Precursors_cci+

ESA Climate Change Initiative (CCI)

D2.2 End to End ECV Uncertainty Budget (E3UB) – Version 2

Title	D2.2 End to End ECV Uncertainty Budget
Reference	Precursors_cci+_D2.2_E3UB_02_04
Issue	02
Revision	04
Status	Final
Date of issue	11/05/2025
Document type	Deliverable

Generated by KNMI

Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

This work has received funding from the European Space Agency under Contract No 4000138243 and is the result of the cooperation between the partners of the Precursors_cci+ consortium.

	FUNCTION	NAME	DATE
LEAD AUTHOR(S)		Folkert Boersma Isolde Glissenaar Sora Seo Isabelle De Smedt Nicolas Theys Thomas Danckaert Maya George Lieven Clarisse	30/01/2025
REVIEWED BY	Science Leaders	M. Van Roozendael Folkert Boersma	13/05/2025 13/05/2025
	ESA Project Officer	Simon Pinnock	../../….
DISTRIBUTED TO	ESA Precursors_cci+ Consortium		

DOCUMENT CHANGE RECORD

Issue	Revision	Date	Modified items	Observations
2	0	06/06/2024	Initial template	This is D2.2 for cycle 2 of the ESA CCI+ Precursor project
	1	12/11/2024	First version	Completed the NO_2 L3 part
	2	17/01/2025	Second version	Updated with sections for HCHO, CO and NH3 L3, and input orginal uncertainty description from cycle 1.
	3	30/01/2025		Updated with section for NO2 morning, SO2, and CHOCHO
	4	11/05/2025	Fourth version	RIDs have been addressed

INHOUDSOPGAVE

Executive Summary **7**

1. Purpose and scope **8**

1.1. Purpose *8*

1.2. Scope *8*

1.3. Applicable documents *8*

1.4. Reference documents *9*

1.5. List of acronyms *9*

2. Terminology **12**

3.1 Tropospheric nitrogen dioxide (L2-NO2) *13*

3.1.1. Uncertainty propagation 13

3.1.2. NO_2 slant column density uncertainty estimates 14

3.1.3. NO_2 stratospheric correction uncertainty estimate 16

3.1.4. Tropospheric air mass factor uncertainty estimates 17

3.1.5. Overall NO_2 retrieval uncertainties 18

*3.2. **Formaldehyde (L2-HCHO)*** *19*

3.2.1. Uncertainty propagation 19

3.2.2. HCHO slant column density uncertainty estimates 19

3.2.3. Tropospheric air mass factor uncertainty estimates 20

3.2.4. Background correction uncertainty estimates 20

3.2.5. Overall HCHO retrieval uncertainties 21

*3.3. **Sulfur dioxide (L2-SO2)*** *23*

3.3.1. Uncertainty propagation 23

3.3.2. SO_2 slant column density uncertainty estimates 24

3.3.3. Tropospheric air mass factor uncertainty estimates 24

3.3.4. Overall SO_2 retrieval uncertainties 25

*3.4. **Glyoxal (L2-CHOCHO)*** *26*

3.4.1. Uncertainty propagation 26

3.4.2. Slant column errors 26

3.4.3. Air mass factor errors 27

3.4.4. Background correction error 28

3.4.5. Global budget: Total precision and trueness 28

*3.5. **Carbon monoxide (L2-CO)*** *31*

3.5.1. CO uncertainties on global scale 31

3.5.2. CO uncertainties on local scale 33

*3.6. **Ammonia (L2-NH3)*** *36*

3.6.1. Uncertainty propagation 36

3.6.2. Uncertainty on the input parameters 36

3.6.3. Uncertainty characterization of L2-NH_3 38

*4.1. **Uncertainty assessment for nitrogen dioxide (L3-NO2) from OMI and TROPOMI*** *41*

4.1.1. Procedure 41

4.1.2. Spatial averaging 42

4.1.3. Temporal averaging 45

4.1.4 Analysis of L3 uncertainty budget for TROPOMI 51

4.2 Uncertainty assessment for nitrogen dioxide (L3-NO2) from morning sensors 54
4.2.1 Overview 54
4.2.2 Data gridding approach 54
4.2.3 L3 uncertainty budget estimate and data quality for GOME-2A 58

4.3 Uncertainty assessment for formaldehyde (L3-HCHO) 65
4.3.1 Random and systematic uncertainties provided in the Level 2 data products 65
4.3.2 Averaging Procedure 65
4.3.3 L3 uncertainty estimate and data quality 65
4.3.4 Recommendations regarding the use of HCHO L3 data 70

4.4 Uncertainty assessment for sulphur dioxide (L3-SO2) 71
4.4.1 Uncertainty provided in the Level 2 data products 71
4.4.2 Averaging procedure 71
4.4.3 L3 uncertainty estimate and data quality 72

4.5 Uncertainty assessment for glyoxal (L3-CHOCHO) 74
4.5.1 Procedure 74
4.5.2 L3 uncertainty estimate and data quality 74

4.6 Uncertainty assessment for carbon monoxide (L3-CO) 76
4.6.1 Procedure 76
4.6.2 Analysis of the L3 uncertainty budget 77

4.7 Uncertainty assessment for ammonia (L3-NH3) 80
4.7.1 Uncertainty on the Level 2 80
4.7.2 Averaging Procedure 81
4.7.3 Examples 82
4.7.4 General recommendations 85

5 References 87

Executive Summary

This document provides the theoretical uncertainty assessment of the ECV Precursor level-2 and level-3 datarecords generated for NO_2, HCHO, SO_2, CO, NH_3 and glyoxal in the ESA CCI+ ECV Precursor project.

The purpose of the level-2 part is to characterize the uncertainties in the entire retrieval process, from the impact of instrumental noise and degradation to uncertainties associated with the required input datasets. This leads to estimates of uncertainty magnitudes and patterns in time and space, which informs users about the relative strength of the level-2 satellite retrievals. A distinction is made between random and systematic contributions to the overall (pixel-specific) uncertainties. The document explains how the uncertainties are estimated, what their magnitudes are, and how the uncertainties should be interpreted to better understand the fitness-for-purpose of the ESA CCI+ ECV Precursor data.

These level-2 uncertainties are then input, together with known systematic aspects of the retrievals, to estimate the uncertainties in the level-3 products. The purpose is to explain how uncertainties are propagated when level-2 data are gridded in space and averaged in time to level-3 products, in the form of monthly mean values at a specific, regular geographic grid. The act of aggregating and averaging tends to reduce the random components of the level-2 uncertainties, whereas the systematic components of the retrieval uncertainties remain. For NO_2, two approaches are presented: one new method, developed for OMI and TROPOMI, wherein error correlations in time and space are quantified and accounted for in the spatio-temporal averaging of the L2 data, and one conventional method, that makes expert opinion assessments of the fractions of the uncertainty that are random and systematic.

1. Purpose and scope

1.1. Purpose

This document describes the uncertainties associated with the level-2 and level-3 (L3) processors in the Precursors_cci+ project. It addresses uncertainty estimates in the level-2 products (at pixel level) and in the level-3 (monthly mean, gridded) data for the tropospheric NO_2, HCHO, CO, SO_2, NH_3 and glyoxal retrievals.

1.2. Scope

The scope of version 2 of the E3UB is to describe and quantify the uncertainty budget for L2 and L3 data of ECV Precursors. This includes the general uncertainty propagation approach, the most important algorithm sensitivities, uncertainty estimates on ancillary data, and how these drive the overall pixel based L2 uncertainty budget. Examples will be given on spatio-temporal patterns in L2 uncertainties. It then proceeds with the general uncertainty propagation approach and the most important considerations in generating L3 data, such as the role of flagging, representativity, and minimum required coverage. Examples will be given on spatio-temporal patterns in L3 uncertainties.

We focus here on the following questions relevant to L2 and L3 users of ESA CCI+ ECM Precursor retrievals:

- What is the origin of the uncertainties in the L2 retrieval algorithms?
- Which uncertainty contributions are random, and which are systematic, and how do they propagate in the L2 retrieval algorithms?
- What is the budget of the uncertainties in the final ECV Precursor L2 data?
- What is the distribution of uncertainty patters in space and time?
- What does this mean for user applications of the ECV Precursor data?

- What is the budget of the uncertainties in the final ECV Precursor L3 data?
- How are L2 errors correlated in space and time, and what fraction of those errors can be expected to vanish when averaging over space (a grid cell) or period (e.g. a month)?
- What fraction of the uncertainties in the L3 data is due to uncertainties in the L2 data, and what fraction comes from incomplete sampling (due to e.g. clouds)?
- What is the distribution of L3 uncertainty patterns in space and time?
- What does this mean for user applications of the ECV Precursor data?

1.3. Applicable documents

[AD-1] Data Standards Requirements for CCI Data Producers. Latest version at time of writing is v1.2: ref. CCI-PRGM-EOPS-TN-13-0009, 9 March 2015, available online at: https://climate.esa.int/media/documents/CCI_DataStandards_v2-3.pdf

[AD-2] CCI Data Policy v1.1. Available online at: https://climate.esa.int/sites/default/files/CCI_Data_Policy_v1.1.pdf

1.4. Reference documents

[RD-1] GCOS Climate Monitoring Principles, November 1999. Available online at: https://gcos.wmo.int/en/essential-climate-variables/about/gcos-monitoring-principles

[RD-2] Guideline for the Generation of Satellite-based Datasets and Products meeting GCOS Requirements, GCOS Secretariat, GCOS-128, March 2009 (WMO/TD No. 1488). Available online at: https://library.wmo.int/index.php?lvl=notice_display&id=12884#.Yw4rL7RByUk

[RD-3] Quality assurance framework for earth observation (QA4EO): http://qa4eo.org

[RD-4] The Global Observing System for Climate: Implementation Needs, GCOS-200, October 2016. Available online at: https://gcos.wmo.int/en/gcos-implementation-plan

[RD-5] Status of the Global Observing System for Climate, GCOS-195, October 2015. Available online at: https://library.wmo.int/index.php?lvl=notice_display&id=18962#.Yw4r8LRByUk

[RD-6] Hollmann, R., et al., The ESA climate change initiative: Satellite data records for essential climate variables. American Meteorological Society. Bulletin, Vol. 94, No. 10, 2013, p. 1541-1552.

[RD-7] Joint Committee for Guides in Metrology, 2008, Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM), JGCM 100: 2008. Available online at: https://www.iso.org/sites/JCGM/GUM-JCGM100.htm

[RD-8] Merchant, C., et al., 2017, Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data Discuss., vol. 9, p511-527

1.5. List of acronyms

AC-SAF	Satellite Application Facility on Atmospheric Composition Monitoring
ADP	Algorithm Development Plan
AK	Averaging Kernel
AMF	Air-mass factor
ATBD	Algorithm Theoretical Basis Document
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
BIRA-IR	BIRA-IASB Infrared Team
BIRA-SYN	BIRA-IASB Synergy Team
BIRA-UVVIS	BIRA-IASB UV-Vis Team
BIRA-MOD	BIRA-IASB Tropospheric Modeling Team
CAMS	Copernicus Atmospheric Monitoring Service
C3S	Copernicus Climate Change Monitoring Service
CCI	ESA Climate Change Initiative
CCI+	Climate Change Initiative Extension (CCI+), is an extension of the CCI over the period 2017-2024.
CEOS	Committee on Earth Observation Satellites
CMUG	Climate Modeling User Group
CO	Carbon monoxide
COBRA	COvariance-Based Retrieval Algorithm
CRDP	Climate Research Data Package
CRG	Climate Research Group

DLR	German Aerospace Centre
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-range Weather Forecast
ECV	Essential Climate Variable
ENVISAT	Environmental Satellite (ESA)
EO	Earth Observation
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
FRM	Fiducial Reference Measurement
GCOS	Global Climate Observation System
GOME	Global Ozone Monitoring Instrument (aboard ERS-2)
GOME-2	Global Ozone Monitoring Instrument – 2 (aboard MetOp-A, -B and -C)
IASI	Infrared Atmospheric Sounding Interferometer
HRI	Hyperspectral Range Index
KNMI	Royal Netherlands Meteorological Institute
LEO	Low Earth Orbit
LUT	Look-up table
Metop	Meteorological Operational Platform (EUMETSAT)
MOPITT	Measurement of Pollution in the Troposphere
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detection of Atmospheric Composition Change
NH_3	Ammonia
NN	Neural Network
NO_2	Nitrogen dioxide
NRT	Near-Real Time
OCRA	Optical Cloud Recognition Algorithm)
OMI	Ozone Monitoring Instrument (aboard EOS-Aura)
PCA	Principal Component Analysis
QA4ECV	Quality Assurance for Essential Climate Variables
QA4EO	Quality Assurance framework four Earth Observation
R&D	Research and Development
ROCINN	Retrieval of Cloud Information using Neural Networks
SAF	Satellite Application Facility
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric
S5P	Sentinel-5 Precursor
SoW	Statement of Work
STREAM	STRatospheric Estimation Algorithm from Mainz
SZA	Solar Zenith Angle
TEMIS	Tropospheric Emission Monitoring Internet Service
TIR	Thermal Infrared spectral range
TROPOMI	Tropospheric Monitoring Instrument (aboard Sentinel-5 Precursor)
TOA	Top-of-atmosphere
TOAR-II	Tropospheric Ozone Assessment Report Phase-II
ULB	Université Libre de Bruxelles
IUP-UB	Institute of Environmental Physics, University of Bremen
UPAS	Universal Processor for UV/Vis Atmospheric Sensors

UV-Vis	Ultraviolet and visible spectral range
WP	Work Package

2. Terminology

The uncertainty characterization in the ECV Precursors CCI follows, in general, the approach and recommendations from Merchant et al. (2017), applied widely in the CCI programme, and used previously in the QA4ECV project. The terminology follows international standards from metrology. We note the following definition of terms:

- *the measurand*: the quantity to be measured (here often the tropospheric vertical column density);
- *measurement*: the process of experimentally obtaining measured values that can be attributed to a quantity. In this project the *retrieval process*, or *retrieval* is sometimes understood to be the measurement;
- the *measured value*: the result of a measurement or a retrieval obtained to quantify the measurand;
- the *error*: the measured value minus the true value of the measurand. In practice the error is unknowable, except when the measured or retrieved value can be compared with a reference value of negligible uncertainty;
- the *uncertainty*: a non-negative parameter characterizing the dispersion (spread) of the quantity value attributed to the emasurand, given the measured value and an understanding of the measurement. The uncertainty estimates are central in this document.

For the trace gas products discussed here, the spatial resolution for the L3 gridding has been selected such, that multiple L2 pixels fall in one cell on one day, irrespective of whether these pixels are valid or not. In practice this means that the following grid resolutions are considered practical for TROPOMI, IASI, OMI, GOME-2, SCIAMACHY, and GOME:

	0.2°x0.2°	0.5°x0.5°	1.0°x1.0°	2.0°x2.5°
TROPOMI	✓	✓	✓	✓
IASI	✓	✓	✓	✓
OMI		✓	✓	✓
GOME-2			✓	✓
GOME				✓

3. Level-2 algorithms

3.1 Tropospheric nitrogen dioxide (L2-NO2)

3.1.1. Uncertainty propagation

The retrieval procedure for tropospheric NO_2 consists of three distinct steps, and each contributes to the overall uncertainty. We therefore evaluate and propagate the uncertainties associated with each step: (1) the uncertainty in the total slant column originating from the spectral fitting, (2) the uncertainty in the background (stratospheric) correction, and (3) the uncertainty in the tropospheric air mass factor estimates, and (4) the overall algorithm uncertainty assessment resulting from the uncertainty propagation implemented for each individual pixel following the principles laid out in Boersma et al. [2004] and Boersma et al. [2018]. Specifically, the tropospheric NO_2 vertical column density is retrieved with the following equation:

$$N_{v,trop} = \frac{N_s - N_{s,strat}}{M_{trop}} \tag{3.1-1}$$

The main intermediate quantities of the retrieval algorithm are the slant column density (N_s), the tropospgeric air mass factor (M_{trop}), and the estimates for the stratospheric correction ($N_{s,strat}$). Through propagation of the uncertainties in all three intermediate quantities, we arrive at an uncertainty estimate for the tropospheric column at the algorithm level via Equation (3.1-2):

$$\sigma^2_{N_{v,trop}} = \left(\frac{\sigma_{N_s}}{M_{tr}}\right)^2 + \left(\frac{\sigma_{N_{s,strat}}}{M_{tr}}\right)^2 + \left(\frac{(N_s - N_{s,strat})\sigma_{M_{tr}}}{M^2_{tr}}\right)^2 \tag{3.1-2}$$

In Eq. (3.1.2), the (squared) air mass factor uncertainty is written as:

$$\sigma^2_{M_{tr}} = \left(\frac{\partial M_{tr}}{\partial A_s}\sigma_{A_s}\right)^2 + \left(\frac{\partial M_{tr}}{\partial f_{cl}}\sigma_{f_{cl}}\right)^2 + \left(\frac{\partial M_{tr}}{\partial p_{cl}}\sigma_{p_{cl}}\right)^2 + (\alpha M_{tr})^2 \tag{3.1-3}$$

where $\frac{\partial M_{tr}}{\partial A_s}$ represents the local sensitivity of the air mass factor to the surface albedo, σ_{A_s} the estimate of the uncertainty in the surface albedo, and so on. The last term on the right-hand side of Eq. (3.1.3) represents the contribution from uncertainty in the a priori NO_2 profile shapes and is estimated as ±10% (i.e. α=0.1) of the tropospheric air mass factor based on studies that replace a priori profile shapes with observed NO2 profiles or with simulated profiles from higher-resolution models in strongly polluted regions. Outside of those areas, in regions with no or modest pollution, the Round Robin pointed out that AMF uncertainty resulting from a priori profile shapes may be as high as 30% (i.e. α=0.3).

The contribution from uncertainties in the a priori profile is eliminated when NO_2 data users apply the averaging kernel (e.g. Rodgers (2000); Eskes and Boersma (2003); Boersma et al. (2016)). The last term on the right-hand side in Eq. (3.1-3) may then be omitted. Using averaging kernels, or recalculation of tropospheric air mass factors with the model simulations of interest (e.g. Visser et al. (2019)), allows for a better satellite-to-model comparison, by ensuring that the model is sampled consistent with the satellite retrievals, because identical assumptions are made on vertical sensitivity, and differences between the model's vertical NO2 distribution and the satellite a priori vertical NO_2 distribution then cancel.

3.1.2. NO_2 slant column density uncertainty estimates

Within the S5P and QA4ECV-projects, and in the Round Robin of this ECV Precursor project, much work has been done to characterize the magnitude and spatial patterns in NO2 slant column uncertainties from various sensors, especially for GOME-2A, OMI, and TROPOMI, and to better understand the sources of error that drive the uncertainty in the NO_2 slant columns. We summarize here the state-of-science in our understanding of the uncertainties in the NO_2 slant column densities (σ_{N_s}) for all sensors in this project.

Table 3.1.1 NO_2 slant column (statistical) uncertainties (typical, per pixel), and their trends, recently estimated in the QA4ECV, S5P, and ECV Precursor projects.

