Acronyms and abbreviations

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

ATBD	Algorithm Theoretical Basis Document
DOAS	Differential Optical Absorption Spectroscopy
FRM_4DOAS	Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations
MAX-DOAS	Multi AXis Differential Optical Absorption Spectroscopy
MMF	Mexican MAX-DOAS Fit
MAPA	Mainz Profile Algorithm
SZA	Solar zenith angle
VCD	Vertical Column Density
WP	Work Package
SD	Standard deviation

1 Introduction

1.1 Purpose of the ATBD

This Algorithm Theoretical Basis Document (ATBD) describes the algorithms and parameters proposed to provide users of the tropospheric NO_2 product processed by the Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations (FRM4DOAS) centralised processing system with information on the homogeneity of the atmospheric NO_2 field during the measurements. Such information can be important for satellite validation but also in its own right.

The aim is to implement the simple algorithms outlined here in the FRM4DOAS processor and to include the output in the FRM4DOAS product. It is independent of the already implemented uncertainty estimates and quality flagging and is for information purposes only. Any filtering for certain criteria will have to be applied by the users of the product.

1.2 Document overview

The document starts with a brief consideration of which quantities could be used to characterize the spatial inhomogeneity of the NO_2 distribution and which choices were made for the implementation in Section 2. This is followed by a mathematical description of the parameters selected for implementation in Section 3. Section 0 provides some details for the implementation of the parameters, followed by the analysis of the uncertainty components in Section 5 and a short summary in Section 6.

2 Design considerations and choices

MAX-DOAS measurements can provide information on the spatial variability of the absorber fields in several different ways:

1. Directly by taking measurements in **different azimuthal directions**. Any differences in the retrieved tropospheric columns indicate variations in the absorber distribution. Depending on the location of the MAX-DOAS instrument, there are two possible situations: a) a persistent gradient, when an instrument is located, for example, in the suburbs of a city and has viewing directions towards the city and to background directions, or b) variable values from air masses transported towards the instrument. In the first case, similar differences would be found in many measurements whereas in the second case, the differences between the measurements in the different azimuthal directions would depend on the wind direction and speed, as well as the on distance of the air mass with the increased absorber concentration from the instrument.
2. Indirectly by evaluating the variation of the observed column over **time**. In most cases, rapid variations of the observed columns are linked to transport of air masses having different absorber concentrations, and therefore, temporal changes are an indication of spatial variability. Slow changes on the other hand could also be related to emission changes or local build-up or dilution of pollution for example from changes in boundary layer height.
3. Another criterion linked to spatial inhomogeneity is the **difference between UV and vis** retrievals. As Rayleigh scattering is more effective in the UV part of the spectrum, UV measurements integrate over a shorter horizontal range than measurements in the visible part of the spectrum. Any inconsistencies

between the UV and vis columns are therefore indicative of a change in concentration along the viewing direction. However, this was not investigated in this work package.

Not all MAX-DOAS measurements have observations in several azimuthal directions, and not all instruments cover both the UV and the visible part of the spectrum. Therefore, it makes sense to include several different parameters in the FRM4DOAS products to cover as many situations as possible.

An important question is on which quantity to perform the tests. The obvious choice is to use the vertical columns provided by MAPA and MMF, as these are the final products delivered to the users. In the current version of the FRM4DOAS product, MMF profiles and vertical columns are the main output, while MAPA data is mainly used for cross-validation and flagging. A disadvantage of using the vertical columns is that the representative measurement volume of the retrieval results is difficult to assess, as it combines measurements at different elevation angles, which are sensitive to different horizontal distances from the instrument. The largest sensitivity to horizontal variations is expected for the lowest elevation angles, which would therefore be another possible choice for the quantity to base the variability criterion on. However, the length of the light path depends on clouds, aerosols, and the relative azimuth towards the sun, and it would be difficult to include these dependencies in the evaluation of the variability from slant columns. Therefore, here only vertical columns are used.

When deciding on which quantities to use for detecting horizontal variability, one has two different options:

1. Use of **absolute variations**. This approach would emphasize large changes in the columns, which in the case of satellite validation would lead to large absolute uncertainties in the comparison. It is robust for small columns.
2. Use of **relative variations**. This approach is more appropriate when the internal consistency of measurements is to be described, as variations by, for example, a factor of two would be seen as large, independent of the absolute columns

Here, it was decided to provide both flavours of variability to the users so that they can pick what best fits their needs.

Another choice to be made is the reference to which to refer for the variability. In particular for the time-based criteria, different choices are conceivable:

1. Comparison between **successive observations**. This is a simple method, but can lead to amplification of noise if the individual measurements have significant uncertainties.
2. Comparison to the **median values** over a certain time period. This is a more robust approach, and in many cases, no systematic temporal variation is expected for tropospheric absorbers, making the median a good estimator. However, the question is which time period to use for the median.
3. Comparison to a **modelled** expected time evolution, for example by a fitted polynomial. This approach is already implemented for the zenith-sky measurements in the FRM4DOAS processor, as total columns have a strong dependency on the SZA for species with significant stratospheric contributions. For tropospheric absorbers, a systematic variation is usually not expected.

Here, the approach is to use median values as the reference.

Finally, the question is how to provide the selected parameters:

1. As **flags**, indicating, for example, rather homogeneous or suspected inhomogeneous or certainly inhomogeneous scenarios. This type of parameter is straightforward to use, but cannot easily be changed after processing of the product.
2. As **continuous indicators**, giving the variability, for example in per cent or molecules per cm^2. These indicators can then be used with different thresholds for different applications, making them more flexible. On the other hand, the users have to decide how to apply them.

Here, it was decided to opt for maximum flexibility and, therefore, continuous indicators are used.

3 Definition of inhomogeneity parameters

The following six indicators are used in the current version of the test processor:

3.1 Temporal variation

A running standard deviation in a window ±60 minutes, leading to a total window duration of 120 minutes, around the current measurement with azimuth *azi* and time *t* is calculated for all measurements with the same azimuth direction:

$$inhomogene\ time\ vcd\ absolute = sd\left(VCD_{azi,t-60\ \mathrm{minutes}}, \dots, VCD_{azi,t+6\quad minutes}\right) \tag{1}$$

60 minutes before and after were chosen to ensure that a suitable standard deviation and median can be calculated in almost every case. Previous versions used 240-minute time windows, but it turned out that 120 minutes are sufficient for all stations. For stations scanning in several azimuth directions, shorter time windows might lead to cases where the median would be calculated out of one or two VCD, which could lead to cases where some indicators only return NaN, because at least three values are required for the calculation.