	Processor	**σ_{N_s} (random)**	**Trend**	**Reference**
GOME	QA4ECV (405-465 nm)	0.40×10^{15} molec. cm^{-2}		Boersma et al. (2004)
SCIAMACHY	QA4ECV (405-465 nm)	0.50×10^{15} molec. cm^{-2}		Boersma et al. (2008)
GOME-2A	QA4ECV AC-SAF[1]	0.73×10^{15} molec. cm^{-2} 0.47×10^{15} molec. cm^{-2}	3% yr^{-1} 5.2%yr^{-1}/1.8% yr^{-1} [2]	Zara et al. (2018) Seo et al. (2023)
GOME-2B	AC-SAF[a]	0.48×10^{15} molec. cm^{-2}	1.2% yr^{-1}	Seo et al. (2023)
GOME-2C	AC-SAF[a]	0.60×10^{15} molec. cm^{-2}	4.5 % yr^{-1}	Seo et al. (2023)
OMI	QA4ECV Coll4	0.76×10^{15} molec. cm^{-2}	1 % yr^{-1}	Zara et al. (2018); Anglou et al. (2023)
TROPOMI	Coll3	0.6×10^{15} molec. cm^{-2}	1.5 % yr^{-1}	van Geffen et al. (2020; 2022)

One lesson from earlier work is that the uncertainty in NO_2 slant column densities is spatially coherent and is mostly driven by random sources of error such as instrument noise. Uncertainty levels indeed depend on the reflectance level, with the darkest scenes having somewhat higher NO_2 uncertainties, and brighter scenes somewhat lower. The largest uncertainties typically occur close to areas with very high reflectance resulting from very bright (convective) clouds. On average, all-sky scenes (including cloudy areas) have some 10% lower NO2 uncertainties than clear-sky scenes (Zara et al., 2018), but we focus on the latter here, since these are the scenes relevant for tropospheric NO_2 retrievals.

To assess the NO_2 slant column uncertainties from three GOME-2 sensors on MetOp-A/B/C throughout their mission periods, we used a posteriori statistical approach that quantifies the spatial variability of the slant columns over clean Pacific regions used in previous studies (Boersma et al., 2007; Zara et al., 2018). This statistical method is useful to quantify the NO_2 slant column uncertainties derived by the instrument level-1 noise and spectral fit quality by limiting the contributions from other components such as anthropogenically pollution sources. Figure 3.1.1 shows a temporal evolution of the statistical NO2 slant column uncertainties from the AC-SAF GOME-2A (2007-2021), GOME-2B (2013-2022), and GOME-2C (2019-2022). All three GOME-2 instruments show a positive trend due to the propagation of measurement noise over time. For GOME-2A, a relatively fast increase rate of NO_2 slant column

[1] The AC-SAF uncertainty estimates are reported here for 'all-sky' conditions, all other are for 'clear-sky' conditions, which are known to be substantially higher because of lower reflectances (Zara et al., 2018).
[2] Trends in GOME-2A NO2 slant column uncertainties before and after the second throughput test (September 2009). The first period is from January 2007 to September 2009, and the second is from October 2009 to December 2019.

uncertainties by 5.2 % yr^{-1} is found until the second throughput test in September 2009. The second throughput test, which initially resulted in an additional loss of signal-to-noise, has positive effects on the quality of the level-1 data by weakening the long-term drift of the instrument's slit function. Therefore, following the test, GOME-2A NO_2 slant column uncertainties increase at a slower pace of 1.8% yr^{-1} until 2019. From 2020 until its lifespan, the quality of GOME-2A NO_2 fits deteriorated rapidly due to a fast degradation of level-1 data, resulting in a sharp increase in NO_2 slant column uncertainties at a rate of 17.5% yr^{-1}. The AC-SAF GOME-2B shows a comparable but slightly lower increase rate in NO_2 slant column uncertainties (1.2% yr^{-1}) over the 10-year period from 2013 to 2022 when compared with GOME-2A, due to their identical spectral retrieval settings. In contrast, GOME-2C has a relatively higher trend of increase in NO_2 slant column uncertainties during its initial phase of operation compared to GOME-2A and GOME-2B. This trend may be attributed to a combination of factors, including detector degradation and differences in spectral fit settings for GOME-2C in the AC-SAF algorithm.

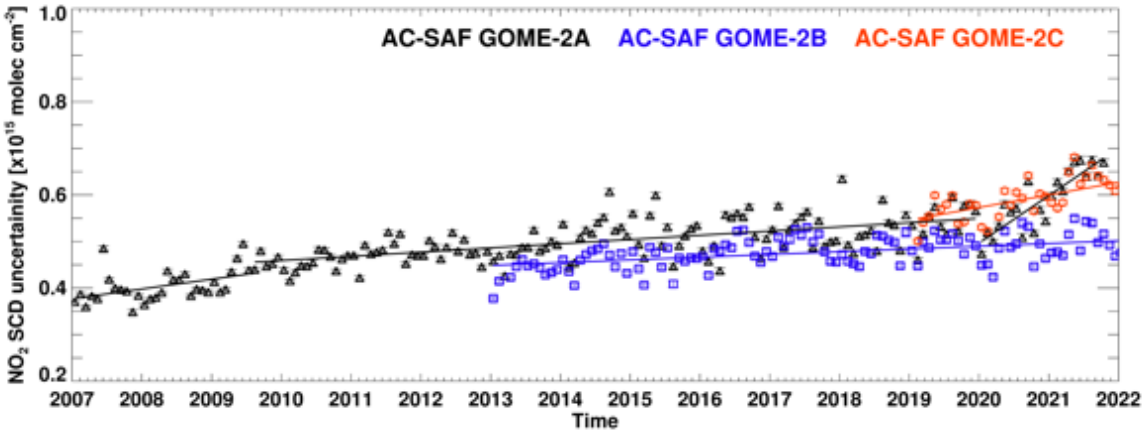


Figure 3.1.1. Temporal evolution of the statistical NO_2 slant column uncertainties from GOME-2 on MetOp-A/B/C retrieved from the AC-SAF algorithm. These uncertainties are calculated using the slant column retrievals over the clean Pacific region on day 15 of each month. Solid lines represent the linear regressions fitted for each instrument's periods. Error bars indicate one standard deviation (1σ).

Figure 3.1.2 below shows the evolution of the TROPOMI NO_2 slant column statistical uncertainty over time. The red line shows a 21-day running mean. Between 1 August 2019 and 1 July 2023 the TROPOMI NO_2 SCD statistical uncertainty has increased by 6%, corresponding to a trend of 1.5 % yr^{-1}. The result is an update of the material presented in van Geffen et al. (2022).

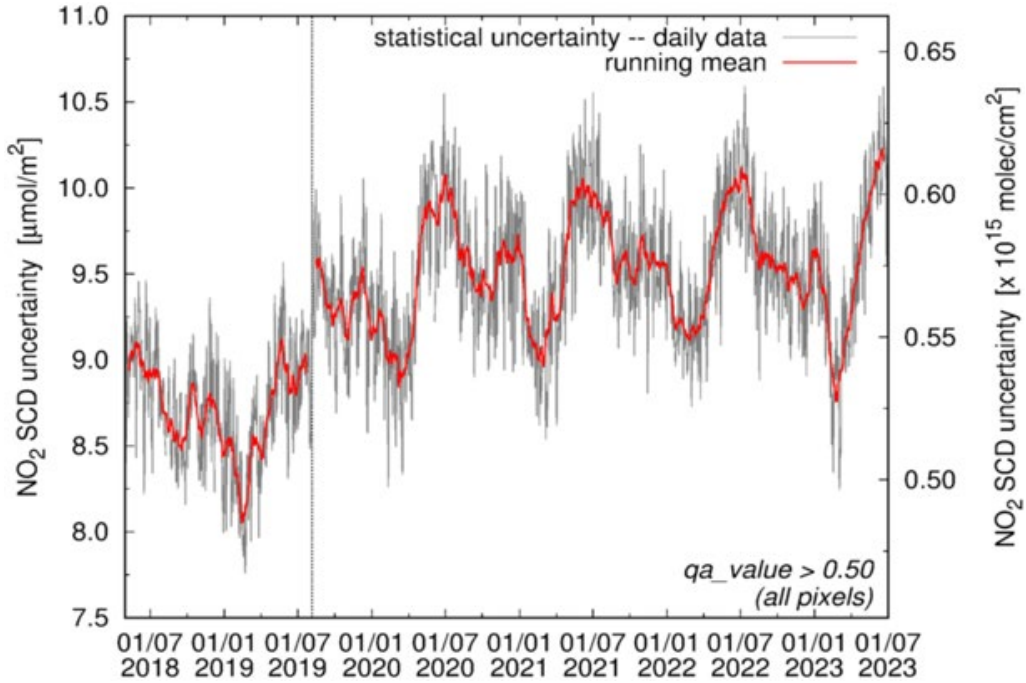


Figure 3.1.2 Statistical uncertainty in TROPOMI NO_2 slant columns as a function of time. The thin vertical grey line indicates the day that the along-track size of the TROPOMI pixel was reduced, which resulted in smaller pixels and therefore larger per-pixel uncertainties.

Nevertheless, there are also known systematic effects, such as the choice to include or omit a term describing the I_0-offset in the NO_2 spectral fitting. Earlier work by Oldeman (2018) has showed that including the I_0-correction leads to 1×10^{15} molec. cm^{-2} smaller TROPOMI NO_2 SCDs and up to 0-30% (average: 8%) lower NO_2 SCD uncertainties for clear-sky ocean scenes but has a near-zero effect over land and over cloudy scenes. Since some of the spectral fitting approaches include the I_0-correction (GOME, SCIAMACHY, GOME-2), but others do not, we should account for a systematic uncertainty of 1×10^{15} molec. cm^{-2}.

3.1.3. NO_2 stratospheric correction uncertainty estimate

Stratospheric NO_2 columns and associated tropospheric residuals can be estimated with different approaches. The KNMI data assimilation approach, applied in the QA4ECV (Boersma et al., 2018) and the operational OMI and TROPOMI retrieval algorithms, and the STREAM scheme, used by DLR for AC-SAF retrievals provide more insight into the NO2 stratospheric correction uncertainties.

The theoretical uncertainty estimates for the data assimilation (derived from observation minus forecast statistics of the assimilation scheme, Dirksen et al. (2018)) and for STREAM are both on the order of 0.2×10^{15} molec. cm^{-2} (Beirle et al., 2016). Comparisons in the Round Robin indicate that, on a day-to-day basis, the differences between the stratospheric NO_2 columns from data assimilation and STREAM were generally smaller than 0.5×10^{15} molec. cm^{-2} which can be considered an upper limit of the systematic uncertainty in the stratospheric NO_2 estimate. There is generally good agreement in spatiotemporal variability in stratospheric NO_2 between data assimilation and STREAM. STREAM, with the settings used by DLR, attributes generally a higher fraction of the NO_2 SCDs to stratospheric NO_2,

so that tropospheric NO_2 columns retrieved with data assimilation are higher by up to 1×10^{15} molec. cm^{-2}, and tropospheric NO_2 columns retrieved with STREAM can be occasionally negative.

Table 3.1.2 summarizes the random and systematic uncertainties which have been investigated specifically for GOME-2A (in the ESA CCI ECV Precursor project and Liu et al. (2019)), for OMI (in QA4ECV), and for TROPOMI (in the S5P project). Those studies gave very similar conclusions, so we hold the uncertainty estimates representative for all sensors and both the data assimilation and STREAM methods studied here.

Table 3.1.2 NO_2 stratospheric slant column (statistical) uncertainties (typical, per pixel).

	$\sigma_{N_{s,strat}}$ (random)	$\sigma_{N_{s,strat}}$ (systematic)
All sensors	0.2×10^{15} molec. cm^{-2}	$<0.5\times10^{15}$ molec. cm^{-2}

3.1.4. Tropospheric air mass factor uncertainty estimates

The theoretical uncertainties in the tropospheric AMF are calculated with Eq. (3.1.3), which, for every pixel, accounts for the sensitivity of the AMF to changes in the local forward model parameter value, evaluated around the specific value for the parameter at that pixel. The uncertainties in the surface albedo, surface pressure height, cloud fraction, cloud pressure are estimated from the literature, or estimated from intercomparisons with independent data. Typical values in the NO_2 AMF calculation are σ_{A_s}=0.015, $\sigma_{f_{cl}}$=0.025, $\sigma_{p_{cl}}$ =50 hPa. Because the retrieved cloud paramaters generally depend on (or are reasonably consistent with) the same assumed surface albedo as used for the AMF calculation, AMF uncertainties related to clouds and surface albedo will be dampened (Boersma et al. (2004); Boersma et al. (2018)).

AMF uncertainties due uncertainties in the cloud fraction and cloud pressure are 5%-10% in magnitude (Boersma et al., 2018), although they can be much larger when the cloud retrieval algorithm suffers from systematic errors such as in the early TROPOMI mission phase, when cloud pressures were biased low before FRESCO+ wide was implemented (Riess et al. (2022); van Geffen et al. (2022)). AMF uncertainties due to uncertainties in surface albedo are 10%. AMF uncertainties from NO_2 profile uncertainty are 10%-30%, based on the Round Robin comparison of tropospheric AMFs calculated with profiles from either TM5-MP or CAMS for strongly polluted regions (10%) and regions with enhanced background NO_2 (30%). Recent work from Douros et al. (2023) suggests that for TROPOMI, with its high resolution, the 10% profile uncertainty should probably be increased to 20% for polluted situations, reflecting the strong variability and heterogeneity in NO_2 profiles at the size of a TROPOMI pixel (3.5×5.5 km^2).

Table 3.1.3 provide a general estimate of the AMF uncertainties that are expected in the tropospheric NO_2 retrievals under polluted conditions. In such situations, the AMF uncertainty dominates the overall uncertainty budget. Aerosol-related uncertainties are not explicitly accounted for in the uncertainty analysis, because cloud retrievals are sensitive to scattering aerosols that prevail over many polluted regions, and effectively account for aerosol effects in the AMF calculation (Boersma et al., 2004; Boersma et al. 2011). It is known that for situations with strongly absorbing aerosols, the cloud retrievals do not capture the effect of aerosols, and substantial errors can occur (Castellanos et al. (2015); Liu et al. (2019;2020)). In a general sense, the effect of scattering aerosols on AMF calculations dominates over the effect of absorbing aerosols, whose impact is episodically significant over biomass burning regions, deserts, and some industrial hotspots (e.g. Sand et al., 2021).

Table 3.1.3 Contributions to the tropospheric AMF uncertainty due to various contributions. Especially the contribution from the a priori NO_2 profile uncertainty will be larger for the high-resolution sensors OMI and especially TROPOMI.

	Uncertainty in model parameter	Corrresponding contribution to $\sigma_{M_{tr}}/M_{tr}$
Surface albedo	σ_{A_s}=0.015	±10%
Cloud fraction	$\sigma_{f_{cl}}$=0.025	5-10%
Cloud pressure	$\sigma_{p_{cl}}$ =50 hPa	5-10%
A priori NO_2 profile shape		10-30% (sensor dependent)
Overall AMF uncertainty ($\sigma_{M_{tr}}$)		20-35% (sensor dependent)

3.1.5. Overall NO_2 retrieval uncertainties

The overall uncertainty in the tropospheric columns is driven by the propagation of uncertainties via Eq. (3.1.2). The overall uncertainty depends on the details of the individual retrieval and therefore differs from one pixel to the next. For small tropospheric NO_2 columns, such as over pristine or background areas, the overall uncertainty is dominated by the uncertainty from the spectral fitting process, whereas over polluted regions with high tropospheric NO_2 columns, the retrieval uncertainty is dominated by the AMF uncertainty.

Table 3.1.4 summarizes the assessment of the main contributions to the overall uncertainty budget for tropospheric NO_2 retrievals. These uncertainties should be interpreted as representative of typical single-pixel uncertainties for retrievals over polluted regions encountered by users. Expressions for systematic or structural uncertainty are under investigation and will be provided in cycle 2 of the project.

Table 3.1.4 Estimation of the uncertainty contributions to the overall, per-pixel tropospheric NO_2 column uncertainty over polluted regions.

Uncertainty contribution	**Random**	**Systematic**
From NO_2 slant column ($\frac{\sigma_{N_S}}{M_{tr}}$)	0.4-0.8×10^{15} molec. cm^{-2}	< 1.0×10^{15} molec. cm^{-2}
From stratosphere ($\frac{\sigma_{N_{S,strat}}}{M_{tr}}$)	0.2×10^{15} molec. cm^{-2}	< 0.5×10^{15} molec. cm^{-2}
From tropospheric AMF ($\frac{\sigma_{N_S}}{M_{tr}}$)	20-35% (sensor dependent)	TBD
Approximation for σ_{N_v}	$\sigma_{N_v} = 0.5 \times 10^{15} + 0.3N_v$	TBD

3.2. Formaldehyde (L2-HCHO)

3.2.1. Uncertainty propagation

The tropospheric HCHO vertical column density is retrieved with the following equation:

$$N_v = \frac{N_s - \overline{N_{s,0}}}{M} + \overline{N_{v,0}} \quad (3.2\text{-}1)$$

A formulation of the uncertainty can be derived analytically by error propagation, starting from the equation of the vertical column (Eq. (**3.2-1**) which directly results from the different retrieval steps. If we assume normal probability distributions, the total uncertainty on the tropospheric vertical column can be expressed as [Boersma et al., 2004, De Smedt et al., 2018]:

$$\sigma_{N,v}{}^2 = \left(\frac{\partial N_v}{\partial N_s}\sigma_{N,s}\right)^2 + \left(\frac{\partial N_v}{\partial M}\sigma_M\right)^2 + \left(\frac{\partial N_v}{\partial N_{v,0}}\sigma_{N,v,0}\right)^2 \quad (3.2\text{-}2)$$

$$\sigma_{N,v}{}^2 = \frac{1}{M^2}\left(\sigma_{N,s}{}^2 + \frac{\Delta N_s{}^2}{M^2}\sigma_M{}^2\right) + \sigma_{N,v,0}{}^2 \quad (3.2\text{-}3)$$

where $\sigma_{N,s}$, σ_M, and $\sigma_{N,v,0}$ are respectively the errors on the slant column, the air mass factor, and the background vertical column in the reference sector (indicated by suffix 0).

For each of these three categories, more details on the implementation of the uncertainty values in the L2 algorithm can be found in the TROPOMI HCHO ATDB (De Smedt et al., 2018). A summary of the sources of uncertainties and their estimated size are presented the next sections.

3.2.2. HCHO slant column density uncertainty estimates

The random uncertainty of the HCHO slant columns is estimated based on the slant column errors of the DOAS fits, and compared to the standard deviations of the slant columns in the remote Equatorial Pacific (as illustrated on Figure 3.2.1 and summarized in Table 3.2-1 for each sensor). To this random component, we consider an additional systematic error of 20% of the background-corrected slant columns.

Table 3.2-1 HCHO slant column random uncertainties (typical, per pixel) recently estimated in the QA4ECV, S5P, and ECV Precursor projects.

	$\sigma_{N,s,rand}$	Reference
GOME	2.7×10^{15} molec.cm^{-2}	Zara et al., 2018 De Smedt et al., 2015, 2018, 2021.
SCIAMACHY	$6\text{-}8\times10^{15}$ molec.cm^{-2}	
GOME-2	$7\text{-}13\times10^{15}$ molec.cm^{-2}	
OMI	8×10^{15} molec.cm^{-2}	
TROPOMI	$6\text{-}7\times10^{15}$ molec.cm^{-2}	

3.2.3. Tropospheric air mass factor uncertainty estimates

The uncertainties on the air mass factor depend on input parameter uncertainties and on the sensitivity of the air mass factor to each of them. This contribution is broken down into the squared sum [Boersma et al., 2004, De Smedt et al., 2018]:

$$\boldsymbol{\sigma_M}^2 = \left(\frac{\boldsymbol{\partial M}}{\boldsymbol{\partial A_s}} \cdot \boldsymbol{\sigma_{A,s}}\right)^2 + \left(\frac{\boldsymbol{\partial M}}{\boldsymbol{\partial f_c}} \cdot \boldsymbol{\sigma_{f,c}}\right)^2 + \left(\frac{\boldsymbol{\partial M}}{\boldsymbol{\partial p_{cloud}}} \cdot \boldsymbol{\sigma_{p,cloud}}\right)^2 + \left(\frac{\boldsymbol{\partial M}}{\boldsymbol{\partial s}} \cdot \boldsymbol{\sigma_s}\right)^2 + (\mathbf{0.15}\boldsymbol{M})^2 \quad (3.2\text{-}4)$$

The contribution of each parameter to the total air mass factor uncertainty depends on the observation conditions. The air mass factor sensitivities ($M' = \frac{\partial M}{\partial parameter}$), i.e. the air mass factor derivatives with respect to the different input parameters, can be derived for any particular condition of observation using the altitude-dependent AMF LUT, and using the model profile shapes. It should be stressed that the uncertainty on the profile shape is estimated using one parameter describing the shape of the profile: the profile height, i.e., the pressure below which resides 75% of the integrated HCHO profile. The errors $\sigma_{A,s}, \sigma_{f,c}, \sigma_{p,cloud}, \sigma_{s,h}$ are typical uncertainties on the surface albedo, cloud fraction, cloud pressure and profile shape, respectively. The fifth term on the right of Eq. (**3.2-4**) represents the structural uncertainty contribution due to possible errors in the AMF model itself (Lorente et al., 2017). We estimate this contribution to 15% of the air mass factor.