The units of this indicator are molecules per cm^2.

Additionally, a relative indicator is calculated using the running median in the same timeframe for normalization.

$$inhomogene\ time\ vcd\ relative = abs\left(\frac{sd\left(VCD_{azi,t-60\ \mathrm{minutes}}, \dots, VCD_{azi,t+60\ minutes}\right)}{median\left(VCD_{azi,t-60\ \mathrm{minutes}}, \dots, VCD_{azi,t+60\ minutes}\right)}\right) \tag{2}$$

This indicator is unitless and has values between 0 and infinity, 1 corresponding to a variation of 100%.

As an additional criterion for the calculation of the indicator, at least three VCDs must be in the given time window; the number of values used for calculation is saved in an additional variable, which can be accessed by users.

3.2 Azimuthal variation

Azimuthal variation is computed considering sequences of VCDs from scans in different azimuth direction, limited to one VCD per azimuth direction in a given time frame. For each VCD, the matching VCDs of other azimuth directions are identified, and the standard deviation of this group of VCDs is calculated. To reduce the influence of temporal variation for each azimuth, only the VCD closest in time to the current VCD, and with a time difference of less than 60 minutes, is used in the calculation.

$$inhomogene\ azimuth\ vcd\ absolute = sd(VCD_{azi\ 1}, VCD_{azi\ 2}, \ldots, VCD_{azi\ n}) \tag{3}$$

The units of this indicator are molecules per cm^2.

As was the case for the relative temporal in-homogeneity indicator, the relative azimuth in-homogeneity indicator also uses the median of all VCDs in each set for normalization.

$$inhomogene\ azimuth\ vcd\ relative = abs\left(\frac{sd(VCD_{azi\ 1}, VCD_{azi\ 2}, \ldots, VCD_{azi\ n})}{median(VCD_{azi\ 1}, VCD_{azi\ 2}, \ldots, VCD_{azi\ n})}\right) \tag{4}$$

This indicator is unitless and has values between 0 and infinity, 1 corresponding to a variation of 100%.

The number of values taken into account is saved in an additional variable and the indicators are only calculated if there are at least two different azimuth directions.

3.3 Combination of azimuthal and temporal variation

To combine temporal and azimuthal variation into a single indicator, the rolling standard deviation in a window of ±60 minutes around the current VCD at time *t*, independent of the azimuth direction, is calculated:

$$inhomogene\ time\ azimuth\ vcd\ absolute = sd(VCD_{t-60\ \text{minutes}}, \ldots, VCD_{t+60\ minutes}) \tag{5}$$

The units of this indicator are molecules per cm^2.

For the relative indicator, this is again normalized by the respective median of the same group of values:

$$inhomogene\ time\ azimuth\ vcd\ relative = abs\left(\frac{sd(VCD_{t-60\ \text{minutes}}, \ldots, VCD_{t+60\ minutes})}{median(VCD_{t-60\ \text{minutes}}, \ldots, VCD_{t+60\ minutes})}\right) \tag{6}$$

This indicator is unitless and has values between 0 and infinity, 1 corresponding to a variation of 100%.

As an additional criterion for the calculation of the indicator, at least three VCDs must be in the given time window; the number of values used for the calculation is saved in an additional variable.

4 Calculation of representative VCDs

For each type of inhomogeneity indicator (time, azimuth, and the combination of time and azimuth), a set of representative VCDs is calculated. For the calculation, the VCDs are filtered using the respective inhomogeneity indicator, either absolute for VCDs that are below the station's yearly median or relative for VCDs above the station's yearly median. The thresholds are 1e15 molec/cm² for the absolute indicators and 20% for the relative indicators. If the VCD's inhomogeneity indicator is above the given threshold, it is not used for the calculation of the representative column, as it is considered to be inhomogeneous.

Each representative VCD is computed from sequence of scans with different azimuth directions ($azi_1, azi_2, \ldots, azi_n$). For each filtered VCD, the filtered VCDs of other azimuth directions are identified and the mean of this group of filtered VCDs is calculated and serves as the representative column. For each azimuth only the filtered

VCD closest in time to the current filtered VCD and with a time difference lower than 60 minutes is used for the calculation.

$$representative\ VCD = mean(VCD_{azi\ 1}, VCD_{azi\ 2}, \ldots, VCD_{azi\ n}) \quad (7)$$

The error of the representative VCDs is calculated through error propagation using the uncertainty of the individual VCDs (σ_i):

$$error\ of\ representative\ VCD = \frac{1}{n}\sqrt{\sum_{i=1}^{n} \sigma_i^2} \quad (8)$$

For cases where only a single valid VCD remained after filtering, the representative VCD is this single VCD with its respective uncertainty value. For stations with only a single azimuth direction, the representative column are the VCDs that are not flagged as inhomogeneous, without any averaging.

For each inhomogeneity indicator, a new subgroup is created within the respective MAPA or MMF group. Each row in this subgroup represents a representative VCD on an individual time grid, independent of the scan index or other time grids in the file. Each row includes:

1. the representative column,
2. the error of the representative column,
3. the datetime of the earliest used VCD,
4. the average datetime of all used VCDs,
5. the datetime of the last used VCD, and
6. 15 variables indicating which scan index numbers were used to calculate the representative VCD.

These 15 variables are filled only as needed. For example, if a representative VCD is calculated from 3 individual VCDs, variables 0 to 2 are populated with the respective scan index values, while the remaining variable are left empty (set to NaN).