Estimates of the air mass factor errors and of their impact on the vertical column uncertainties are listed in Table 3.2-1. They are based on the application of equation (**3.2-4**) to HCHO columns retrieved from OMI and TROPOMI measurements.

Table 3.2-1 Summary of the different error sources considered in the air mass factor error budget.

Error source	Type	Parameter uncertainty	Estimated uncertainty on HCHO VCD
Surface albedo	systematic	0.02	10%
Cloud fraction	systematic	0.05	10%
Cloud height	systematic	50hPa	20%
Profile shape height	systematic	100hPa	30-60%
Structural uncertainty	systematic	15%	5-15%

3.2.4. Background correction uncertainty estimates

This error includes contributions from the model background vertical column, from the error on the air mass factor in the reference sector, and from the amplitude of the normalization applied to the HCHO columns. The uncertainty on the air mass factor in the reference sector is calculated as in Eq. (**3.2-4**) and saved during the background correction step. Uncertainty on the model background has been estimated as the monthly averaged differences between two different CTM simulations in the

reference sector: IMAGES (Stavrakou et al., 2009a), TM5 (Huijnen et al., 2010) and CAMS Rea (Innes et al., 2019).

Table 3.2-2 Estimated errors on the reference sector correction.

Error source	Type	Uncertainty on HCHO VCD	Evaluation method – reference
Model background	systematic	0.5-1.5x10^{15} molec.cm^{-2}	Difference between IMAGES and TM5 model
Amplitude of the column normalisation	systematic	0-4x10^{15} molec.cm^{-2}	Sensitivity tests using GOME-2 and OMI data.

3.2.5. Overall HCHO retrieval uncertainties

Here are presented respectively our current estimates of the precision (random error uncertainty) and the trueness (systematic error uncertainty) of the HCHO vertical columns. These estimates are mainly based on OMI and TROPOMI HCHO retrievals.

Random Uncertainty (precision)
The random uncertainty of the vertical columns is provided in the L2 products on an individual pixel basis and for each sensor. It is calculated as the uncertainty of the slant columns errors divided by the air mass factors (first term of equation 3.2-3). Indeed, in the case of HCHO, the uncertainty of the slant columns completely dominates the vertical column precision. The precision can be improved by averaging the data, for example when grids are created (see section 4.3). At a given spatio-temporal resolution, the precision will depend on the initial SCD precision, on the air mass factors, and on the number of observations included in the average (sensor dependent, see section 4.3.3.1 Random and systematic uncertainties).

Figure 3.2.1 shows the estimated precision of OMI and TROPOMI HCHO columns at different spatial and temporal scales, for continental emissions and in the remote reference sector (from De Smedt et al., 2021). The theoretical noise corresponds to single-measurement precision divided by the square root of observations. The horizontal blue line at 1.3x10^{15} molec.cm^{-2} represents the Copernicus user requirement. TROPOMI significantly improves the precision of the HCHO observations at short temporal scales and for low HCHO columns. Compared to OMI, the precision of the TROPOMI HCHO columns is improved by 25 % for individual pixels and by up to a factor of 3 when considering daily averages in 20 km radius circles.

Systematic Uncertainty (trueness)
Table 3.2-4 provides estimates of the systematic uncertainty on the HCHO tropospheric columns, for two different NMVOC emission regimes (low/high HCHO columns). This evaluation was made for the QA4ECV project, based on OMI data. More recent validation studies based on TROPOMI data agree with those estimates (Vigouroux et al., 2020, De Smedt et al., 2021). The largest contribution to the systematic uncertainty on the vertical columns comes from air mass factor uncertainty and is mainly related to the auxiliary data used in the AMF calculation. Therefore, we do not expect large differences between the different sensors. The largest contribution is from the a priori profile uncertainty. In the case where the satellite averaging kernels are used for comparisons with external HCHO columns, the a priori profile contribution can be removed from the relative comparison uncertainty budget, leading to a total uncertainty in the range of 25% to 50%.

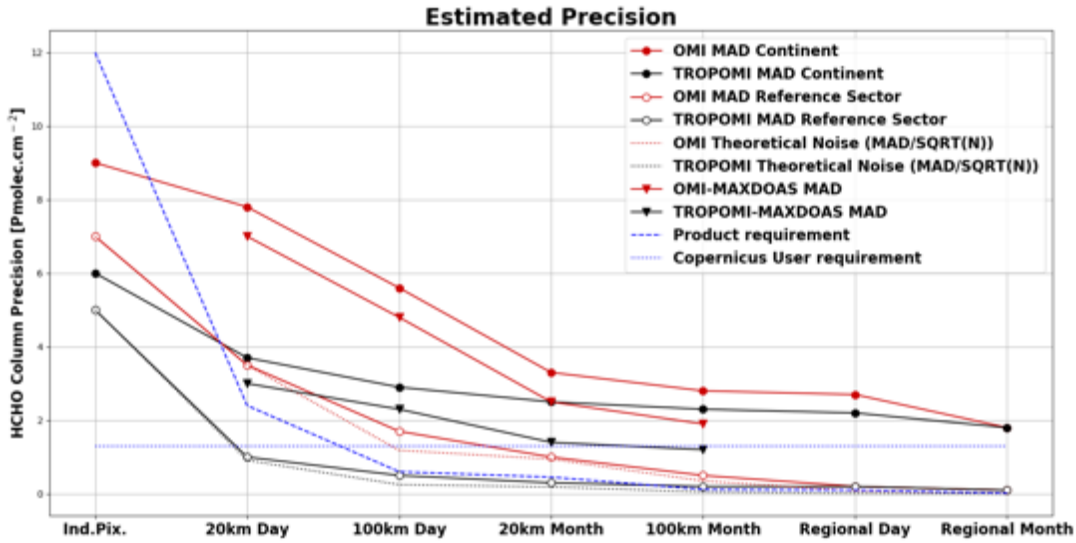


Figure 3.2.1: Estimated precision of OMI (in red) and TROPOMI (in black) HCHO columns at different spatial and temporal scales (20km, 100km, regions, day/month) (from De Smedt et al., 2021). The median deviation of the satellite HCHO columns are provided for continental emissions (plain circles) and in the remote reference sector (white circles). Validation estimates are plotted at 20km and 100km (MAD of differences between satellite and MAX-DOAS columns, triangles). The theoretical noise (dotted lines) corresponds to single measurement precision divided by the square root of observations. The dashed blue line is the TROPOMI product requirement, based on a single measurement precision of $12x10^{15}$ molec.cm^{-2}. The horizontal blue line at $1.3x10^{15}$ molec.cm^{-2} represents the COPERNICUS user requirement. [Pmolec.cm^{-2} = $1x10^{15}$ molec.cm^{-2}].

Table 3.2-4: Estimated HCHO vertical column uncertainty budget for monthly averaged low and elevated columns (higher than $1x10^{16}$ molec.cm^{-2}). Contributions from the three retrieval steps are provided, as well as input parameter contributions.

HCHO vertical error uncertainty	Remote regions / low columns	Elevated column regions / periods
Contribution from systematic slant columns uncertainties	25%	15%
Contribution from systematic air mass factors uncertainties	75%	30%
• from a priori profile errors • from model errors • from albedo errors • from cloud top pressure errors • from cloud fraction errors	• 60% • 35% • 20% • 20% • 15%	• 30% • 15% • 10% • 10% • 05%
Contribution from background correction uncertainties	40%	10%
Total	**90%**	**35%**
Total without smoothing error	**50%**	**25%**

3.3. Sulfur dioxide (L2-SO2)

3.3.1. Uncertainty propagation

The total uncertainty (accuracy and precision) on the SO_2 columns produced by the algorithm, it is composed of many sources of error (see e.g., Theys et al., 2017). Several of them are related to the instrument and propagate into the uncertainty on the slant column. Other types of error can be considered as forward model errors and are related to the representation of the physics in the algorithm and can affect the slant column results or the air mass factors.
Another important point is that one should differentiate systematic and random components of a given error source V:

$$\sigma_V^2 = \frac{\sigma_{V(rand)}^2}{n} + \sigma_{V(syst)}^2 \quad (3.3\text{-}1)$$

here *n* is the number of pixels considered. However, these are hard to separate in practice. Therefore, error contributions are (tentatively) classified as either "random" or "systematic" errors, depending on their tendencies to average out in space/time or not. These error estimates are labelled as "precison" and "trueness" in the TROPOMI data files.

As a matter of fact, the total retrieval uncertainty on the SO_2 vertical columns can be derived by propagation of errors from the two indepent retrieval steps:

$$\sigma_{N_v}^2 = (\frac{\sigma_{N_S}}{M})^2 + (N_v.\frac{\sigma_M}{M})^2 \quad (3.3\text{-}2)$$

where $\sigma_{VCD}, \sigma_{SCD}, \sigma_{AMF}$ are the errors on the vertical column, slant column and air mass factor, respectively. The AMF error can be broken down into a squared sum of terms (Boersma et al., 2004):

$$\sigma_M^2 = \left(\frac{\partial M}{\partial A_s} \cdot \sigma_{\text{alb}}\right)^2 + \left(\frac{\partial M}{\partial p_{cl}} \cdot \sigma_{\text{ctp}}\right)^2 + \left(\frac{\partial M}{\partial f_{cl}} \cdot \sigma_{feff}\right)^2 + \left(\frac{\partial M}{\partial s} \cdot \sigma_s\right)^2 \quad (3.3\text{-}3)$$

where σ_{As}, σ_{pcl}, σ_{fcl}, σ_s are typical uncertainties on the albedo, cloud top pressure, cloud fraction and profile shape, respectively.
The contribution of each parameter to the total air mass factor error depends on the observation conditions. The air mass factor sensitivities ($\frac{\partial M}{\partial parameter}$), i.e., the air mass factor derivatives with respect to the different input parameters, can be derived for any condition of observation using the altitude-dependent AMF LUT and using the a-priori profile shapes. It should be stressed that the uncertainty on the profile shape is estimated using one parameter describing the shape of the profile: the profile height, i.e., the pressure below which resides 75% of the integrated SO2 profile.

Note that this error analysis is complemented by the total column averaging kernel (AK) as described in Eskes and Boersma (2003):

$$AK(p) = \frac{m'(p)}{M} \quad (3.3\text{-}4)$$

(m' is the weighting function, Eq. 3.6 in the ATBD SO_2 part) which characterizes the sensitivity of the retrieved column to a change in the true profile. For certain applications, the AK is particularly useful to apply an apriori substitution using another SO_2 profile. In that case, the systematic uncertainty on

the AMF due to the uncertainty on the profile is zero, and it is recommended to use in the output files the "trueness_kernel" field, that leaves out the profile error from the AMF error calculation.

3.3.2. SO_2 slant column density uncertainty estimates

Error sources that contribute to the total uncertainty on the slant column originate both from instrument characteristics and uncertainties/limitations in the COBRA slant column fitting algorithm. A summary is given in Table 3.3.1, at low/high SZA (latitudes). For the random component of the slant column errors, the error on the slant columns provided by the COBRA fit is considered (referred as SCDE) as it is assumed to be dominated by and representative of the different random sources of error.

For the systematic errors on the slant column, the numbers provided in Table 3.3.1 have been determined based on sensitivity tests and based on the actual data. For instance, systematic SCD offsets are sometimes found over clean areas, and are accounted for in Table 3.3.1 in the form of an absolute term. The relative term accounts for the uncertainty on the SO_2 cross-section and other spectral interferences.

Table 3.3.1 SO_2 slant column uncertainties (typical, per pixel) recently estimated in the S5P and ECV Precursor projects. Note: 1DU=1 Dobson unit (1DU: 2.69 x 10^{16} molecules/cm²).

	σ_{NS} (SZA=30°/60°)	Reference
GOME	Random: 0.15 DU /0.3 DU Systematic: 0.05 DU + 10% /0.1 + 15%	-
SCIAMACHY	tbd	-
OMI	Random: 0.2 DU /0.4 DU Systematic: 0.05 DU + 10% /0.1 + 15%	-
TROPOMI	Random: 0.2 DU /0.4 DU Systematic: 0.05 DU + 10% /0.1 + 15%	Theys et al. (2017, 2021)

3.3.3. Tropospheric air mass factor uncertainty estimates

For SO_2, the main error source is related to the air mass factor, as it depends on many input parameters that are not always well constrained. Using Eq. 3.3.3, one can calculate a total AMF error based on typical uncertainties on the input parameters. Here we use fixed values set to σ_{As}=0.02, σ_{pcl}=50 hPa, σ_{fcl}=0.05, σ_s=75 hPa. The resulting uncertainty budget is based on the error analysis from the TROPOMI SO_2 retrieval algorithm (Theys et al., 2017), but is also consolidated with results from the CCI+precursors round Robin SO_2 exercise.

Generally speaking, it is difficult to ascertain whether a source of error is purely random or systematic. Several sources have the tendency to reduce by spatial or temporal averaging and are labelled as 'random'. Conversely, other error sources that do not average out are classified as 'systematic'.

Errors due to uncertainties on cloud fraction and cloud pressure are typically random errors, and account for 5% and 20% on the VCDs, respectively. Inversely, applying a cloud correction to the AMF or not, has a systematic effect of about 10% on the VCDs.

Errors due to uncertainties on the surface albedo is of ~15% (systematic), while the SO2 profile shape related error ranges from 20-50% and is considered as the largest systematic error source. However, for some applications, this error can be mitigated by using the column averaging kernels (see section 3.3.1).

Finally, additional systematic errors are due to the effect of atmospheric aerosols on the measurement sensitivity (15-25%) and other forward model errors (structural errors ~15%).

3.3.4. Overall SO_2 retrieval uncertainties

Table 3.3.2 summarizes the assessment of the main contributions to the global error budget on SO_2 retrieval from COBRA.

Table 3.3.2 Estimation of the error sources of the SO_2 COBRA retrieval (preliminary) for typical polluted conditions, at low/high SZA (latitudes). Blue and red fields indicate random (precision) and systematic (trueness) errors. Total errors are computed assuming all contributions are mutually uncorrelated.

Error source	**Error on VCD**	
	SZA=30°	**SZA=60°**
Instrument signal-to-noise	0.6 DU	1.2 DU
SO_2 absorption cross-section, spectral interferences +instrumental features	0.1 DU + 10%	0.2 DU + 15%
Surface albedo*	15%	15%
SO_2 Profile shape	20-50%	20-50%
Cloud correction	10%	10%
Cloud fraction	5%	5%
Cloud pressure	20%	20%
Aerosols (except in case of volcanic eruptions)	< 15%	< 25%
Structural (AMF)	15%	15%
Total random error	0.6 DU + 20%	1.2 DU + 20%
Total systematic error	0.1 DU + 45%	0.2 DU + 55 %
Total systematic error kernel	0.1 DU + 27%	0.2 DU + 35 %

*snow/ice free scenes

3.4. Glyoxal (L2-CHOCHO)

3.4.1. Uncertainty propagation

The total uncertainty on the retrieved glyoxal tropospheric columns is determined by many different factors, including instrument noise, radiative transfer modelling errors, errors in the absorption cross sections, or uncertainty in a priori information. We provide an overview of the different sources of uncertainty and estimate their contribution, as described in Lerot at al. (2021).

Following (Boersma et al., 2004; Lerot et al., 2010; De Smedt et al., 2008), the total error on the retrieved vertical column N_V can be expressed as:

$$\sigma_{N_V}^2 = \frac{1}{M^2}\left(\sigma_{N_s}{}^2 + N_V^2\,\sigma_M{}^2 + \sigma_{N_{s,0}}{}^2 + N_{v,0,ref}{}^2\sigma_{M,0}{}^2 + M_0{}^2\sigma^2{}_{N_{v,0,ref}}\right) \quad (2.4\text{-}1)$$

where σ_{N_s}, σ_M and $\sigma_{N_{v,0,ref}}$ are the errors on the slant column, the air mass factor and the reference value used in the background correction, respectively. $\sigma_{N_{s,0}}$ and $\sigma_{M,0}$ represent the errors on the slant column and air mass factor for pixels belonging to the reference Pacific sector used for the background correction. The three last terms together represent the error associated to the background correction.

We distinguish random errors, which tend to zero when averaging observations, from systematic errors, which are correlated in time and/or space. However, it is difficult to make this distinction in practice, as the various systematic error components have different correlation lengths and timescales (Boersma et al., 2004). For the glyoxal error budget, we consider the propagation of the instrumental shot noise in the slant column retrieval as the main random error. For weak absorbers such glyoxal, this error component dominates for individual measurements and depends on the observation geometry, the scene reflectivity, and on the instrument itself. The other error components are treated as systematic, though this assumption may lead to conservative systematic error estimates and to an underestimation of the product scatter, depending on the time and spatial resolution of interest. Specifically, uncertainties associated to the input parameters needed for the AMF calculation are directly related to the resolution of the used databases and may appear as random at coarser resolution. This has been discussed by Vigouroux et al. (2020) who attributed part of the scatter in formaldehyde vertical column TROPOMI/MAX-DOAS differences to a random component of the AMF errors.

The following sections discuss the different error components in detail.

3.4.2. Slant column errors

The retrieval noise for individual observations is limited by the instrument signal-to-noise ratio. The corresponding random errors on the slant column retrieval is estimated using the root mean square optical depth residual and the covariance matrix of the linear fit (Danckaert et al., 2017).

There are also systematic errors associated to the DOAS spectral fit, that are mainly dominated by slit function and wavelength calibration uncertainties, absorption cross-section uncertainties, by interferences with other species (O_4, liquid water, Ring, …), or by stray light corrections. Those contributions to the slant column errors are difficult to assess and can only be approximated from sensitivity tests (Lerot et al., 2010). In general, this term can be as high as 2–3×10^{14} molec.cm^{-2}.

However, the use of a radiance as the DOAS fit reference and the application of the background correction removes a large part of the systematic error in the slant column fit. Therefore, we pragmatically estimate this remaining systematic slant column error $\sigma_{N_{s,syst}}$ to 1×10^{14} molec.cm^{-2}.

3.4.3. Air mass factor errors

The errors on the air mass factor depend on input parameter uncertainties and on the sensitivity of the air mass factor to each of them. This contribution is broken down into the squared sum (Boersma et al., 2004; Lerot et al., 2010; De Smedt et al., 2008):

$$\sigma_{M,syst}^2 = \left(\frac{\partial M}{\partial A_s} \cdot \sigma_{A_s}\right)^2 + \left(\frac{\partial M}{\partial s} \cdot \sigma_s\right)^2 + (0.15M)^2 \qquad (3.4\text{-}2)$$

where σ_{A_s}, σ_s are typical uncertainties on the surface albedo and profile shape, respectively. They are estimated from the literature or derived from comparisons with independent data.
The contribution of each parameter to the total air mass factor error depends on the observation conditions. The air mass factor sensitivities ($M' = \frac{\partial M}{\partial parameter}$), i.e. the air mass factor derivatives with respect to the different input parameters, can be derived for any particular condition of observation using the altitude-dependent AMF LUT, created with VLIDORT, and using the model profile shapes. In practice, a LUT of AMF sensitivities has been created using reduced grids from the AMF LUT and a parameterisation of the profile shapes.