5 Implementation

The algorithms are currently implemented in Python 3 and use pandas data structures and functions for processing. Pandas was chosen because it offers high performance and provides functions for easy data selection and manipulation of the tabular data in the FRM4DOAS files. Pandas also conveniently allows working with missing data or NaN values without taking extra steps. Within the algorithm, the VCDs and ancillary data are stored in a pandas data frame, which is a two-dimensional table, where each column represents a variable (VCD, timestamp, azimuth angle, corresponding quality flag, …) and each row represents the output of the MMF or MAPA algorithms for one profile retrieval. The Data frame is sorted according to timestamp, and all azimuth directions are combined into a single table. For all algorithms only VCDs with a qa_value < 2 are considered.

For the temporal variation, the main part of the algorithm then looks like the following pseudo-code:

```
for azimuth in azimuths:
```

```
subset = data_df.query("telescope_azimuth_angle == @azimuth")
subset_sd = subset.rolling("120min", on="datetime_vcd",
                            center=True, min_periods=3).sd()
subset[column+"_sd"] = subset_sd[column]

subset_median = subset.rolling("120min", on="datetime_vcd",
                            center=True, min_periods=3).median()
subset[column+"_median"] = subset_median[column]

subset_num = subset.rolling("120min", on="datetime_vcd",
                            center=True, min_periods=3).count()
subset[column+"_number"] = subset_num[column]

for index, row in subset.iterrows():
  data_df.loc[index, "flag_rel"] = abs([column+"_sd"]/row[column+"_median"])
  data_df.loc[index, "flag_abs"] = abs(row[column+"_sd"])
  data_df.loc[index, "flag_number"] = abs(row[column+"_number"])
```

The algorithm iterates over all azimuth directions and extracts the data of one azimuth direction. The “rolling” function iterates over all rows and selects all appropriate data in a time window 60 minutes before and 60 minutes after the current row. The functions “.sd()” ,“.median()” and “.count() then calculate the standard deviation, median, and number of observations using the values selected by “rolling”. If no data is found for the given time window, the functions return a NaN value for the respective row.

The result is then stored in a new variable with the same dimensions as “subset”. Each subset is then stored at the appropriate scan indices in the main dataframe.

The pseudo code for the azimuthal variation looks like the following lines:

```
data_df["nearest_values"] = data_df.apply(lambda row:
                                compute_nearest_values(row, data_df, azimuths, column), axis=1)
for index, row in data_df.iterrows():
  if len(row["nearest_values"]) > 1:
    data_df.loc[index, "flag_abs"] = np.nansd(row["nearest_values"])
    data_df.loc[index, "flag_rel"] = np.nansd(row["nearest_values"]) / np.nanmedian(row["nearest_values"])
    data_df.loc[index, "flag_number"] = np.count_nonzero(row["nearest_values"])
  else:
    data_df.loc[index, "flag_abs"] = np.nan
    data_df.loc[index, "flag_rel"] = np.nan
    data_df.loc[index, "flag_number"] = np.nan
```

For each measurement, the respective measurements of other azimuth directions are identified and stored as a new variable in the data frame. In the next step, the stored values are used to calculate the standard deviations, medians, and number of observations, and these results are then saved alongside the original data. In cases where no adjacent VCDs from other azimuths are identified, the respective indicators are set to NaN.

Finally, the pseudo code for the combined temporal and azimuthal variation:

```
data_df_sd = data_df.rolling("120min", on="datetime_vcd", center=True, min_periods=3).sd()
data_df_median = data_df.rolling("120min", on="datetime_vcd", center=True, min_periods=3).median()
data_df_number = data_df.rolling("120min", on="datetime_vcd", center=True, min_periods=3).count()
data_df[column+"_roll_sd"] = data_df_sd[column]
data_df[column+"_roll_median"] = data_df_median[column]
data_df[column+"_roll_number"] = data_df_number[column]

for index, row in data_df.iterrows():
    data_df.loc[index, "flag_rel"] = abs(row[column+"_roll_sd"]/row[column+"_roll_median"])
    data_df.loc[index, "flag_abs"] = abs(row[column+"_roll_sd"])
    data_df.loc[index, "flag_number"] = abs(row[column+"_roll_number"])
```

This works in a similar way as the inhomogeneity indicator for time only, but as all azimuths are considered at the same time, the outer loop over all possible azimuths is removed.

6 Uncertainty

To assess the influence of filtering using the different inhomogeneity indicators on the systematic uncertainty components, the mean VCDs for each station have been calculated for unfiltered VCDs and for inhomogeneity filtered, representative VCDs. Filtering for inhomogeneity is done in a similar way as in the calculation of the representative columns. The VCDs were filtered based on their absolute inhomogeneity indicator when they exceeded the station's median VCD and on the relative indicator when below the station's median VCD, with thresholds of 1e15 molec/cm² and 20%, respectively. Calculations were performed for both MAPA and MMF. For simplicity, in the tests for MAPA only the version with O_4 scaling factor of 1.0 was used. The different viewing directions of the Mainz station were processed in a combined setup.

For the MAPA algorithm, filtering for inhomogeneity reduces the mean VCDs for almost all stations, which indicates that the filtering using the inhomogeneity indicators introduces a bias towards lower values for these stations. The only exception is the Xianghe_1669_2 channel, where slightly higher average VCDs are found when filtering for inhomogeneity in time. In all cases, the differences are small, and the mean VCDs have similar magnitudes.

The MMF algorithm shows similar results, with bias towards lower values when filtering, but an exception for Xianghe_1669_2 and Thessaloniki_1683_1, where the mean VCDs increase when filtering for specific inhomogeneities.

Table 1: Mean values of unfiltered and representative MAPA vertical column densities based on inhomogeneity indicators. The VCDs were filtered based on their absolute inhomogeneity indicator when they exceeded the station's median VCD and with the relative indicator when below the stations' median VCD, with thresholds of 1e15 molec/cm² and 20% respectively. If there is no value in the first data column (Mean representative VCD inhomogeneity azimuth) the station only has a single azimuth direction and the azimuthal inhomogeneity cannot be calculated. Sometimes these stations have short periods of time (e.g. a day out of several years of data) with more than one azimuth, in these cases the azimuthal inhomogeneity filtered mean is also not shown in the table, as it covers a much smaller time frame than the other indicators.

Station	**Mean representative**	**Mean representative**	**Mean representative**	**Mean unfiltered VCD**

	VCD inhomogeneity azimuth [molec/cm²]	VCD inhomogeneity time and azimuth [molec/cm²]	VCD inhomogeneity time [molec/cm²]	 [molec/cm²]
ATHENS_1674_1_uv	2.46e+15	1.56e+15	2.34e+15	4.92e+15
ATHENS_1674_1_vis	2.78e+15	1.83e+15	2.79e+15	5.66e+15
BREMEN_1673_1	4.82e+15	4.60e+15	4.82e+15	9.58e+15
BREMEN_1673_2	5.31e+15	4.99e+15	5.07e+15	9.91e+15
CABAUW_1690_1		5.11e+15	5.11e+15	7.37e+15
CABAUW_1693_1		4.65e+15	4.79e+15	1.02e+16
CABAUW_1693_2		5.74e+15	5.74e+15	1.07e+16
DEBILT_1675_1		8.35e+15	8.35e+15	1.51e+16
DEBILT_1684_1		5.72e+15	5.72e+15	1.75e+16
DEBILT_1688_1		8.66e+15	8.66e+15	2.83e+16
HEIDELBERG_1679_1		5.02e+15	5.02e+15	8.65e+15
HEIDELBERG_1679_2		6.30e+15	6.30e+15	9.92e+15
MAINZ_1676	8.95e+15	9.14e+15	8.79e+15	9.90e+15
SAN.PIETRO.CAPOFIUME_1695_1	4.46e+15	4.25e+15	4.34e+15	7.38e+15
SAN.PIETRO.CAPOFIUME_1695_2	3.70e+15	3.58e+15	3.60e+15	6.92e+15
THESSALONIKI_1683_1	7.51e+15	7.23e+15	8.00e+15	9.16e+15
THESSALONIKI_1686_1	7.08e+15	6.36e+15	4.88e+15	9.46e+15
THESSALONIKI_1699_1	8.03e+15	7.31e+15	6.37e+15	9.71e+15
THESSALONIKI_1699_2	6.86e+15	6.18e+15	5.39e+15	8.21e+15
TORONTO_1694_1		7.79e+15	8.39e+15	1.24e+16
TORONTO_1694_2		7.15e+15	7.50e+15	1.13e+16
UCCLE_1670_1	6.13e+15	6.65e+15	6.70e+15	1.25e+16
UCCLE_1670_2	6.38e+15	6.10e+15	6.27e+15	1.37e+16
XIANGHE_1669_1		1.48e+16	1.69e+16	1.78e+16

XIANGHE_1669_2		1.83e+16	2.16e+16	1.92e+16

Table 2: Same as Table 1 but for MMF values.

Station	**Mean representative VCD inhomogeneity azimuth [molec/cm²]**	**Mean representative VCD inhomogeneity time and azimuth [molec/cm²]**	**Mean representative VCD inhomogeneity time [molec/cm²]**	**Mean unfiltered VCD [molec/cm²]**
ATHENS_1674_1_uv	3.99e+15	3.02e+15	3.80e+15	4.99e+15
ATHENS_1674_1_vis	3.64e+15	2.54e+15	3.55e+15	5.49e+15
BREMEN_1673_1	6.18e+15	5.88e+15	5.71e+15	7.37e+15
BREMEN_1673_2	6.55e+15	6.21e+15	6.16e+15	7.67e+15
CABAUW_1690_1		5.93e+15	5.93e+15	7.01e+15
CABAUW_1693_1		6.96e+15	6.96e+15	8.00e+15
CABAUW_1693_2		6.83e+15	6.83e+15	7.79e+15
DEBILT_1675_1		8.82e+15	8.82e+15	9.96e+15
DEBILT_1684_1		6.96e+15	6.95e+15	8.86e+15
DEBILT_1688_1		1.06e+16	1.06e+16	1.24e+16
HEIDELBERG_1679_1		6.48e+15	6.48e+15	7.95e+15
HEIDELBERG_1679_2		7.15e+15	7.15e+15	8.34e+15
MAINZ_1676	8.41e+15	8.17e+15	8.23e+15	9.06e+15
SAN.PIETRO.CAPOFIUME_1695_1	5.40e+15	5.33e+15	5.33e+15	6.10e+15
SAN.PIETRO.CAPOFIUME_1695_2	5.14e+15	5.11e+15	5.11e+15	6.05e+15
THESSALONIKI_1683_1	1.02e+16	1.02e+16	1.03e+16	8.71e+15
THESSALONIKI_1686_1	7.79e+15	7.21e+15	5.50e+15	8.79e+15
THESSALONIKI_1699_1	9.18e+15	8.67e+15	7.40e+15	1.06e+16
THESSALONIKI_1699_2	7.88e+15	7.18e+15	6.26e+15	8.25e+15
TORONTO_1694_1		8.01e+15	8.21e+15	9.75e+15

TORONTO_1694_2		7.44e+15	7.63e+15	8.97e+15
UCCLE_1670_1	7.06e+15	7.47e+15	7.53e+15	8.91e+15
UCCLE_1670_2	7.48e+15	7.30e+15	7.45e+15	9.14e+15
XIANGHE_1669_1		1.29e+16	1.36e+16	1.58e+16
XIANGHE_1669_2		1.79e+16	1.93e+16	1.70e+16

To assess the influence of filtering for inhomogeneity on the random uncertainty component introduced by filtering for inhomogeneity, the azimuthal standard deviations of the representative columns are compared to the azimuthal standard deviations of the unfiltered columns. The results are shown in Figure 1 and Figure 2 for the MAPA algorithms, and in Figure 4 and Figure 5 for the MMF algorithm. For stations with only a single azimuth direction the rolling standard deviation in time has been compared for unfiltered VCDs and inhomogeneity in time filtered VCDs (MAPA:Figure 3, MMF: Figure 6).

For most cases the standard deviations of the representative VCDs are lower than the standard deviation for the respective unfiltered data, which indicates a reduction in the random uncertainty of the representative columns. Notable exception is Mainz where the reduction in standard deviation is not as strong and also there are times, where both unfiltered and filtered standard deviation have a similar magnitude.