A reasonable uncertainty on the albedo is $\sigma_{A_s} = 0.02$ (Kleipool et al., 2008). Using this value and the AMF sensitivity LUT allows to estimate for every pixel the first term of Eq. (3.4.2) and gives the AMF error component related to the surface reflectivity. Of course, in practice, more substantial errors can be introduced if the real albedo differs considerably from what is expected, for example in the case of the sudden snowfall or ice cover.

The uncertainty associated to the a priori profile shapes (the smoothing error) used in the retrieval are more difficult to assess, especially due to the scarcity of independent glyoxal profile measurements. Those errors are nevertheless large owing to the large uncertainty in our current knowledge of the glyoxal production mechanisms. To assess this error component and the sensitivity of the air mass factor to the profile shape, we have introduced one profile effective pressure s_p, i.e. the pressure below which resides 50% of the integrated glyoxal profile. $\frac{\partial M}{\partial s}$ is approximated by $\frac{\partial M}{\partial s_p}$ and σ_s is taken equal to 50hPa. The latter value corresponds to the typical standard deviation of the model profile shape effective heights over polluted regions. Note that the error related to the a priori profiles can be removed with the use of the averaging kernels when comparing the satellite glyoxal columns to independent glyoxal profile measurements from models or MAX-DOAS. That is the reason why an estimate of the total systematic error without this smoothing error is also provided in the L2 product:

$$\sigma_{M,syst,kernel}^2 = \left(\frac{\partial M}{\partial A_s} \cdot \sigma_{A_s}\right)^2 + (0.15M)^2 \qquad (3.4\text{-}3)$$

The third term in Eq. (3.4.2) accounts for possible error in the AMF model itself (e.g. wavelength dependence, neglect of aerosols,…) and is estimated to be 15% of the air mass factor (Lorente et al., 2017).

As explained in the CCI+ precursors ATBD, no cloud correction is applied to compute the AMFs. Therefore, it is highly recommended to apply a stringent filter based on the cloud fraction (<20%) on the retrieved glyoxal vertical columns. Nevertheless, even with such a filter, an error on the AMF caused by residual cloud contamination may remain. This error, which strongly depends on the cloud coverage and altitude, is expected to average out in space and/or time when combining several observations. To consider this, we add a small random contribution to the air mass factor as:

$$\sigma_{M,rand}^2 = (0.05M)^2 \qquad (3.4\text{-}4)$$

The designed spatial and temporal threshold requirements for CHOCHO are respectively 100km and 1 month. See table 3.4.1 for the typical CHOCHO VCD precision estimates.

3.4.4. Background correction error

Although the background correction is designed to overcome systematic features/deficiencies of the slant column fitting, some errors are also associated to this procedure. In particular, systematic errors on the reference slant columns and their air mass factors are propagated to the computed correction values. Also, there is an uncertainty related to the used reference value of the glyoxal vertical column in the reference sector $N_{v,0,ref}$. Considering these different components, the background correction error can be estimated as:

$$\sigma_{bck}^2 = \frac{1}{M^2}\left({\sigma_{N_{s,0}}}^2 + {N_{v,0,ref}}^2{\sigma_{M,0}}^2 + {M_0}^2{\sigma_{N_{v,0,ref}}}^2\right) \qquad (3.4\text{-}5)$$

where $\sigma_{N_{s,0}}$ is the systematic slant column error fixed to 1x10^{14} molec/cm² (see section 3.4.2), M_0 and $\sigma_{M,0}$ are the air mass factors and associated systematic errors in the background correction reference sector. $\sigma_{N_{v,0,ref}}$ represents the error associated to the reference vertical column density $N_{v,0,ref}$ and is fixed to 5x10^{13} molec/cm².

3.4.5. Global budget: Total precision and trueness

Combining the different error components discussed above, the total random error (or precision) is computed as

$$\sigma_{Nv,rand}^2 = \frac{\sigma_{Ns,rand}^2 + \sigma_{M,rand}^2 N_v^2}{M^2} \qquad (3.4\text{-}6)$$

Table 3.4.1 provides typical values for each instrument.

The instrumental signal-to-noise ratio is the main factor introducing random errors in the product. The TROPOMI instrument has a signal-to-noise ratio around 1600 in the visible spectral range for low albedo scenes. Figure 3.4.1 illustrates the glyoxal VCD random error estimates for one day of TROPOMI observations (1st June 2018). This random error is generally less than 0.8-1x10^{15} molec/cm² over dark

surfaces and much lower for bright surfaces (clouds and snow-covered scenes). Previous studies (Alvarado et al., 2014; Chan Miller et al., 2014; Lerot et al., 2010) have shown that this random error is slightly larger for the older instruments GOME-2 and OMI (in the range 1-1.5 x 10^{15} molec/cm²). For individual satellite ground pixels, the random error is the dominant source of error on the total vertical column. This clearly points to the importance of the spatial/temporal averages, which reduce the random error by the square root of the number of observations and facilitate the detection of meaningful glyoxal signal.

Table 3.4.1 Typical CHOCHO VCD precision estimates (Lerot et al., 2021).

Instrument	VCD precision [molec. cm^{-2}]
TROPOMI	6 x 10^{14} – 1 x 10^{15}
OMI	7 x 10^{14} – 1.2 x 10^{15}
GOME-2A	5 x 10^{14} – 7 x 10^{14}
GOME-2B	7 x 10^{14} – 1 x 10^{15}

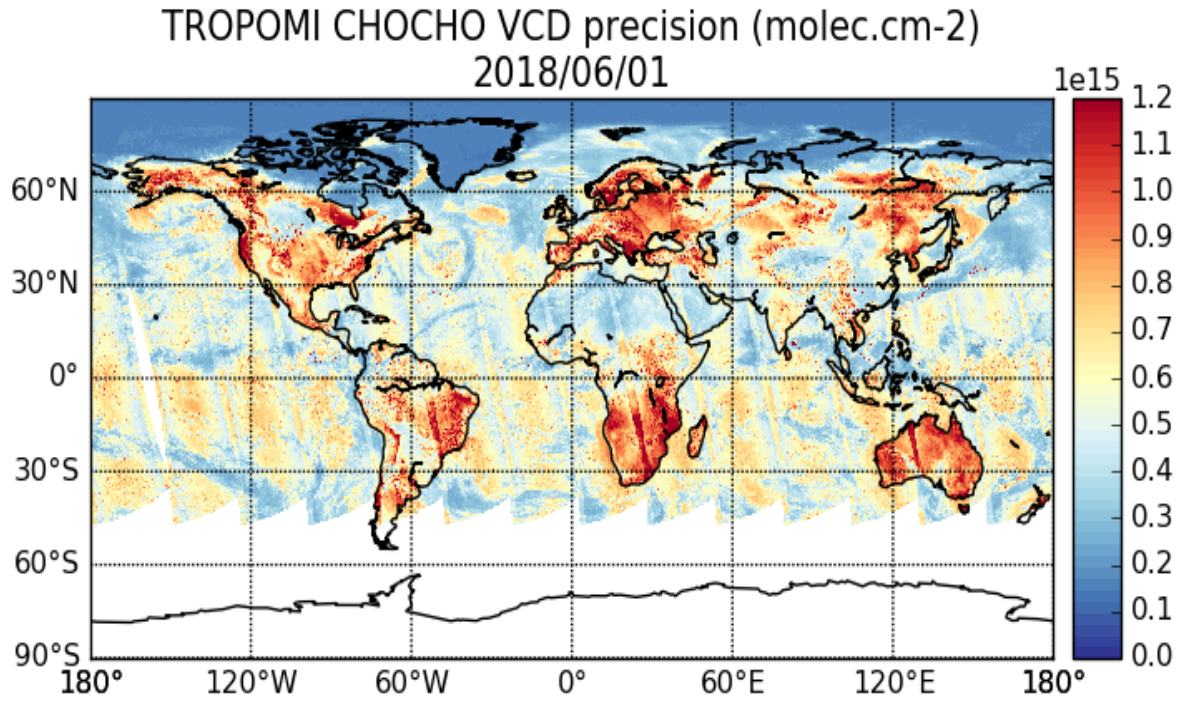


Figure 3.4.1: TROPOMI glyoxal tropospheric column random errors for 1 day, keeping all observations to illustrate the impact of scene brightness.

The total systematic error (trueness) is computed as

$$\sigma^2_{Nv,syst} = \sigma^2_{bck} + \frac{\sigma^2_{Ns,syst}}{M^2} + \frac{\sigma^2_{M,syst} \times N_v^2}{M^2} \quad (3.4\text{-}7)$$

Similarly, the total systematic error without the smoothing error component is calculated as

$$\sigma^2_{Nv,kernel,syst} = \sigma^2_{bck} + \frac{\sigma^2_{Ns,syst}}{M^2} + \frac{\sigma^2_{M,kernel,syst} \times N_v^2}{M^2} \quad (3.4\text{-}8)$$

Note that the conversion of the AMF error into an absolute vertical column error (3rd terms of equations (3.4-7) and (3.4-8) requires this error to be multiplied by the corresponding vertical column. Because of the high level of noise in the product, using the retrieved column for this would lead to a strong overestimation of the systematic error. To circumvent this, we use instead pre-computed climatological glyoxal noise-free VCDs. Figure 3.4.2: shows the mean estimates of the TROPOMI glyoxal VCD trueness for the month of June 2018, which are generally in the range 1-3x10^{14} molec/cm² (20-60% for emission regimes). Note that pixels strongly contaminated by clouds (cloud fraction > 20%) or covered by snow/ice have been discarded. Systematic errors are expected to be larger for those pixels as spectral interferences are larger and the information content becomes limited (cloud shielding effect).

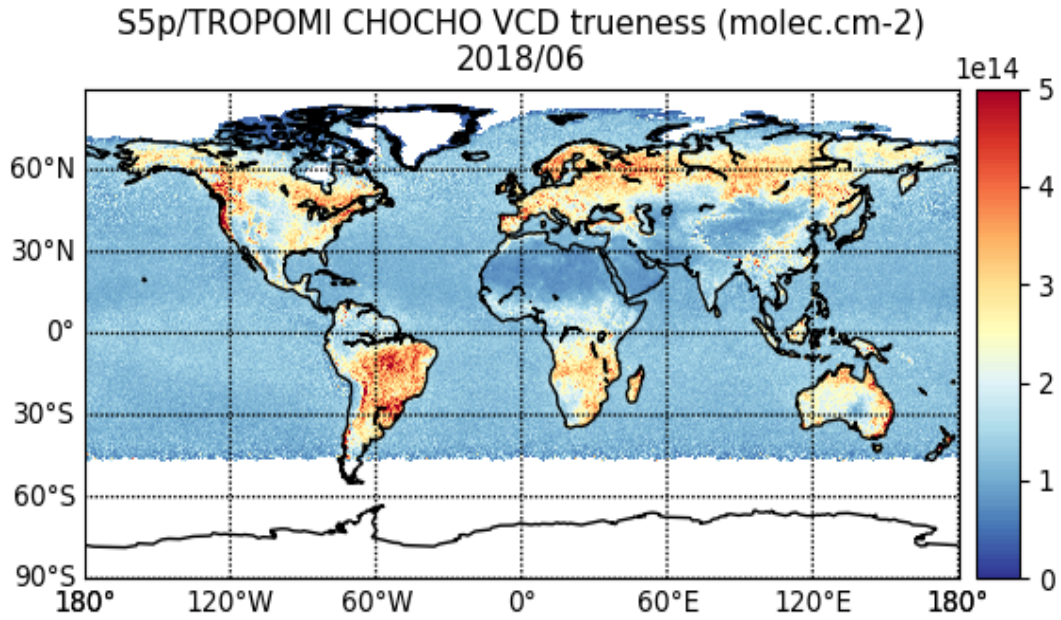


Figure 3.4.2: Mean TROPOMI glyoxal VCD trueness for June 2018.

Table 3.4.2 Glyoxal tropospheric VCD trueness.

Error source	Order of magnitude [molec. cm^{-2}]
Slant column retrieval systematic uncertainty	5 x 10^{13} – 1 x 10^{14}
Air mass factor uncertainty	1 x 10^{13} – 1 x 10^{14}
Background correction uncertainty	5 x 10^{13} – 1 x 10^{14}
Total systematic uncertainty	1 x 10^{14} – 3 x 10^{14}

3.5. Carbon monoxide (L2-CO)

IASI CO L2 data have been retrieved using the Fast Optimal/Operational Retrieval on Layers for IASI (FORLI), a dedicated radiative transfer and retrieval software for IASI. It relies on a scheme based on the Optimal Estimation theory (Rodgers, 2000). General formulation is given in FORLI Algorithm Theoretical Basis Document (ATBD, Astoreca et al., 2014) and in Hurtmans et al. (2012). The variance-covariance matrix representing the total statistical error after the retrieval is given in Equation 28 of Hurtmans et al. (2012):

$$\hat{\mathbf{S}} = (\mathbf{K}^T \mathbf{S}_\eta^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1}$$

From the different computed matrices, a detailed error budget can theoretically be performed, and in particular different contributions to the total error can be isolated :

- from the limited vertical sensitivity (the so-called smoothing error),
- from the measurement noise,
- and from uncertainties on fitted (here the water vapor column) or fixed (e.g. surface emissivity, temperature profile) parameters. In the routine processing of the error matrix, the error introduced by uncertainties on the fixed parameters is not taken into account.

In Rodgers (2000) formalism, the terminology “error” is used. When retrieving carbon monoxide concentrations from IASI measurements using FORLI, the user gets an “Total Error” with the CO total column, for every pixel. As previously said, this total error consists of the smoothing error and the measurement error, the smoothing error being dominant.

In this document, we performed an analysis of the IASI CO “total error” or uncertainty. First on global scale and then on local scale.

3.5.1. CO uncertainties on global scale

To apprehend CO uncertainties spatial distribution, we first present monthly total uncertainty global distributions, for four months of 2020 in Figure 3.5.1, for daytime and nighttime measurements. Uncertainties at high latitudes are larger than at mid-latitudes and at the Tropics, espacially in Boreal winter, when CO concentrations are the largest. Uncertainties are slighly larger during nighttime compared to daytime. When looking at land/sea differences, uncertainties appear to be larger over sea during the day, and larger over land during the night. For daytime measurements, uncertainties are ranging from 0.2 to 5.6 10^{17} molecules/cm^2. For nighttime, uncertainties are ranging from 0.53 to 5.3 10^{17} molecules/cm^2. IASI carbon monoxide total columns are retrieved with an uncertainty generally about 10–15% at mid and tropical latitudes, and 30-40% in the polar regions.

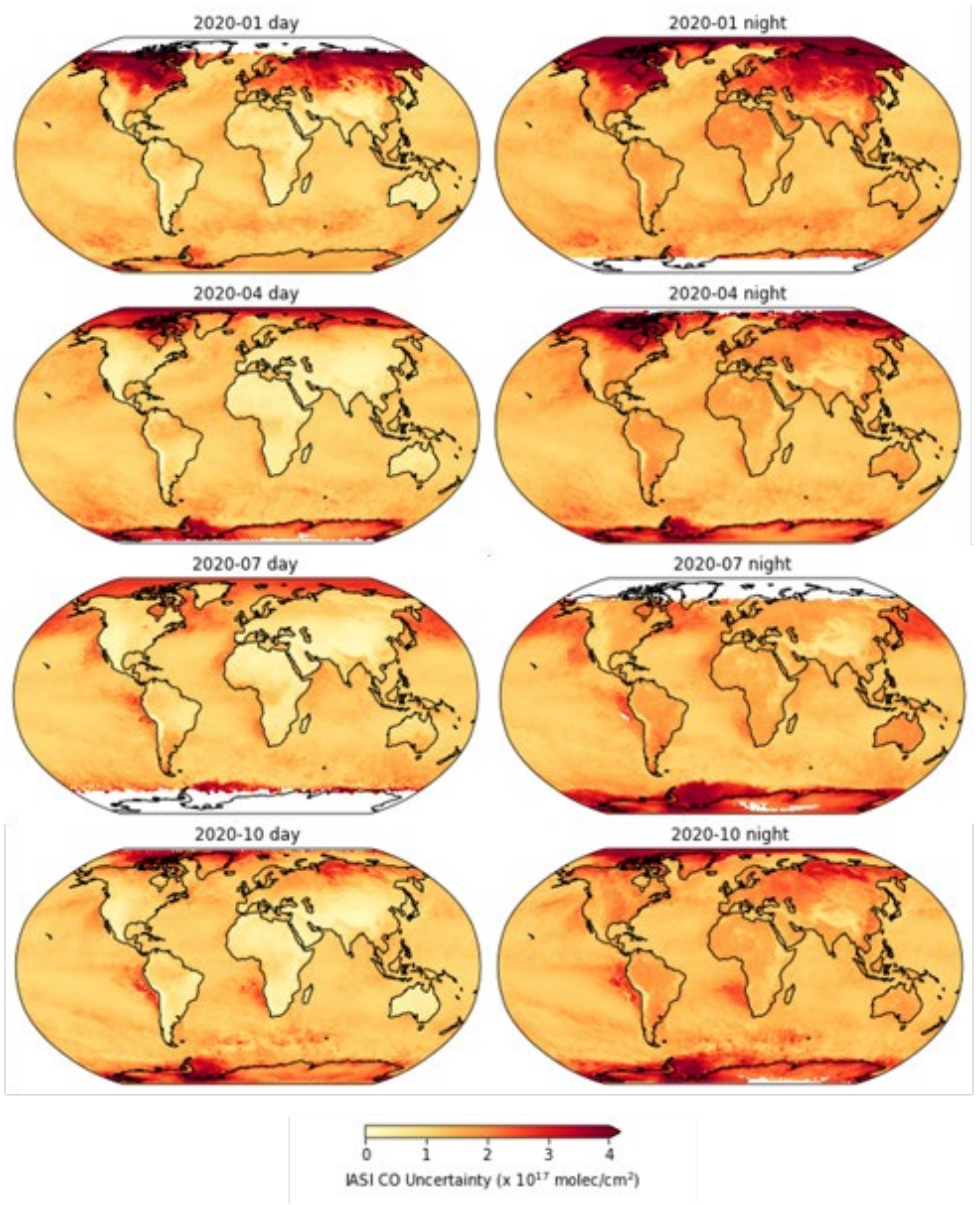


Figure 3.5.1 Monthly CO uncertainty distributions from January, April, July and October 2020, for day time measurements on the left and night time measurements on the right.

For statistics, we focused on different latitude bands and present CO uncertainty time series for Metop-A and B, from January 2008 til December 2021 (see Fig. 3.5.2). On global scale, mean uncertainty for daytime measurements is 1.4 10^{17} molecules/cm^2 with a standard deviation of 0.1 10^{17} molecules/cm^2 for both Metop-A and Metop-B. For nighttime, the mean uncertainty is larger: 1.7 10^{17} molecules/cm^2 [std: 0.1 10^{17} molecules/cm^2]. For the 20°N-60°N latitude band, mean uncertainty is 1.1 10^{17} molecules/cm^2 [0.1 10^{17}] for daytime, and 1.6 10^{17} molecules/cm^2 [0.1 10^{17}] for night time. For the tropical band [20°S-20°N] latitude, mean uncertainty is the smallest compared to the other latitude bands: 1.0 10^{17} molecules/cm^2 [0.02 10^{17}] for daytime, and 1.2 10^{17} molecules/cm^2 [0.02 10^{17}] for night time. Finally, for the [60°S-20°S] band, mean uncertainty is 1.3 10^{17} molecules/cm^2 [0.03 10^{17}] for daytime, and 1.4 10^{17} molecules/cm^2 [0.05 10^{17}] for night time.

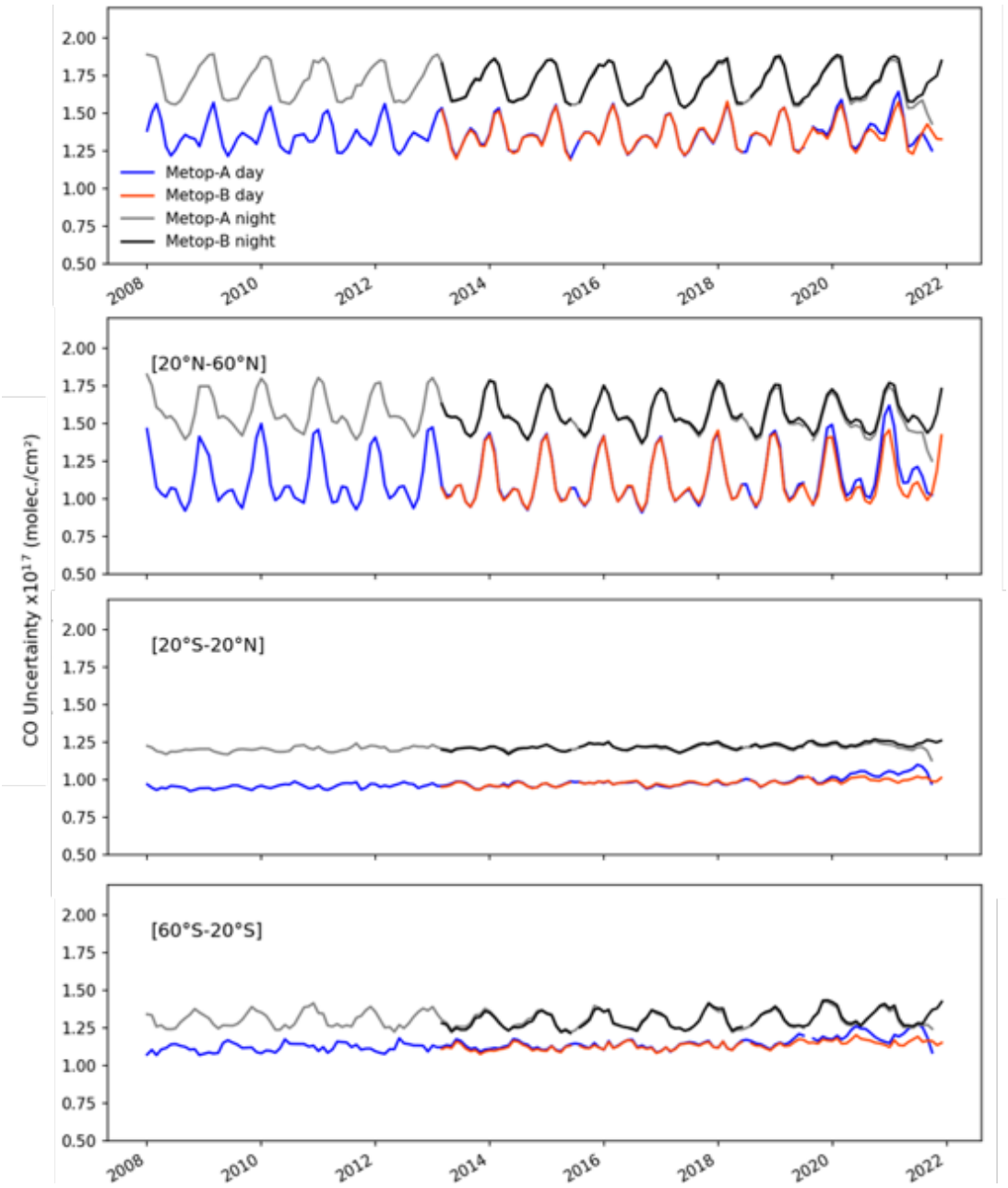


Figure 3.5.2: CO uncertainty time series for the globe (top) and for different latitudes bands: [20°N-60°N], [20°S-20°N] and [60°S-20°S]. For day time and night time measurements.

3.5.2. CO uncertainties on local scale

In order to get uncertainty estimates at local scale, we focused on fifteen locations on the globe, corresponding to TCCON stations (Fig. 3.5.3). Uncertainty time series are presented in Figure 3.5.4. CO uncertainties are following the CO total column concentrations seasonal cycles. Mean uncertainty for each station is given in Table 3.5.1. Over the fifteen stations, mean uncertainties are ranging from 0.63 10^{17} molecules/cm^2 [std 0.27 10^{17}] for Pasadena in the USA (3.2%) to 1.58 10^{17} molecules/cm^2 [std 0.78 10^{17}] (8.4%) in Sodankyla in Finland.

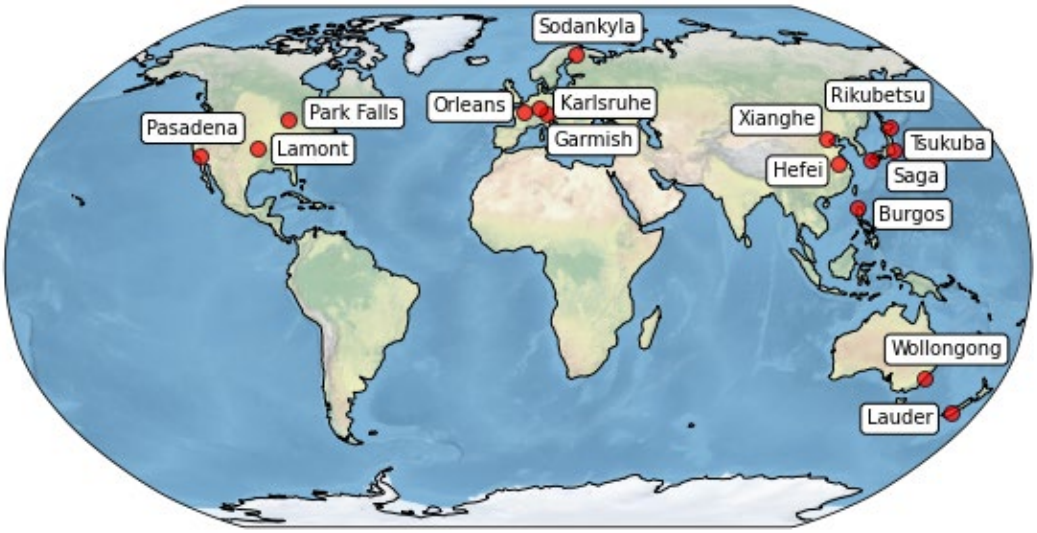


Figure 3.5.3: Location of the fifteen TCCON stations.

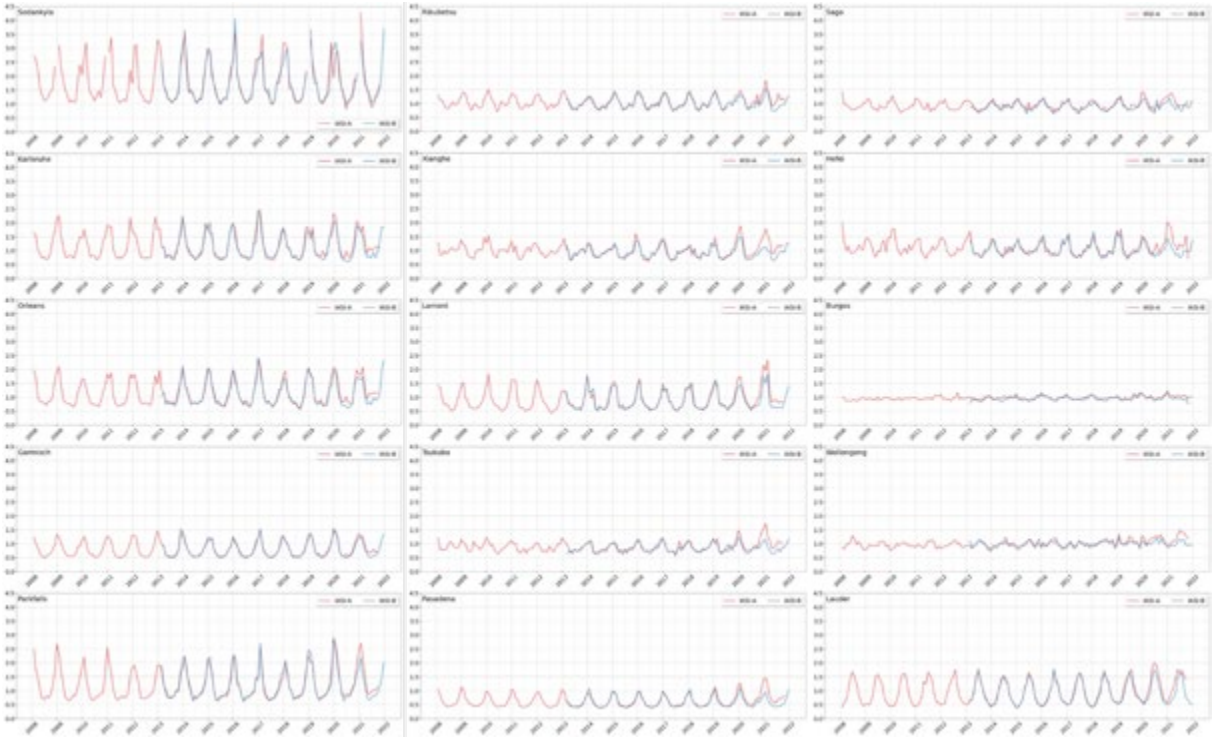


Figure 3.5.4: IASI CO uncertainty (x10^{17} molecules/cm^2) time series at the fifteen TCCON stations presented in Fig. 3.5.3, for Metop-A in red and Metop-B in blue.

Table 3.5.1: Mean uncertainties and standard deviation at the different TCCON stations.

Stations	Mean uncertainty ($x10^{17}$ molec./cm2)	Standard deviation ($x10^{17}$ molec./cm^2)	Mean uncertainty (%)	Standard deviation (percentage point)
Burgos	0.96	0.16	4.5	1.0
Garmisch	0.80	0.32	4.9	2.2
Hefei	1.10	0.36	3.8	1.7
Karlsruhe	1.12	0.56	5.5	3.2
Lamont	0.91	0.49	4.6	2.7
Lauder	0.97	0.50	8.9	4.5
Orleans	1.13	0.56	5.7	3.2
ParkFalls	1.13	0.62	6.2	4.2
Pasadena	0.63	0.27	3.2	1.4
Rikubetsu	1.08	0.30	5.3	2.0
Saga	0.94	0.25	3.9	1.2
Sodankyla	1.58	0.78	8.4	4.8
Tsukuba	0.91	0.29	3.8	1.5
Wollongong	1.01	0.27	7.8	2.1
Xianghe	1.01	0.41	3.7	2.4

3.6. Ammonia (L2-NH3)

The uncertainty assessment of NH_3 has been recently completely reviewed and overhauled (Clarisse et al., 2023). Most of the text below is taken verbatim from this reference.

3.6.1. Uncertainty propagation

In previous ANNI versions, an estimated uncertainty $\sigma_{\hat{X}}$ was calculated for each individual measurement $\hat{X}$ via (Ku, 1966)

$$\sigma_{\hat{X}}^2 = \sum_i \left(\frac{\partial \hat{X}}{\partial p_i}\right)^2 \sigma_{p_i}^2 \quad (3.6\text{-}1)$$

with σ_{p_i} the uncertainties of the different input parameters p_i. This formula assumes uncorrelated uncertainties, but as this cannot always be justified, in ANNI v4, we switch to the more general (Tellinghuisen, 2001)

$$\sigma_{\hat{X}}^2 = \vec{\boldsymbol{J}}^{\top} \boldsymbol{S_p} \vec{\boldsymbol{J}}, \quad (3.6\text{-}2)$$

With $\boldsymbol{S_p}$ the error covariance matrix of the input parameters (with covariances $S_{p,ij} = \sigma_{p_i p_j}$) and $\vec{\boldsymbol{J}}$ the Jacobian of the retrieval, with components $\frac{\partial \hat{X}}{\partial p_i}$.

In the ANNI retrieval framework, the input parameters include the skin temperature, the surface pressure, the HRI, the surface emissivity, the zenith angle, the width, and the peak of the Gaussian vertical NH_3 profile, the temperature profile (15 levels) and the water vapour profile (7 levels). After some preliminary analysis, it was concluded that only the correlations between the uncertainties in the temperature profile cannot be neglected. We therefore employ a block diagonal covariance matrix, block diagonal for the elements pertaining to the temperature profile, and diagonal for all other input parameters. As for uncertainty on the vertical profile, this source of uncertainty is removed when applying averaging kernels. For this reason, uncertainties are reported with and without the vertical profile uncertainty, to be used according to whether AVKs are applied.

3.6.2. Uncertainty on the input parameters

As most input parameters come without an uncertainty budget, let alone covariances, we made best-effort estimates of the co(variance) based on the limited information that is available. For now, the same (co)variances were used for the near-real time as for the reanalysed NH_3 product. It is also important to note that the systematic uncertainties of the input parameters vary according to the time and space scales that are considered (Boersma et al., 2004; Merchant et al., 2017; Sayer et al., 2020). Temperature profiles, for example, may be more biased monthly than annually. Here, we estimate systematic uncertainties with a typical L3 gridded data product in mind, i.e., for spatial scales of the order of one degree latitude/longitude or less, and for time periods of the order of one month or less.

The (co)variances, summarized in Table 3.6.1, were determined as follows:

- **HRI.** By definition, the random uncertainty on the HRI equals one. We estimate a systematic uncertainty of 0.1 due to potential and residual interferences (e.g., surface emissivity, VOCs). To account for potential biases in the spectroscopy and generalized error covariance matrix, we add to this an additional 10% on the calculated HRI value.

- **Skin temperature.** Random and systematic uncertainties were set to 1.5 and 0.5 K respectively. These values are in line with the difference between the IASI L2 skin temperature product and the dedicated neural network used for the reanalysis product of ANNI.
- **Emissivity.** For emissivity, which originates from the monthly climatology of (Zhou et al., 2013), an uncertainty of 0.01 and 0.005 was assumed for respectively the random and systematic components.
- **Temperature profile.** Variances were set based on validation results of the IASI level 2 (Eumetsat, 2021): systematic uncertainties of 1 K for the surface level and 0.5 K for the other levels; random uncertainties of 2 K for the surface level and 1 K for the other levels for land observations and 1 K for the surface level and 0.5 K for the other levels for ocean observations. Covariance matrices were then built by appropriate scaling of correlation matrices. These were built, based on a statistical analysis of the differences between collocated ERA5 and IASI L2 profiles. Correlation coefficients were set to 0.5 between neighbouring levels, and 0.25 between levels that are two levels apart. Above 10 km, no strong correlations were observed, and the covariance was therefore assumed to be diagonal for these levels.
- **Water vapour profiles.** Relying again on the IASI level 2 validation report (Eumetsat, 2021), random uncertainties were set to 10% below 3 km and 20% above. Systematic uncertainties were set to half these numbers.
- **Surface pressure.** A random and systematic uncertainty of 500 and 250 Pa was used.
- **NH_3 vertical profiles.** The uncertainties related to NH_3 profile are characterized by uncertainties on the width and the peak of the Gaussian shaped vertical profile. Random and systematic uncertainties of 200 and 100 m were used for both parameters. Given the short lifetime of NH_3 in the atmosphere, these are likely of the right order of magnitude. To obtain better estimates in the future, a thorough analysis using in-situ measurements or modelled profiles would be desirable.

Table 3.6.3. Estimated random and systematic uncertainties of the input parameters.

Component	Random σ_r	Systematic σ_s
HRI	1	0.1 + 10%
Surface temperature (K)	1.5	0.5
Emissivity	0.01	0.005
Temperature profile, land (K)	1 – 2	0.5 – 1
Temperature profile, sea (K)	0.5 – 1	0.5 – 1
Surface pressure (Pa)	500 Pa	250 Pa
Water vapour profile	10 – 20%	5 – 10%
NH_3 profile peak altitude (m)	200	100
NH_3 profile width (m)	200	100

3.6.3. Uncertainty characterization of L2-NH_3

It is useful, remembering the general form $\hat{X}^a = \text{HRI}/\text{SF}^a$ of the retrieval, to rewrite the propagation of uncertainty in terms of the uncertainty of the nominator and denominator (see also Boersma et al., 2004 and van Geffen et al., 2022). Neglecting the small dependence of the SF on the HRI, we obtain

$$\sigma_{\hat{X}}^2 = \left(\frac{\partial \hat{X}}{\partial \text{HRI}}\right)^2 \sigma_{\text{HRI}}^2 + \left(\frac{\partial \hat{X}}{\partial \text{SF}}\right)^2 \sigma_{\text{SF}}^2 \tag{3.6-3}$$

or

$$\sigma_{\hat{X}}^2 = \frac{\sigma_{\text{HRI}}^2}{\text{SF}^2} + \frac{\sigma_{\text{SF}}^2}{\text{SF}^2} X^2 \tag{3.6-4}$$

Taking into account both random and systematic uncertainties, we see from Table 1 that the uncertainty on the HRI has an absolute (constant) and a relative (proportional to the value of the HRI) component, so that

$$\sigma_{\hat{X}}^2 = \frac{\sigma_{\text{abs,HRI}}^2}{\text{SF}^2} + \frac{\sigma_{\text{rel,HRI}}^2}{\text{SF}^2} + \frac{\sigma_{\text{SF}}^2}{\text{SF}^2} X^2 \tag{3.6-5}$$

or

$$\sigma_{\hat{X}}^2 = \underbrace{\frac{1^2+0.1^2}{\text{SF}^2}}_{\sigma_{\text{abs}}^2} + \underbrace{\left(0.1^2 + \frac{\sigma_{\text{SF}}^2}{\text{SF}^2}\right) X^2}_{\sigma_{\text{rel}}^2} \tag{3.6-6}$$

3.6.3.1. Absolute uncertainty contribution

The first term is in the optically thin limit independent of the HRI and thus the column, and solely depends on the scene conditions:

$$\sigma_{\text{abs}} = \frac{\sigma_{\text{abs,HRI}}}{|\text{SF}|} = \frac{\sqrt{1+0.1^2}}{|\text{SF}|} \approx \frac{1}{|\text{SF}|}. \tag{3.6-7}$$

It is this term that is used as part of the post-filter to determine whether there is enough intrinsic sensitivity (thermal contrast) to make a valid measurement, i.e., one whose uncertainty is not completely overwhelmed by the instrumental noise. Currently, the post-filter threshold is set to $\sigma_{\text{abs}} < 1.5 \cdot 10^{16}$ molec.cm^{-2}. Note also that a scene-dependent detection threshold of the measurements (typically taken as HRI>3), is conveniently expressed in terms of the absolute uncertainty as $X_{\text{thres}} = 3\sigma_{\text{abs}}$.

The absolute uncertainty contribution is illustrated on the left panels of Figure 3.6.1 for the IASI morning overpass (land observations between 60° S and 60° N), as a function of thermal contrast (TC). As before, we define TC as the brightness temperature of the surface minus the temperature at half the boundary layer height. The absolute uncertainty starts from around $1 \cdot 10^{15}$ molec.cm^{-2} and increases as expected with decreasing thermal contrast, with a global median of $4 \cdot 10^{15}$ molec.cm^{-2}. Observing the inverse proportionality with thermal contrast, the following empirical formula can be used to obtain ballpark estimates of the absolute uncertainty or sensitivity of the IASI NH_3 retrieval (for positive thermal contrasts):

$$\sigma_{\text{abs}} = \frac{3.6 \cdot 10^{16}}{\text{TC}} \frac{\text{molec. K}}{\text{cm}^2} \tag{3.6-8}$$

The constants were determined from a fit of the data shown in Figure 3.6.1. Expressed in terms of Q20 and Q80 quantiles the estimated absolute retrieval uncertainty of IASI (mid-latitude, land, morning overpass) can also be summarized as

$$\sigma_{abs} = [2.5 - 6.6] \cdot 10^{15} \text{molec.cm}^{-2}. \quad (3.6\text{-}9)$$

3.6.3.2. **Relative uncertainty contribution**

The term σ_{rel} is proportional to the column and hence expresses a relative uncertainty for fixed atmospheric conditions:

$$\sigma_{rel} = \sqrt{0.1^2 + \frac{\sigma_{SF}^2}{SF^2}} X \approx \frac{\sigma_{SF}}{|SF|} X \quad (3.6\text{-}10)$$

or

$$\frac{\sigma_{rel}}{(X-B)} = \frac{\sigma_{SF}}{|SF|} \quad (3.6\text{-}11)$$

This term is illustrated on the right panels of Figure 3.6.1. Again, we observe an inverse proportionality with thermal contrast, which can be approximated as

$$\sigma_{rel} = \left(0.07 + \frac{1.6K}{TC}\right) NH_3. \quad (3.6\text{-}12)$$

For typical morning land observations, the relative contribution to the uncertainty starts from around 14% (corresponding to a TC of 20 K). Expressed in terms of Q20 and Q80 quantiles, the estimated relative retrieval uncertainty of ANNI (mid-latitude, land, morning overpass) can be summarized as $\sigma_{rel} = [19 - 36]\%\ NH_3$.