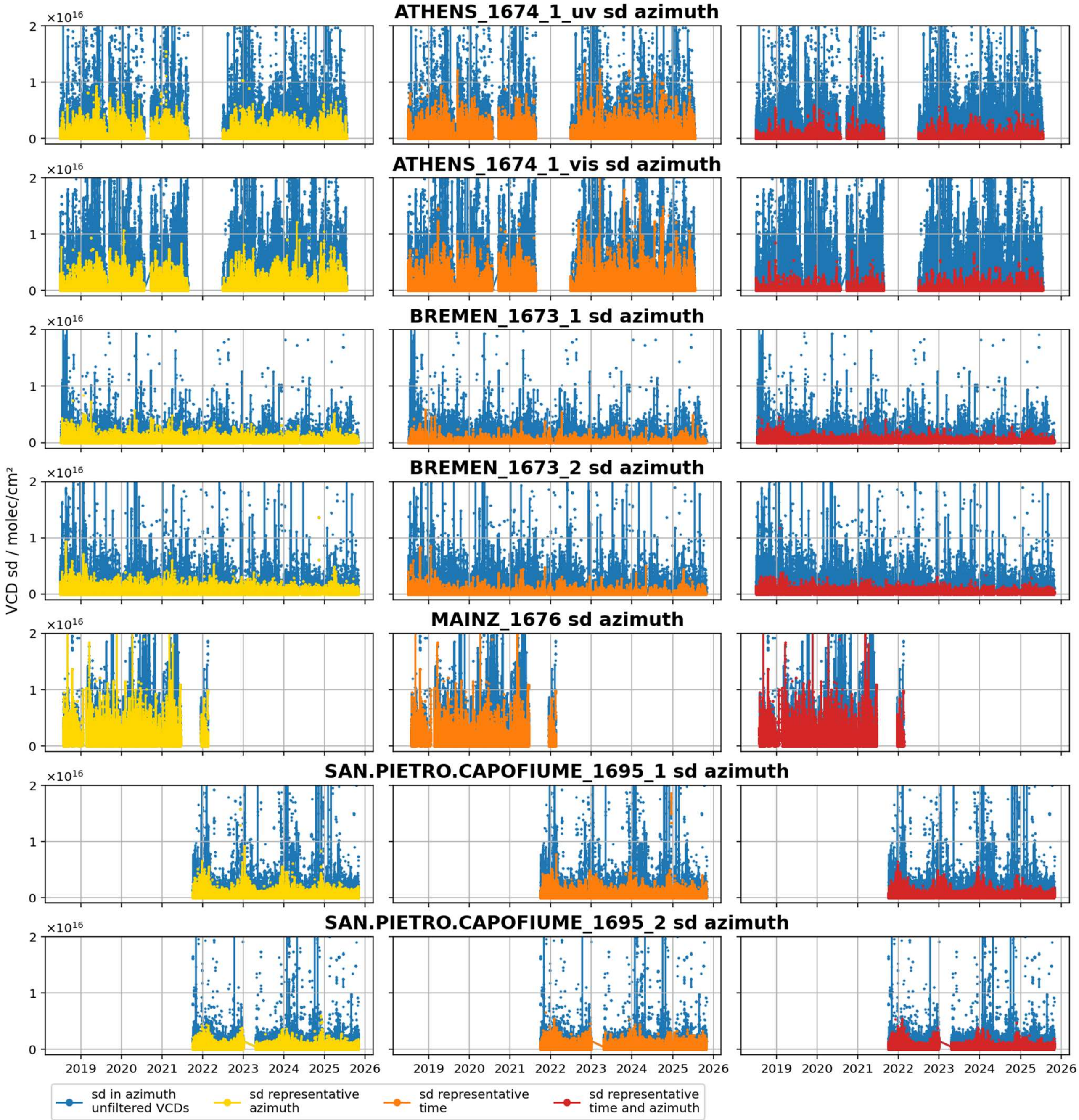

Figure 1: MAPA timeseries of the azimuthal standard deviations of the unfiltered VCDs and the representative columns for stations with more than one azimuth direction.

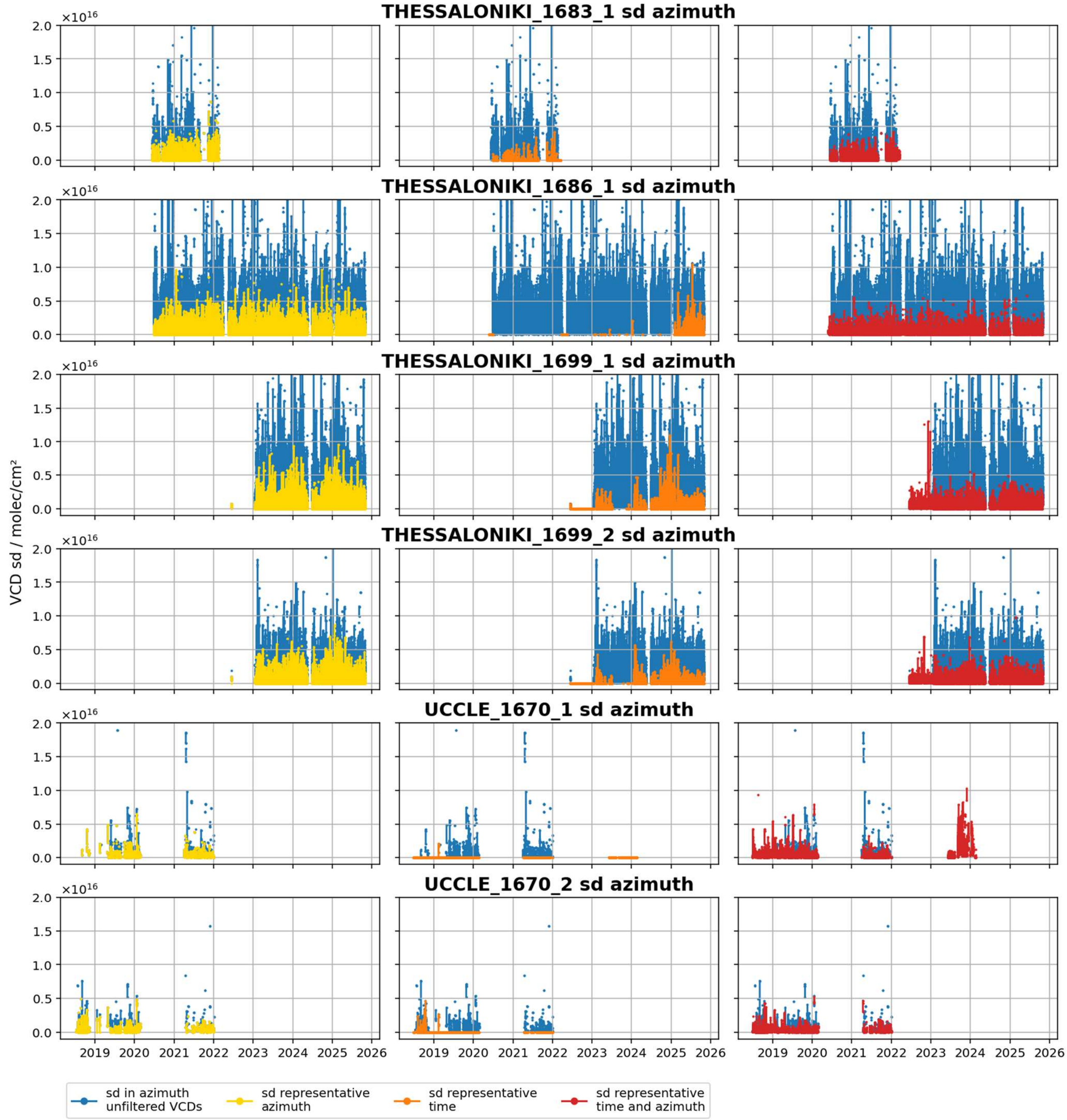

Figure 2: MAPA timeseries of the azimuthal standard deviations of the unfiltered VCDs and the representative columns for stations with more than one azimuth direction.

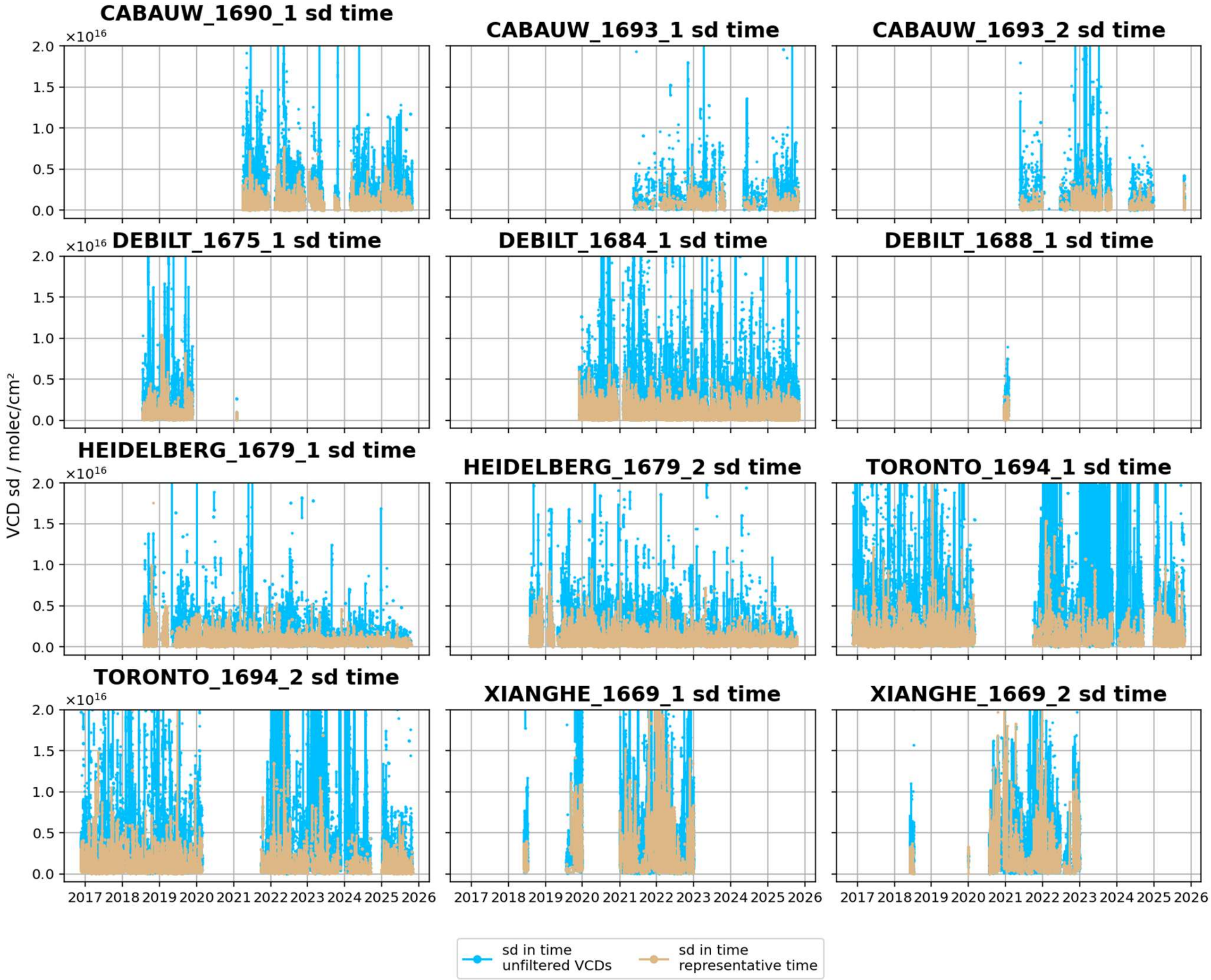

Figure 3: MAPA timeseries of the rolling standard deviation in time for unfiltered and representative columns filtered for inhomogeneity in time for stations with a single azimuth direction.