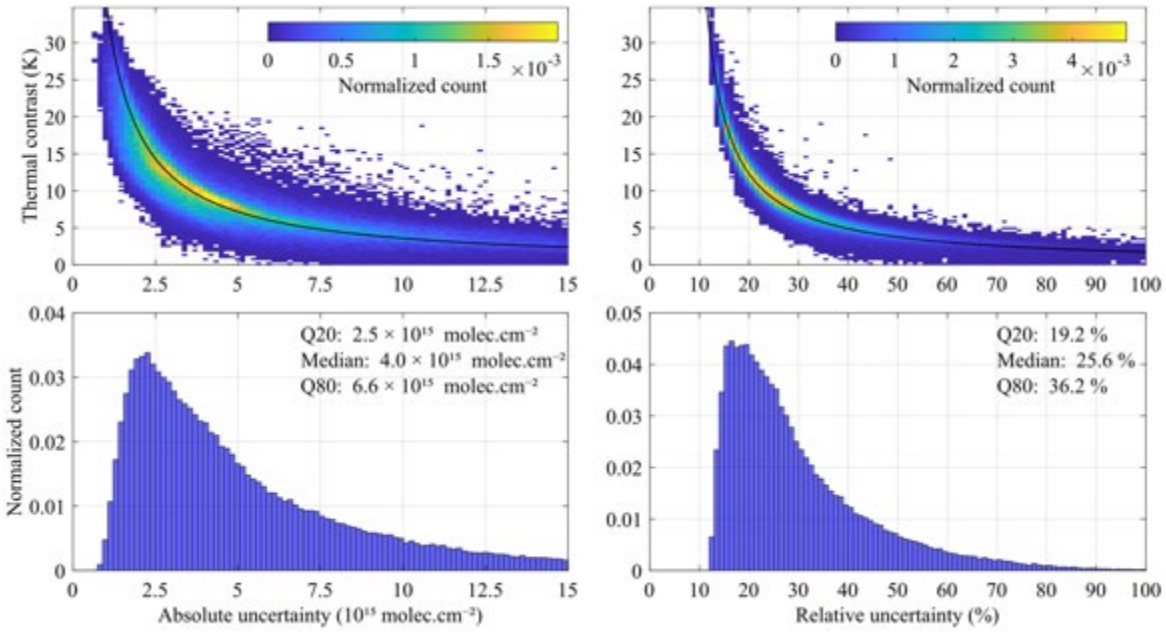


Figure 3.6.1 Illustration of the absolute (left) and relative (right) components of the retrieval uncertainty of NH_3. The top panels illustrate their dependence on thermal contrast, the bottom panels show the normalized count. Data in this plot originates from IASI-B observations on 15 January, April, July and October 2021, morning overpass, land only and between 60° S and 60° N. The approximations from Eq. (3.6.8) and Eq. (3.6.12) are shown in black on the top panels.

4. Level-3 processing

4.1. Uncertainty assessment for nitrogen dioxide (L3-NO2) from OMI and TROPOMI

4.1.1. Procedure

We describe here how we determine the monthly mean Level-3 (L3) TROPOMI tropospheric NO_2 columns and their uncertainties. For OMI, the approach is identical but with some different parameter settings. We start with the v2.3.1 TROPOMI L2 NO_2 retrievals available on https://data-portal.s5p-pal.com/browser/. A small bug with respect to the quality assurance (QA) value over snow and ice was fixed (see Eskes and Eichmann (2023)). The TROPOMI retrievals with qa-values of >0.75 were used, which corresponds to good-quality retrievals over (nearly) cloud-free scenes. Moreover, the retrievals made in the descending part of the orbit were removed, as these showed a different distribution in retrievals likely due to these observations not being included in the data assimilation in the TM5-MP chemistry transport model to retrieve the stratospheric column.

Our starting point are individual retrievals with a retrieved column value x_i (molecules cm^{-2}), the associated retrieval uncertainty σ_i (molecules cm^{-2}) and pixel area a_i (in km^2). We propose a two-step procedure, in which we:

1. calculate spatially averaged (gridded) column values ($x_{o,t}$) – where t stands for an instantaneous column value not averaged in time – along with their associated uncertainties ($\sigma_{o,t}$), and associated coverage degree ($\alpha_{o,t}$), and then
2. provide temporal averaged estimates of the column values ($\underline{x}$), their total uncertainty ($\underline{\sigma}_{total}$), with associated temporal representativity ($\underline{\sigma}_{rt}$) and measurement uncertainties ($\underline{\sigma}_m$).

Figure 4.1.1 provides a schematic picture of the above procedure. After collecting the individual values of the retrieved column, its uncertainty, and the pixel area (x_i, σ_i, a_i), their spatially averaged (instantaneous) counterparts ($x_{o,t}$, $\sigma_{o,t}$, $a_{o,t}$) are calculated for a L3 grid cell. Then, in step 2, these spatial averages are aggregated over time and averaged, leading to the desired L3 product.

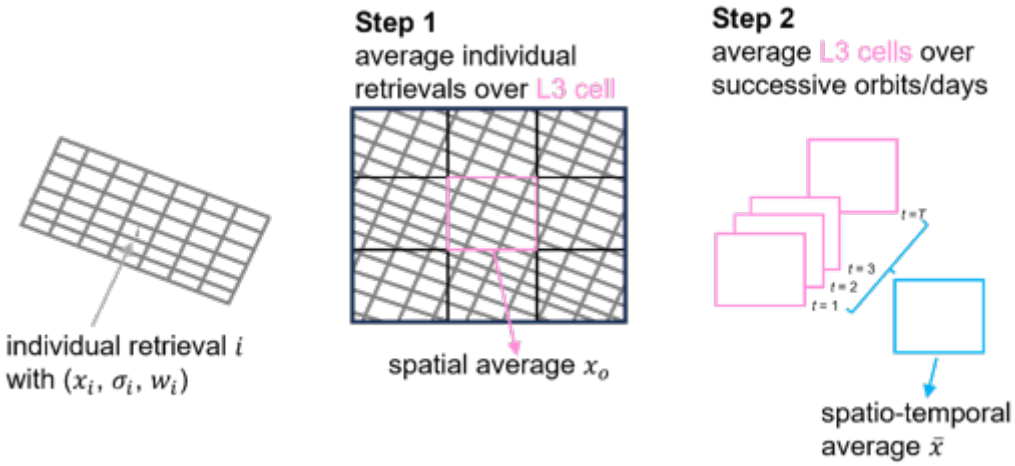


Figure 4.1.1 Schematic example of the procedure to calculate spatio-temporal (gridded) averages of retrieved columns and their associated uncertainties.

4.1.2. Spatial averaging

Step 1 concerns the spatial averaging, sometimes called tiling, of valid retrieved values of the individual pixels (x_i) with weights (w_i) that represent the pixel area (a_i) spatially overlapping with the L3 grid cell ($w_i = a_i$). This results in a best estimated ('tile' or 'superobservation') $x_{o,t}$ of the gridded L3 column:

$$x_{o,t} = \frac{\sum_{i=1}^{N}(w_i x_i)}{\sum_{i=1}^{N} w_i} \quad (4.1\text{-}1)$$

with N the number of valid observations in the L3 cell. We take an area-weighted averaging approach, because the retrieval provides column densities (x_i) that hold for pixels with variable area sizes that (partially) overlap the grid cell the spatial average is taken on (see Figure 4.1.1). The total number of NO_2 molecules accumulated over valid pixels should thus be weighted by the pixel area overlapping with the L3 grid cell (a_i).

The spatial average is determined on a (regular) grid on a per-orbit basis. The spatial average for a grid cell is only determined if the combined individual retrievals x_i cover at least 30% of the grid cell following Rijsdijk et al. (2024).

To assess the overall uncertainty ($\sigma_{o,t}$) in $x_{o,t}$ of these spatial averages we follow the procedure set out in Rijsdijk et al. (2024). The overall uncertainty combines the measurement uncertainties ($\sigma_{m,s}$) and the spatial representativeness uncertainty (σ_{rs}) in quadrature:

$$\sigma_{o,t} = \sqrt{\sum_{s=1}^{S}\left(\sigma_{m,s}\right)^2 + (\sigma_{rs})^2} \quad (4.1\text{-}2)$$

with S the number of independent measurement error sources in the L2 product.

The measurement uncertainty σ_m is the combined uncertainty from error sources ($\sigma_{m,s}$) in the measurement, including the uncertainty in the slant column density measurements (σ_{SC}), the uncertainty resulting from errors in the stratospheric column (σ_{strat}), and the uncertainty in the air mass factor (σ_{AMF}). Each of these are represented by Eq. (4.1-3) (Sekiya et al., 2022; Rijsdijk et al., 2024) that considers that random component of the individual uncertainties $\sigma_{i,s}$ tend to cancel when averaging over many observations (first term), while a fraction c of the individual uncertainties $\sigma_{i,s}$ persists after averaging because they come from systematic spatially correlated contribution to the uncertainties (second term):

$$\sigma_{m,s} = \sqrt{(1-c)\frac{\sum_{i=1}^{N}\left(w_i^2\sigma_{i,s}^2\right)}{\sum_{i=1}^{N} w_i^2} + c\frac{\left(\sum_{i=1}^{N}\left(w_i\sigma_{i,s}\right)\right)^2}{\sum_{i=1}^{N} w_i^2}} \quad (4.1\text{-}3)$$

The spatial correlation factor c in Eq. (4.1-3) is determined for each of the measurement error contributions (slant column density, stratospheric column, and air-mass factor) separately. Following Rijsdijk et al. (2024) the uncertainty in slant column density is assumed to be fully uncorrelated ($c = 0$), the uncertainty in the stratospheric column is assumed to be fully correlated ($c = 1$), and the uncertainty in the AMF is assumed to be partly correlated depending on the size of the

superobservation grid cells ($c = exp(-d/l)$, where d is a distance between observations depending on the grid cell size and $l = 35$ km is the correlation length). See Rijsdijk et al. (2024) for more details on how these correlation fractions are determined.

The spatial representativeness uncertainty σ_{rs} accounts for incomplete sampling (mainly due to cloud cover) of the cell by the N observations available to calculate $x_{o,t}$, and is expressed as a fraction of the standard deviation in the L3 values σ_{x_i} (Rijsdijk et al., 2024). If the entire area of the L3 cell is covered by valid observations, then the representativeness uncertainty is 0. If only a small fraction of the L3 cell is covered, the representativeness uncertainty is large for areas with strong spatial variability in x_i (such as over polluted regions), and smaller for regions with similar values of x_i (such as over clean, background regions). The fraction r is a function of the coverage degree α of the cell by the individual values x_i, the size of the cell itself, and this function will differ for different trace gases and instruments as these have their own spatial distributions:

$$\sigma_{rs} = r\sigma_{x_i} \tag{4.1-4}$$

where σ_{x_i} is the standard deviation in retrievals within the grid cell (x_i) and r is the representativity of the retrieved observations for the fully covered grid cell:

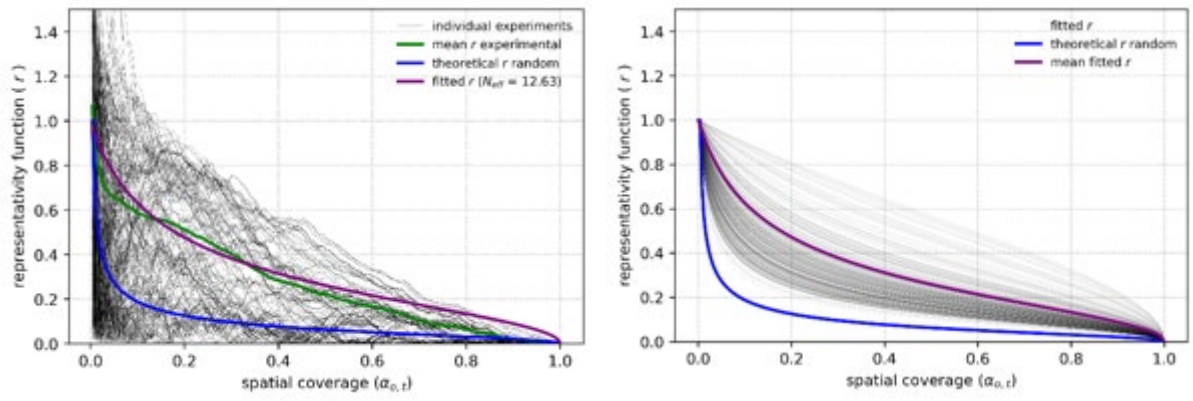
$$r = \frac{1}{\sqrt{N_{eff}\alpha_{o,t} + 1}}\sqrt{1 - \alpha_{o,t}} \tag{4.1-5}$$

where coverage degree $\alpha_{o,t}$ is calculated as the fraction of total valid pixel area (A, in km^2)

$$\alpha_{o,t} = \frac{\sum_{i=1}^{N} w_i}{A} \tag{4.1-6}$$

and N_{eff} is the effective population size, which is dependent on the number of available observations, the sensor, the trace gas, the L3 resolution gridding to, and whether a grid cell is sensitive to systemic sampling or not (Rijsdijk et al., 2024). Systematic sampling can occur when a cloud covers part of the superobservation grid cell. A low value for N_{eff} indicates that a grid cell is sensitive to systematic sampling.

The value for N_{eff} was estimated by Rijsdijk et al. (2024) by repeatedly sampling a fully covered grid cell in a systematic way (starting with one random pixel and then adding neighbouring pixels) and calculating the representativity (the absolute difference between the mean of observed pixels and the true mean of the fully covered grid cell when including all pixels). The representativity function r in equation 5 was then fitted to the mean of these experiments, giving the effective population size N_{eff} for this grid cell in a polluted area (Figure 4.1.2(a)).



(a) (b)

Figure 4.1.2. Results of repeatedly sampling grid cells in a systematic way to calculate the representativity function $\boldsymbol{r}$. (a) One single grid cell (polluted, $x_{o,t} \geq 30\ \mu mol$ m^{-2} or $\geq 1.8\ 10^{15}$ molec. cm^{-2}) sampled 200 times. The grey lines represent 200 individual experiments sampling the same superobservation grid cell. The green line is the mean of the individual experiments and the purple curve is the representativity function in equation 5 fitted to the green line. The blue curve shows the theoretical representativity function (discussed in section 4.1.3.4, equation 4.1-13) in the case of random sampling (result if grid cell is not sensitive to systematic sampling). (b) Results of sampling 100 grid cells in polluted regions. The grey lines represent a collection of purple lines from (a), the purple curve is the mean of the grey curves. The blue line again shows the theoretical random solution.

This method was repeated for multiple superobservation grid cells (Figure 4.1.2(b)). It was shown that polluted regions ($x_{o,t} \geq 30\ \mu mol$ m^2 or $\geq 1.8\ 10^{15}$ molec. cm^{-2}) are more sensitive than unpolluted regions ($x_{o,t} < 30\ \mu mol$ m^2 or $< 1.8\ 10^{15}$ molec. cm^{-2}) and have a lower value for N_{eff} and thus a higher representativity function and uncertainty (Rijsdijk et al., 2024). Overall, N_{eff} depends on the spatial structure of the concentration field within the superobservation and indicates the effect of systematic sampling on a given superobservation.

The ratio N/N_{eff}, where N is the total number of pixels (valid or not) in a grid cell, was determined for NO_2 retrievals from TROPOMI in Rijsdijk et al. (2024) to be linearly dependent on the grid resolution. For a superobservation resolution of 0.2°x0.2° it was established that $N/N_{eff} = 1.376$ for unpolluted areas which are not sensitive to systemic sampling and $N/N_{eff} = 3.933$ for polluted areas which are more sensitive to systemic sampling (Table 4.1.1). These ratios are used to determine the effective population size N_{eff} for each grid cell in the spatial mean. This variable describes how sensitive a region is to systematic sampling, where a higher N/N_{eff} ratio means a region is more sensitive to systematic sampling and thus results in a higher representativity uncertainty. The ratios N/N_{eff} hold for NO_2 fields and would need to be redetermined for other trace gases or sensors following the methods in Rijsdijk et al. (2024).

Table 4.1.1 The ratio $\boldsymbol{N/N_{eff}}$ for TROPOMI NO_2 in unpolluted and polluted ($\boldsymbol{x_{o,t}} \geq 1.8\ 10^{15}$ molec. cm^{-2}) regions for different spatial resolutions of the superobservation grid. The ratio increases linearly with superobservation area (degree^2).

	N/N_{eff} unpolluted areas	N/N_{eff} polluted areas
0.2°x0.2°	1.376	3.933
0.5°x0.5°	1.890	7.392
1.0°x1.0°	3.724	19.746
2.0°x2.5°	13.508	85.634

4.1.3. Temporal averaging

4.1.3.1. Averaging of variables

Step 2 averages the gridbox-mean values $x_{o,t}$ obtained in step 1 over time, again in a weighted fashion to account for different degrees of representativity ($w_{o,t} = 1 - r$). r is high for superobservations with a large representativity uncertainty and low for superobservations with almost no representativity uncertainty. This representativity value depends on coverage – higher coverage results in a smaller representativity error – but also other factors such as sensitivity to systematic sampling. Taking a weighted average like this means that superobservations with a low representativity error get more weight in the temporal average than superobservations with a high representativity error. This results in the following estimate of the spatially and temporally averaged L3 column $\bar{x}$:

$$\bar{x} = \frac{\sum_{t=1}^{T}\left(w_{o,t}x_{o,t}\right)}{\sum_{t=1}^{T} w_{o,t}} \quad (4.1\text{-}7)$$

with T the total number of valid superobservations in the period over which the averaging is performed.

4.1.3.2. Uncertainty estimate

The uncertainty associated with $\bar{x}$ can be estimated in different ways. We propose to calculate a total spatio-temporal averaged uncertainty which combines the propagated measurement uncertainty ($\bar{\sigma}_{\boldsymbol{m}}$, including the spatial representativity uncertainty) and a temporal representativity uncertainty ($\bar{\sigma}_{\boldsymbol{rt}}$):

$$\bar{\sigma}_{total} = \sqrt{(\bar{\sigma}_{\boldsymbol{m}})^{\boldsymbol{2}} + (\bar{\sigma}_{\boldsymbol{rt}})^{\boldsymbol{2}}} \quad (4.1\text{-}8)$$

4.1.3.3. Propagated measurement uncertainty

The spatio-temporally averaged measurement uncertainty $\bar{\sigma}_{\boldsymbol{m}}$ is the combined uncertainty of the measurement error sources $\bar{\sigma}_{m,s}$ including the spatial representativity uncertainty:

$$\bar{\sigma}_{\boldsymbol{m}} = \sqrt{(\bar{\sigma}_{strat})^{\boldsymbol{2}} + (\bar{\sigma}_{SC})^{\boldsymbol{2}} + (\bar{\sigma}_{AMF})^{\boldsymbol{2}} + (\bar{\sigma}_{rs})^{\boldsymbol{2}} + \left(\bar{\sigma}_{apriori}\right)^{\boldsymbol{2}}} \quad (4.1\text{-}9)$$

The propagated slant column density uncertainty ($\bar{\sigma}_{SC}$), stratospheric column uncertainty ($\bar{\sigma}_{strat}$), air-mass factor uncertainty ($\bar{\sigma}_{AMF}$) and spatial representativity uncertainty $(\bar{\sigma}_{rs})^{\boldsymbol{2}}$ are each determined using:

$$\bar{\sigma}_{m,s} = \sqrt{(1-c)\frac{\sum_{\boldsymbol{t=1}}^{\boldsymbol{T}}\left(w_{o,t}{}^{2}\sigma_{m,s_t}{}^{2}\right)}{\left(\sum_{\boldsymbol{t=1}}^{\boldsymbol{T}} w_{o,t}\right)^{\boldsymbol{2}}} + c\frac{\left(\sum_{\boldsymbol{t=1}}^{\boldsymbol{T}}\left(w_{o,t}\sigma_{m,s_t}\right)\right)^{\boldsymbol{2}}}{\left(\sum_{\boldsymbol{t=1}}^{\boldsymbol{T}} w_{o,t}\right)^{\boldsymbol{2}}}} \quad (4.1\text{-}10)$$

where $\boldsymbol{c}$ is the temporal correlation factor of the uncertainties (not necessarily the same as the spatial correlation factor applied in equation (4.1-3)). The last term in equation (4.1-9) represents the

contribution from the uncertainty in the a priori profile shapes, and is approximated as 10% of the tropospheric AMF (Boersma et al., 2018):

$$\bar{\sigma}_{apriori} = \bar{x}\frac{\sigma_{apriori}}{M_{tr}} = \bar{x}\frac{0.1M_{tr}}{M_{tr}} = 0.1\bar{x} \quad (4.1\text{-}11)$$

where M_{tr} is the tropospheric air-mass factor. $\bar{\sigma}_{apriori}$ is not included in the uncertainty calculation when using the averaging kernel in data applications, which removes the dependence on the a priori profile (Eskes and Boersma, 2003).

The temporal correlation factor $\boldsymbol{c}$ is determined for each of the sources of measurement uncertainty (slant column density, stratospheric, and AMF) separately.

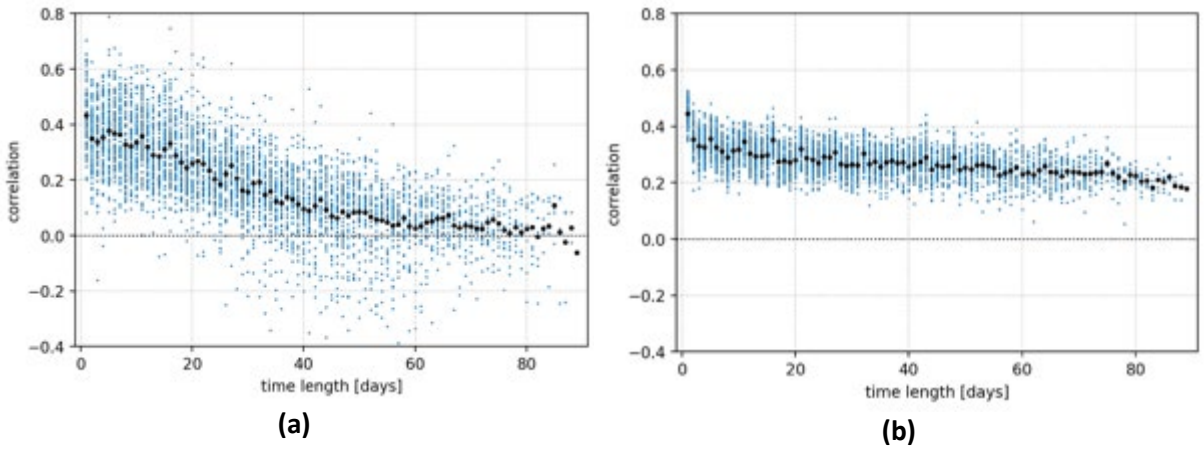


Figure 4.1.3 Temporal correlation in (a) difference of tropospheric NO_2 column for v2.3.1 and v2.4 for polluted areas (AMF uncertainty) and (b) observation-forecast (stratospheric column uncertainty) between each combination of two days within the same month for the period 1 Jan – 31 Mar 2019. Larger points are the mean of correlations with the same time length.

First, the temporal correlation factor for the AMF uncertainty is determined. The a priori NO_2 profile is a large contribution to the uncertainty in the AMF but is shown to become independent when the averaging kernel is used (Eskes and Boersma, 2003). Other large sources of uncertainty in the AMF are the effective cloud cover, the effective cloud height, and the surface albedo. All three of these variables depend on a monthly climatological surface albedo dataset. This introduces a temporal correlation in the uncertainties of daily observations. We estimate the temporal correlation required to determine the uncertainty in spatio-temporal averaged L3 data by comparing versions 2.3.1 and 2.4 of the L2 retrieval. These versions apply different climatological surface albedo datasets, with v2.3.1 applying albedo derived from OMI and GOME-2 and v2.4 applying albedo derived from TROPOMI spectra. As this is the largest change between the two L2 versions, the difference between the versions should be representative of the uncertainty resulting from the climatological surface albedo. As the albedo is a key input for the cloud retrieval, this replacement of climatological albedo dataset also generates cloud fraction and cloud pressure retrieval differences. We determine the difference in tropospheric NO_2 column in 0.2°x0.2° superobservations from v2.3.1 and v2.4 L2 data for each day in the period 1 January to 31 March 2019. For each combination of two days within the same month, we determine

the Pearson correlation coefficient in tropospheric NO_2 for polluted areas (tropospheric column from both versions $> 30\mu mol\ m^{-2}$) (Figure 4.1.3(a)):

$$\boldsymbol{r_{\Delta t}} = \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{I}}\left(\Delta A_{t_1 i} - \overline{\Delta A_{t_1}}\right)\left(\Delta A_{t_2 i} - \overline{\Delta A_{t_2}}\right)}{\sqrt{\sum_{\boldsymbol{i=1}}^{\boldsymbol{I}}\left(\boldsymbol{\Delta} A_{t_1 i} - \overline{\Delta A_{t_1}}\right)^{\boldsymbol{2}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{I}}\left(\Delta A_{t_2 i} - \overline{\Delta A_{t_2}}\right)^{\boldsymbol{2}}}} \quad (4.1\text{-}12)$$

where $\boldsymbol{r_{\Delta t}}$ is the Pearson correlation coefficient between any two days in the time period (blue points in Figure 3-3a), $\Delta A_{t_1} = A_{t_1}^{v2.4} - A_{t_1}^{v2.3.1}$ and $\Delta A_{t_2} = A_{t_2}^{v2.4} - A_{t_2}^{v2.3.1}$ the difference in surface albedo between v2.3.1 and v2.4 for the first day and second day respectively, and I the number of valid superobservation grid cells for both days. For each 'time length' between daily superobservations, we determine the mean correlation (black points in Figure 4.1.3(a)). Then we take the mean of the mean correlations for the time lengths 1-30 days. This results in $c = 0.29$. The same method is carried out for the NH summer period 1 June-31 August, which showed almost the same result ($c = 0.30$). The temporal uncertainty correlation in the AMF is estimated to be $c = 0.30$ for both the NH winter and summer period. When analysing longer time lengths than a month, the correlation disappears. The method was repeated for 1°x1° superobservations, resulting in the same results, giving us confidence that the temporal AMF correlation coefficient is not grid-size dependent.

Next, the temporal correlation factor for the stratospheric uncertainty is determined. In the retrieval of the NO_2 columns the observed slant column is split into its tropospheric and stratospheric parts using data-assimilation in the TM5-MP model. We can assess uncertainty and bias in this model by looking at the difference between the forecasted stratospheric column and the observed stratospheric column (O-F, see e.g. Dirksen et al. (2011)). We take the values for O-F for the period 1 January to 31 March 2019. For each combination of two days within the same month, we determine the Pearson correlation coefficient in stratosphere O-F for clean areas (tropospheric column $< 10\mu mol\ m^{-2}$) (Figure 3(b)). This is again done with equation (12) but now with $\Delta A_{t_1} = A_{t_1}^{observation} - A_{t_1}^{forecast}$ and $\Delta A_{t_2} = A_{t_2}^{observation} - A_{t_2}^{forecast}$ the difference in surface albedo between observation and forecast for the first day and second day respectively. For each `time length' between daily O-F maps, we determine the mean correlation. Then we take the mean of the mean correlations for the time lengths 1-30 days. This results in $c = 0.30$. The same method is carried out for the NH summer period 1 June-31 August, which showed a slightly lower correlation coefficient of $c = 0.21$. We take the higher value of $c = 0.30$ as the temporal uncertainty correlation in the stratospheric uncertainty for both the NH winter and summer period, as a conservative uncertainty estimate. The correlation decreases slightly for longer periods and only disappears over periods longer than 3 months.

The uncertainty in the slant column density is assumed to be fully uncorrelated ($c = 0$). The slant column density uncertainty is largely driven by instrument noise, which is random. The slant column density uncertainty also has a systematic part, resulting from gaps in knowledge such as inaccurate Ring coefficients in the spectral fit or lack of an intensity offset and a correction for vibrational Raman scattering. This systematic component is present in the slant column and influences the observation-forecast discussed in the stratospheric uncertainty. Considering the effect of the systematic error is already included in the stratospheric correlation discussed above, we do not separately quantify the effect of the systematic retrieval error on the slant column density. Instead, the slant column density uncertainty is assumed fully random.

The temporal correlation coefficient of the spatial representativity uncertainty is set at fully random ($c = 0$). The representativity can be correlated through time because of persistent cloud cover, but this will be assessed separately in the temporal representativity uncertainty.

Table 4.1.2 Overview of the spatial and temporal correlations in the components of the uncertainty estimate.

	spatial	temporal
SCD uncertainty	0	0
Stratospheric uncertainty	1	0.3
AMF uncertainty	$e^{-d/35km}$	0.3
Spatial representativity	N/A	0
Temporal representativity	N/A	N/A

The monthly mean L3 NO_2 columns are only representative when enough observations are included in the means. The L3 qa-value is included in the dataset and equals 1 when the data is representative as a monthly mean and equals 0 when it is not. This deviation is made using the count variable, which is the sum of the fractional coverages from the spatial averages $\alpha_{o,t}$ divided by the number of days in the month. If this value is below 0.1 (meaning less than 10% of the month is sampled), the L3 qa-value is set at 0 and the L3 NO_2 columns should not be used in analysis.

4.1.3.4. Temporal representativity uncertainty

The temporal representativity uncertainty ($\overline{\sigma_{rt}}$) is seen as the standard error of the superobservations (spatial means) used to determine the monthly mean. We state that the number of possible observations within a month is finite, and a full monthly observation is reached with at least one observation each day. This is not truly the case, because the satellite orbits the Earth and does not give continuous observations over each location. However, when we say the monthly mean is a representation of midday columns (the sensors applied here make observations around 13:30h local time), in principle we could have a valid observation available every day, meaning a finite number of possible midday observations. A correction factor is applied to the method of calculating the standard error to correct for this (Bondy and Zlot, 1976; Isserlis, 1918). This results in the following theoretical formula for the temporal representativity uncertainty:

$$\overline{\sigma_{rt}} = \frac{\bar{\sigma}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \quad (4.1\text{-}13)$$

where N is the length of the month in days and n is the number of days with at least one valid superobservation. The weighted standard deviation around the temporal mean $\bar{\sigma}$ is determined using:

$$\bar{\sigma} = \sqrt{\frac{\sum_{\boldsymbol{t=1}}^{\boldsymbol{T}}\left(w_{o,t}\left(x_{o,t}-\bar{x}\right)^2\right)}{(T-1)\sum_{t=1}^{T} w_{o,t}}} \quad (4.1\text{-}14)$$

with $\boldsymbol{T}$ the total number of valid orbits in the period over which the averaging is performed.

We apply the method also used in Rijsdijk et al. (2024) and in the spatial means to assess whether this function is suitable for assessing the temporal representativity error in January 2019. We start by taking a completely covered grid cell, which means that at least one spatial mean/superobservation was produced for this grid cell for every day of the month. The representativity error can be quantified

by comparing the mean of this completely covered grid cell to the mean of several samples (days) taken from that grid cell. We start by sampling a single day, which we use to estimate the mean of the grid cell. Then we repeatedly add random days of the month and estimate the mean using the available observations. We perform multiple iterations (200) of this process on the same grid cell and find the relation between representativity error and temporal coverage of the grid cell (Figure 4.1.4(a)). This process is then repeated for 100 grid cells in polluted areas (Figure 4.1.4(b)). We only take polluted grid cells into account (mean NO_2 tropospheric column of $\geq 2\ 10^{15}$ molec/cm^2), as these are expected to have the highest temporal variability and thus more likely to be sensitive to a sampling bias. The mean of the experiments (green) is almost identical as the theoretical solution of equation 4.1-13 (blue), showing that the formula is a suitable method for quantifying the temporal representativity error for random sampling. Similar results are found for June 2019.

As with the spatial representativity (cloud masking one part of the grid cell), there might be a sensitivity to systematic sampling temporally (cloud occurring over subsequent days). This could for example be the case if large weather systems give persistent cloud cover in a given location for part of the month. The occurrence of systematic sampling was tested for the spatial means of January and June 2019. We selected grid cells in the spatial mean that had valid superobservations for ~50% of the given month. On these grid cells we applied a Wald-Wolfowitz test to determine whether the cloudiness was random or systematic (Wald and Wolfowitz, 1943). The Wald-Wolfowitz test is a non-parametric statistical test that checks a randomness hypothesis for a two-valued data sequence, in our case cloudy (no data) or not-cloudy (superobservation available, can be partly cloudy). For both January and June about two-thirds of the grid cells show systematic sampling due to continuous cloud cover for part of the month. This shows that systematic sampling of observations for the temporal mean is relevant. Next, we will investigate whether the temporal variability in the observations makes it sensitive to this systematic sampling. We repeat the method of repeated sampling of a grid cell as outlined above, but now by sampling the grid cell systematically (start by taking the observation of one random day and then repeatedly add adjacent days). Figure 4.1.4(d) shows that systematic sampling produces more variability of grid cells around the theoretical solution than with random sampling, but that the mean of the experiments is close to the theoretical solution and comparing with the spatial representativity curve in Figure 4.1.2 shows that the experimental curves in Figure 4.1.4(d) are closer to the theoretical solution. The fit from equation 4.1-5 in purple does not show a better fit to the experimental mean than the theoretical solution does. This shows that the temporal mean is not very sensitive to systematic sampling and the uncertainty can be assessed with the theoretical solution suggested in equation 4.1-13.

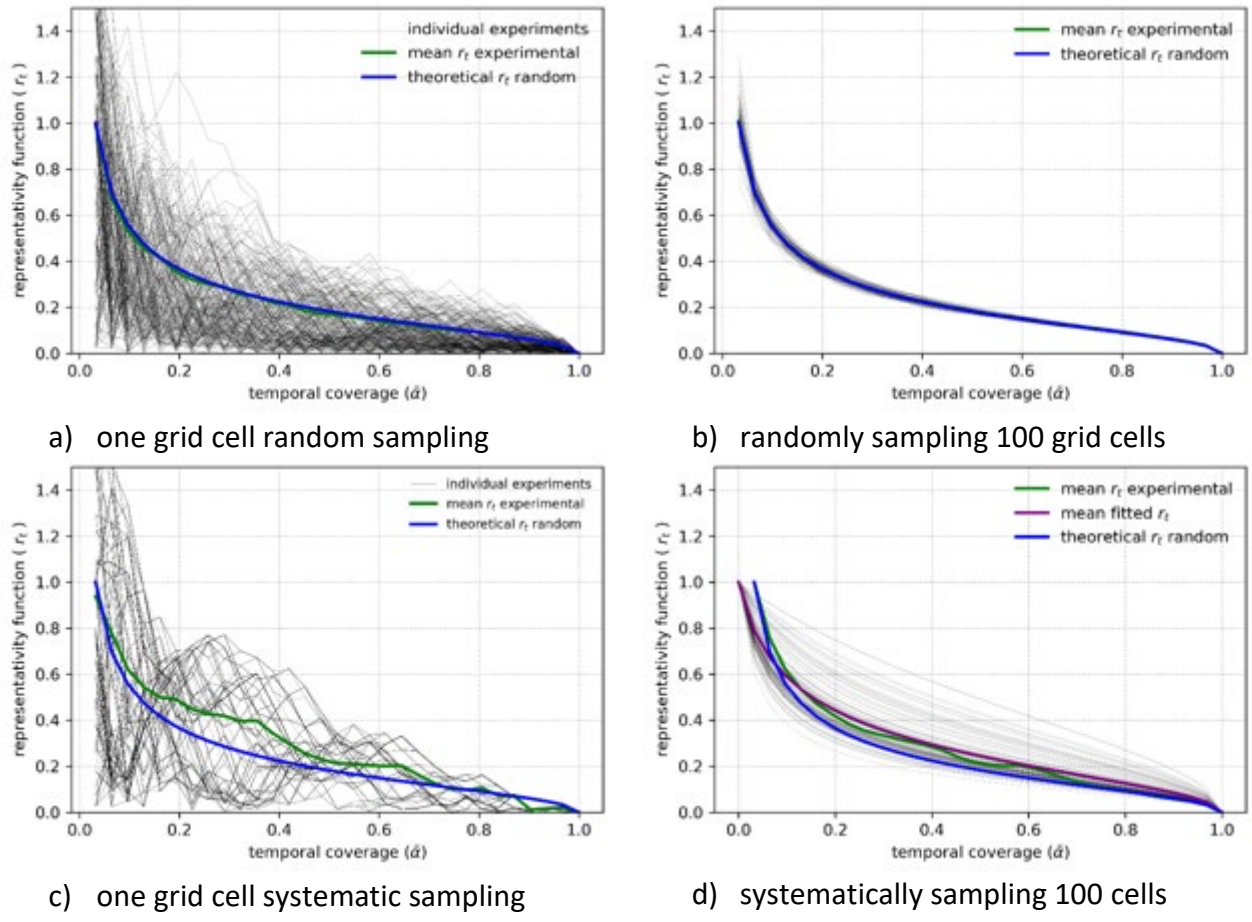


Figure 4.1.4 Results of repeatedly sampling grid cells to calculate $\overline{\sigma_{rt}}$. (a) Repeated random sampling of 31 days of a single grid cell with values $x_{o,t}$. Daily observations in the grid cell are randomly sampled 200 times. The thin grey lines represent the difference between the sampled mean and the actual mean from individual random experiments. The green line is the mean of the samples and the blue line the theoretical result from equation 13. (b) Results of randomly sampling 100 grid cells. The grey lines are the green line from panel a. The green line (overlapping the blue theoretical result) is the mean of the results of 100 grid cells. (c) Systematic sampling of a single grid cell. (d) Results of systematically sampling 100 grid cells. The grey lines are a collection of fits through the green line in panel c with equation 5. The purple line is the mean of the grey lines. The green line is the mean of the green line in panel c and other grid cells. The blue line is the theoretical mean from equation 13. Observations in all panels are from January 2019 and have a mean tropospheric column of $\geq 2 \cdot 10^{15} molec/cm^2$. Only grid cells with observation available every day are sampled.

It should be noted that this method can only be applied to grid cells that provide coverage for all days of the month. This is only true for arid areas with almost no cloud cover. The question arises how representative these regions are. We filter for relatively polluted regions ($\geq$2 10^{15} molec/cm^2) and find more than 100 grid cells to which we can apply this method. These are found mostly in cities in the Middle East and western India, where January is in the middle of the dry season. We assume these polluted areas are representative for polluted areas in other parts of the world and the theoretical solution can be applied globally.

Table 4.1.3 Summary of the concepts and proposed calculations to arrive at spatio-temporally averaged L3 values and associated uncertainties.