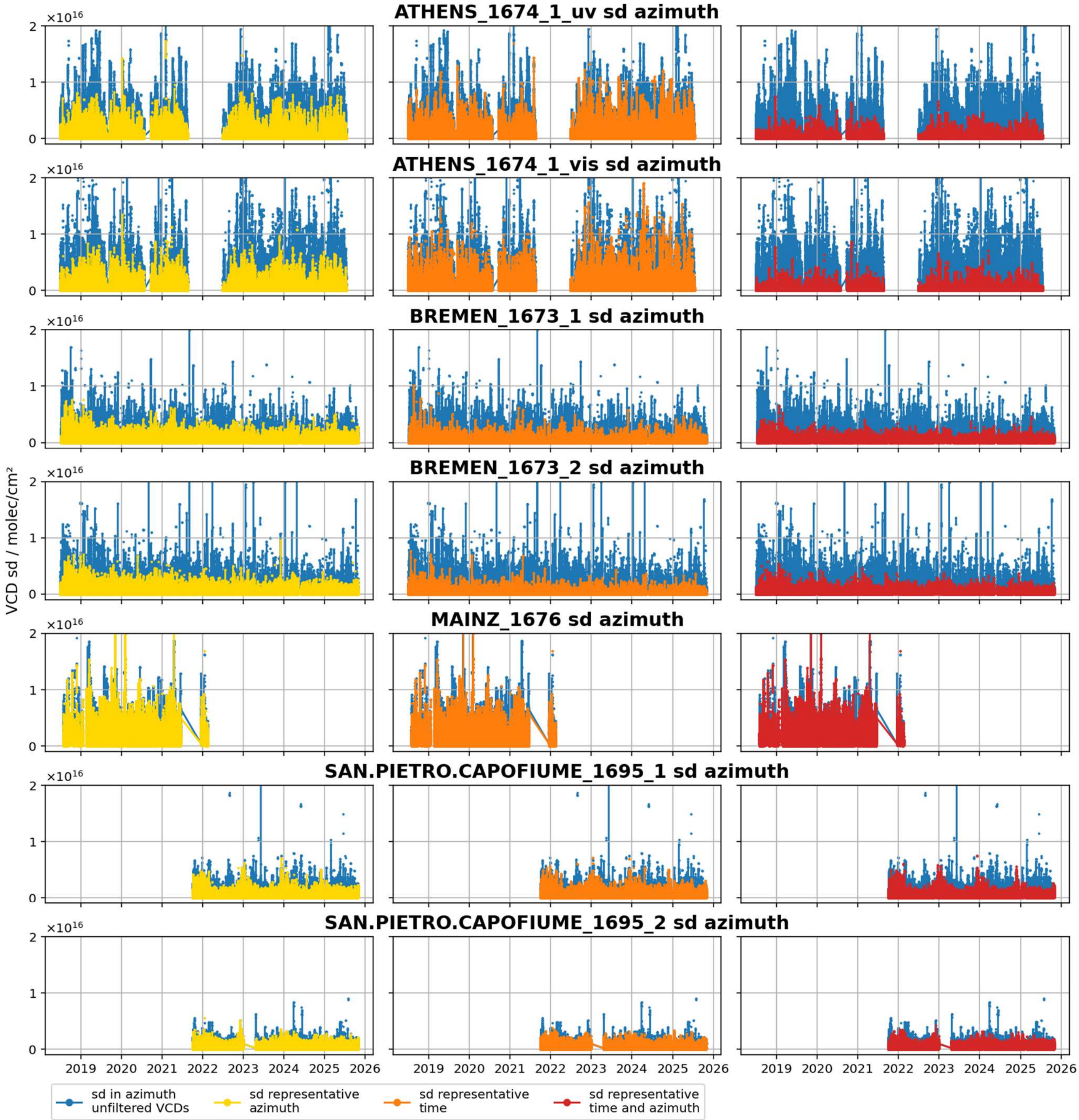

Figure 4: MMF timeseries of the rolling standard deviation in time for unfiltered and representative columns filtered for inhomogeneity in time for stations with a single azimuth direction

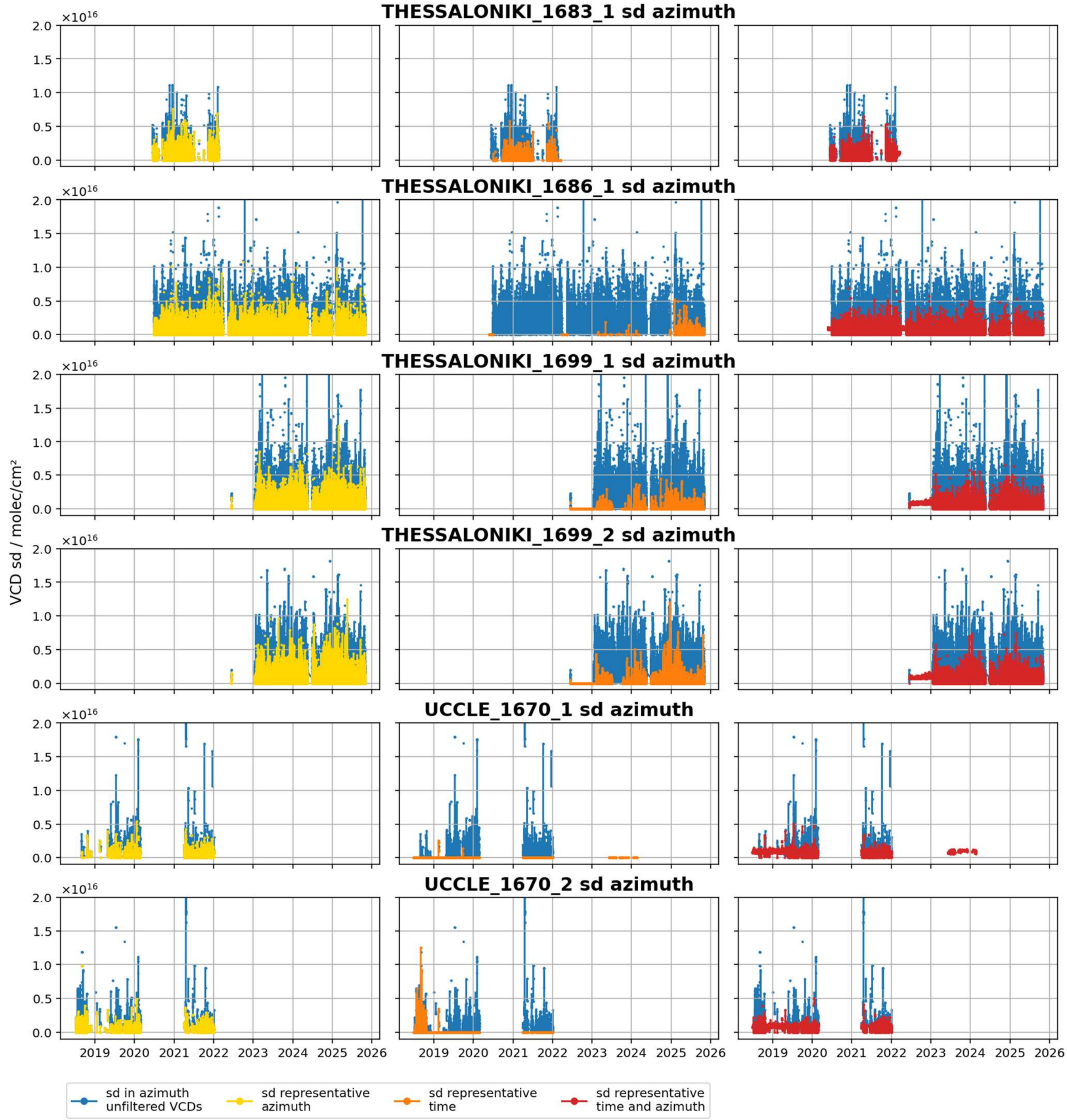

Figure 5: MMF timeseries of the rolling standard deviation in time for unfiltered and representative columns filtered for inhomogeneity in time for stations with a single azimuth direction

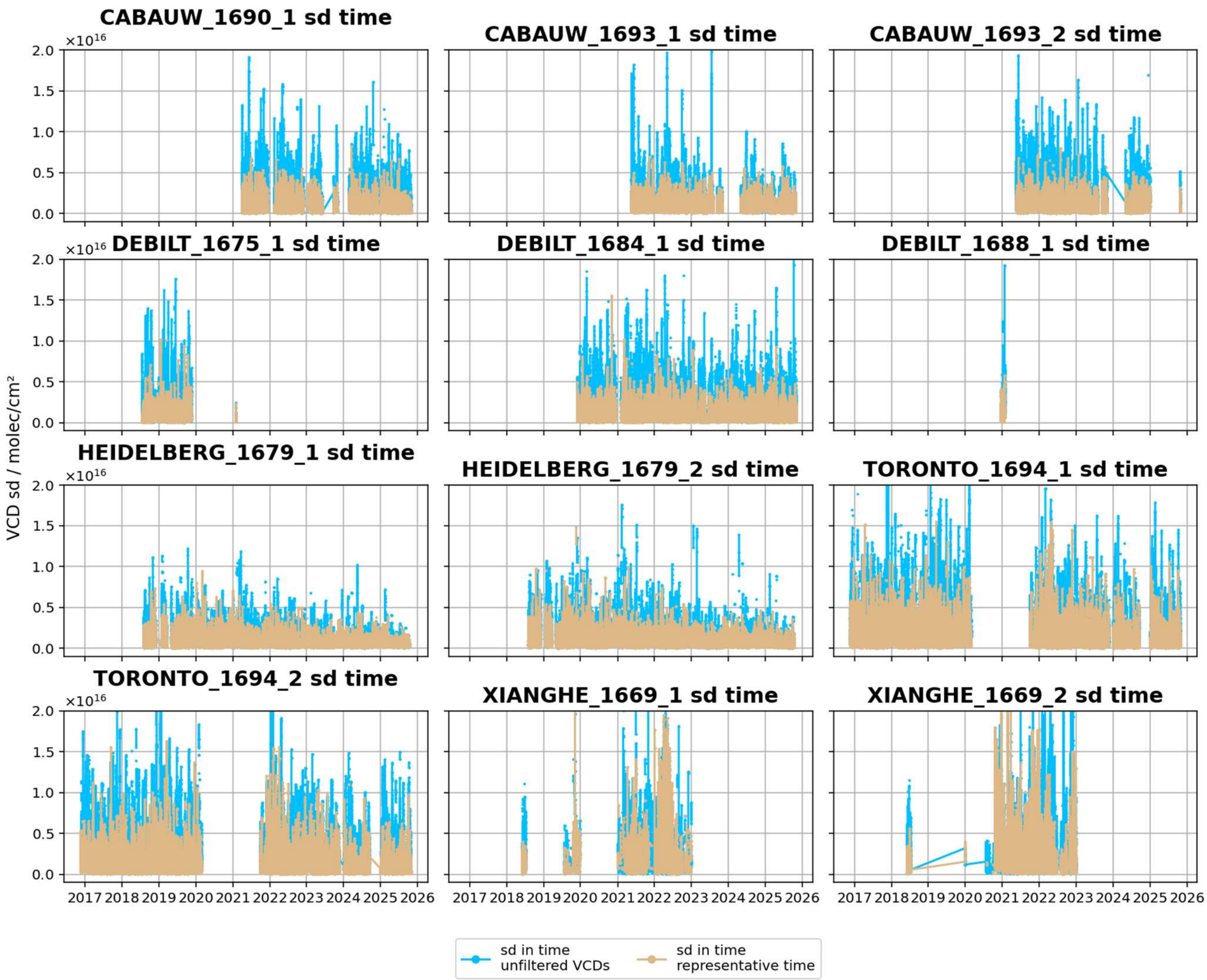

Figure 6: MMF timeseries of the rolling standard deviation in time for unfiltered and representative columns filtered for inhomogeneity in time for stations with a single azimuth direction

7 Summary

Inhomogeneity indicators that represent inhomogeneity in time, azimuth and the combination of time and azimuth were implemented and tested on the data of 12 FRM4DOAS stations, with a total of 27 different channels. Additionally, representative mean VCDs and their respective uncertainty are calculated using the inhomogeneity indicators and predefined thresholds. Filtering the VCDs for inhomogeneity reduces the random uncertainty of the VCDs, but at the same time also introduces a bias towards lower VCDs compared to unfiltered data.