	Level-2 product (i.e. pixel)	Spatial average (superobservation)	Temporal average (L3) at superobs resolution
Column	x_i	$x_{o,t} = \frac{\sum_{i=1}^{N}(w_i x_i)}{\sum_{i=1}^{N} w_i}$	$\bar{x} = \frac{\sum_{t=1}^{T}(w_{o,t} x_{o,t})}{\sum_{t=1}^{T} w_{o,t}}$
Uncertainty	σ_i	$\sigma_{o,t} = \sqrt{\sum_{s=1}^{S}(\sigma_{m,s})^2 + (\sigma_{rs})^2}$	$\bar{\sigma}_{\boldsymbol{m}} = \sqrt{(\bar{\sigma}_{strat})^2 + (\bar{\sigma}_{SC})^2 + (\bar{\sigma}_{AMF})^2 + (\bar{\sigma}_{rs})^2 + (\bar{\sigma}_{apriori})^2}$
Weight		$w_i = a_i$	$w_{o,t} = 1 - \frac{1}{\sqrt{N_{eff}\alpha_{o,t} + 1}}\sqrt{1 - \alpha_{o,t}}$
Fractional coverage	a_i	$\alpha_{o,t} = \frac{\sum_{i=1}^{N} w_i}{A}$	$\bar{\alpha} = \frac{\sum_{t=1}^{T}(w_{o,t}\alpha_{o,t})}{\sum_{t=1}^{T} w_{o,t}}$

4.1.4 Analysis of L3 uncertainty budget for TROPOMI

The results of the spatio-temporal average monthly mean and uncertainties for NO_2 for January and June 2019 are shown in Figure 4.1.5, with local values given in Table 4.1.4. In the urbanised areas in the Northern Hemisphere, there are higher concentrations of NO_2 in the atmosphere during the winter season, and thus we see higher values over Amsterdam and Beijing in January than in June (Table 4.1.4). The total uncertainty when not applying the averaging kernel attributes to 21% and 17% of the total tropospheric column in January for Amsterdam and Beijing respectively. These percentages get lowered to 19 and 14% when applying the averaging kernel. In June the tropospheric NO_2 columns are high in Africa due to wildfires and biomass burning in the dry season.

Figure 4.1.7 sets out the separate contributions to the Level 3 uncertainty. The largest contributions to uncertainty are the stratospheric uncertainty and the AMF uncertainty. The spatial and temporal representativity uncertainty contribute to the total uncertainty in cloudy and polluted regions such as eastern China and Europe. The total uncertainty $\bar{\sigma}_{\boldsymbol{total}}$ as determined in this document is clearly larger than the uncertainty if all contributions were assumed fully uncorrelated (random uncertainty) and smaller than the uncertainty if all contributions were assumed fully correlated (systematic uncertainty) (Figure 4.1.7). The total uncertainty is less than 50% of the systematic uncertainty in most regions and ~60% of the systematic uncertainty in polluted regions. The total uncertainty is between 1.5 times and more than 5 times more than the fully random uncertainty.

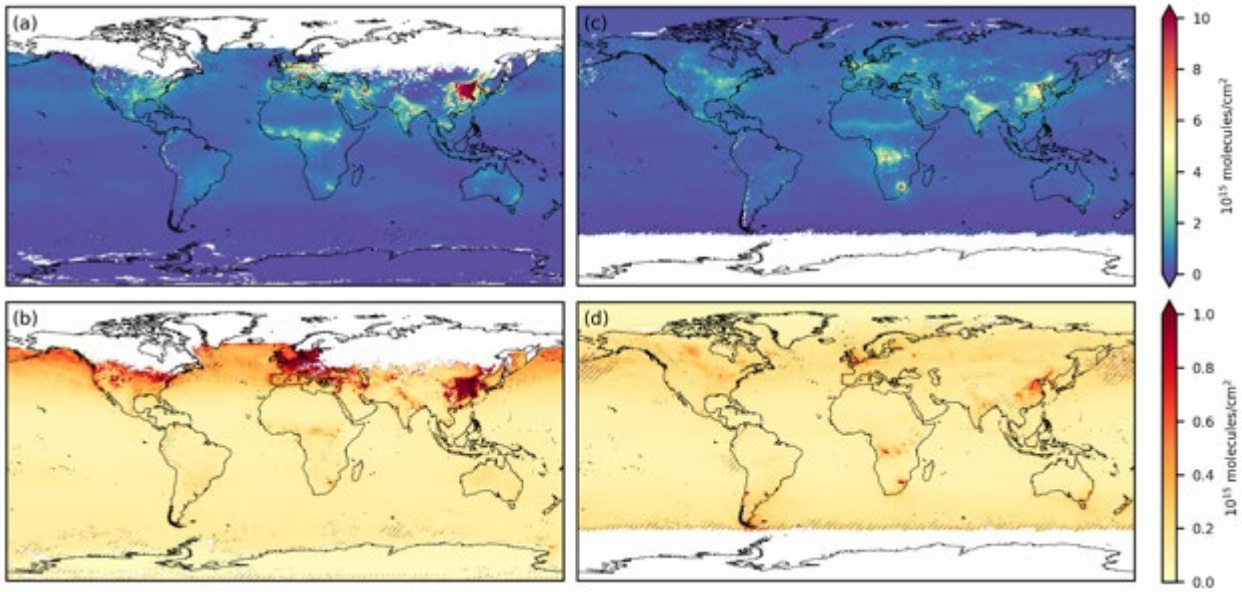


Figure 4.1.5 Example of results of the monthly (a) tropospheric NO_2 column with its (b) total uncertainty ($\bar{\sigma}_{total}$) for January 2019 and the same (c,d) for June 2019 at a grid resolution of 0.2°x0.2°. Hatched where qa=0.

Table 4.1.4 Example values for monthly mean tropospheric NO_2 column and uncertainties at 0.2°x0.2° (with percentage of mean tropospheric column) for a few selected locations for January and June (units: 10^{15} molecules/cm^2). The values for Amsterdam and Beijing were taken from one grid cell in the city centre location (52.3°N 4.9°E for Amsterdam and 39.9°N 116.3°E for Beijing). The value for Africa biomass burning represents the location with the highest observed monthly mean value in the sub-equatorial region on the continent of Africa (in a box between latitudes 16°S and 16°N and longitudes 17.5°E and 42°E). The uncertainty estimates are with the a priori uncertainty, which does not need to be included when using the averaging kernel.

		Amsterdam (0.2°x0.2°)	Beijing (0.2°x0.2°)	Africa biomass burning (32°x40.5°)
$\bar{x}$	Jan	5.45	32.28	7.20
	Jun	4.58	11.04	14.63
$\bar{\sigma}_{total}$	Jan	1.16 (21%)	5.51 (17%)	1.06 (15%)
	Jun	0.78 (17%)	1.73 (16%)	2.10 (14%)

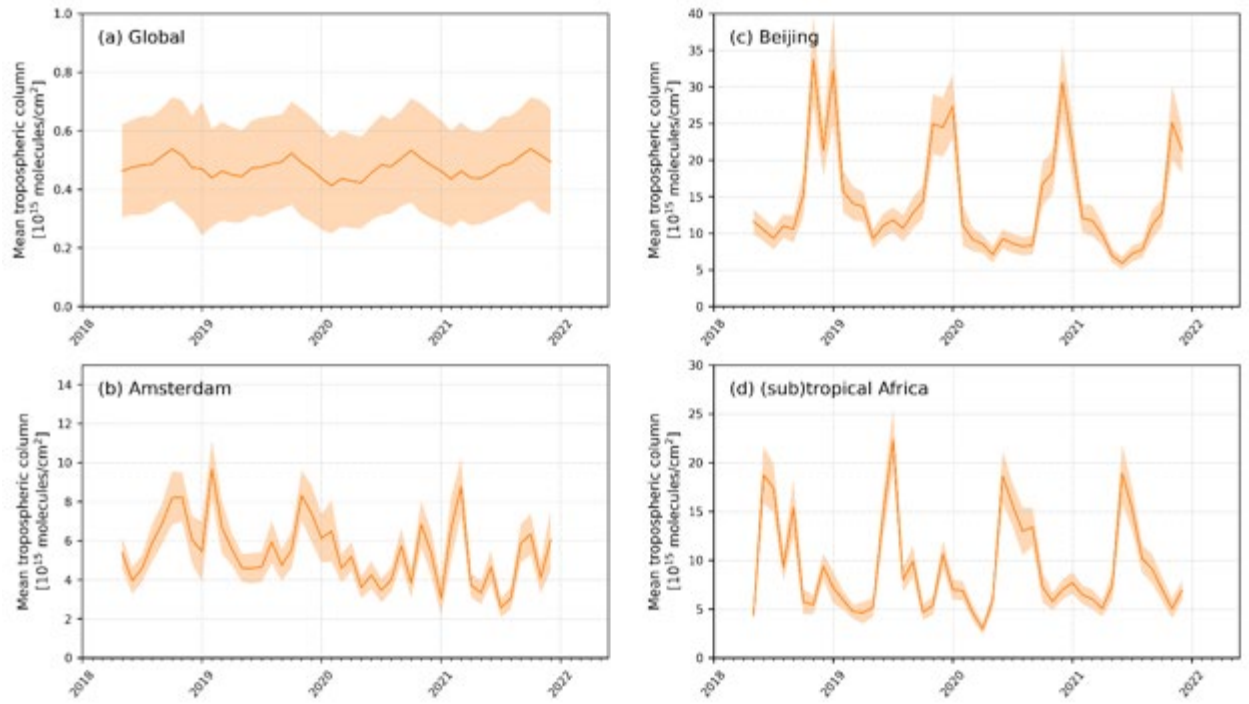


Figure 4.1.6 Time series of the monthly mean tropospheric column NO_2 with the total uncertainty ($\bar{\sigma}_{total}$) for the same locations as in Table 4.1.4. The spatial resolution for (b) and (c) is one grid cell of 0.2°x0.2°, and for (sub)tropical Africa covers multiple grid cells over 32° latitude x 40.5° longitude.

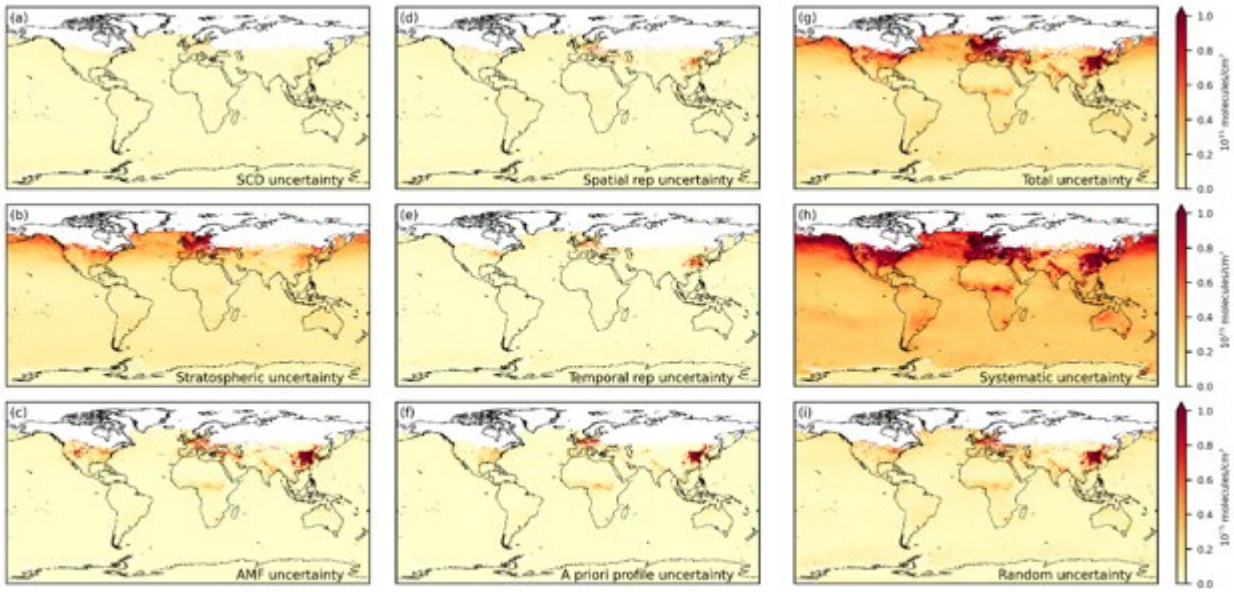


Figure 4.1.7. All sources of uncertainty (a-f) and the combined total uncertainty estimate (g) using the discussed spatial and temporal correlations for January 2019. The total uncertainty was also assessed assuming fully systematic (h) and fully random (i) uncertainty both spatially and temporally.

4.2 Uncertainty assessment for nitrogen dioxide (L3-NO2) from morning sensors

4.2.1 Overview

The level-3 NO_2 data products for morning sensors are developed with the aim of providing user-friendly datasets that support both scientific analysis, such as long-term trend analysis and low-frequency climate variability, and address public interest. Generating level-3 NO_2 data requires binning of level-2 data onto a regular two-dimensional latitude-longitude grid in a specified temporal period, i.e. a day or a month.

In this project, we utilize GOME and SCIAMACHY level-2 NO_2 data from the QA4ECV v1.1 product (Boersma et al., 2018), as well as GOME-2/MetOp-ABC level-2 NO_2 data from the AC-SAF prototype product. In the L2 NO_2 data products, uncertainties are provided separately for each individual observation. The uncertainty estimates are based on error propagation. Table 4.2.1 summarizes the key characteristics of all morning sensor instruments and their level-2 processors. The gridding and uncertainty calculation for level-3 processing are described in following sections.

Table 4.2.1. Summary of characteristics of the morning sensor instruments and level-2 data products.

Instrument	Operational period	Ground-pixel resolution	Equator crossing time	Global coverage	Level-2 processor
GOME	1995 – 2003	320 x 40 km^2	10:30 LT	3 days	QA4ECV v1.1
SCIAMACHY	2002 – 2012	60 x 30 km^2	10:00 LT	6 days	QA4ECV v1.1
GOME-2A	2007 – 2021	80 x 40 km^2; 40 x 40 km^2	09:30 LT	1 day	AC-SAF prototype
GOME-2B	2013 – present	80 x 40 km^2	09:30 LT	1 day	AC-SAF prototype
GOME-2C	2019 - present	80 x 40 km^2	09:30 LT	1 day	AC-SAF prototype

4.2.2 Data gridding approach

Our level-3 (L3) processing for morning sensor NO_2 utilizes a HARP data harmonization toolkit (https://atmospherictoolbox.org/harp/), developed by the European Space Agency for handling scientific earth observation data. For the L3 NO_2 product from morning sensors, only forward-scan pixels are used in the gridding process. Cloudy measurements, where most of tropospheric NO_2 resides below the cloud layer, are excluded since satellite observations could not detect NO_2 columns beneath clouds, resulting in significant uncertainties. Therefore, these cloudy measurements (cloud radiance fraction > 0.5) are not included in the L3 data processing. However, for stratospheric NO_2 columns, no cloud filtering is applied. Additionally, low-quality data with high spectral fitting errors are excluded from the L3 processing to ensure the accuracy of the resulting data products.

Morning sensors, including GOME, SCIAMACHY, and GOME-2, have larger ground pixel sizes compared to afternoon sensors such as OMI and TROPOMI. This can lead to significant grid effects if each measurement is assigned to a single grid cell based solely on the centre coordinates of the ground pixel, without accounting for the pixel geometry and extension. To address this, the gridding process considers the overlapping area between the satellite ground pixel and the latitude-longitude grid. For grid cells that partially overlap with a satellite pixel, a weighting factor is applied based on the ratio of overlap (a satellite pixel fully covering an entire grid cell is considered to have 100 % overlap). The

weighting and count vectors are used to apply this weighted average consistently on the L3 spatial gridding to calculate the mean value and uncertainty as Eq. (4.2.1).

$$x_g = \frac{\sum_{i=1}^{N_{eff}} x_i \times w_i}{\sum_{i=1}^{N_{eff}} w_i} \quad (4.2\text{-}1)$$

where x_g is the gridded column, while x_i represents each individual measurement overlapping within the grid cell. The weighting is denoted w, which is the percentage of the grid cell covered by the satellite pixel, and N_{eff} is the effective population size, determined by the number of available observations of sensors within the L3 grid resolution.

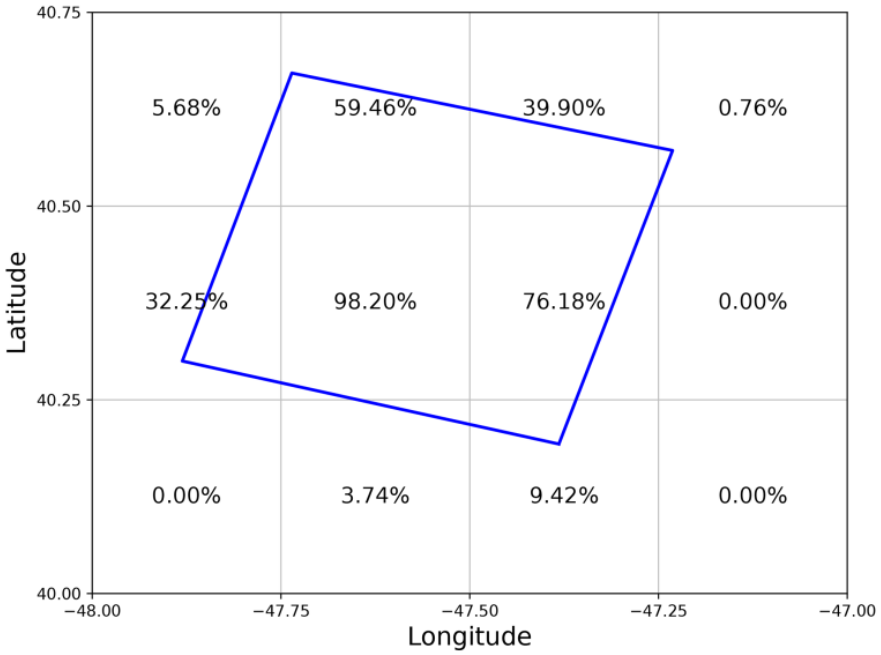


Figure 4.2.1 An example of GOME-2A ground pixel (blue) overlaid on a 0.25°× 0.25° latitude-longitude grid (grey). The percentage of overlap (weighting, w) for each grid box is denoted (image credit: Chan et al., 2023).

Selecting an appropriate resolution for the latitude-longitude grid is essential for processing the L3 products. The L3 product should retain the original spatial features captured in the Level-2 data, which typically has higher spatial resolution. However, it is also important to maintain a reasonable file size to ensure the data remains user-friendly. To select the best spatial resolution for the L3 product, we have analysed the binning results with different resolutions. We investigated the spatial smoothing/averaging effect over polluted regions such as Northeast Asia, as this region is expected to show strong spatial gradients of tropospheric pollutants.

For GOME-2, the results show that gridding data at a higher spatial resolution better preserves the original instrument footprint, while a rather strong smoothing/averaging effect is observed when data is gridded at a lower spatial resolution (i.e. 0.5° x 0.5°, 1° x 1°). While gridded data at a 0.2° x 0.2° resolution exhibit some smoothing/averaging effects, they capture spatial variations reasonably well. The difference between data gridded at 0.1° x 0.1° and 0.2° x 0.2° resolution show small differences. However, 0.5° x 0.5° resolution data show a relatively underestimation of tropospheric NO_2 columns over polluted hotspots. The comparison of GOME-2 data gridded at different spatial sampling resolutions indicates that 0.2° x 0.2° resolution is a balance, preserve the satellite resolution as well as capturing the spatial variations of NO_2. In addition, the absolute values of L3 data show consistent with

the L2 data. Therefore, we concluded that 0.2° x 0.2° resolution is a suitable choice for GOME-2 L3 products.

The consistency of the L3 product across the three GOME-2 sensors is investigated. Figure 4.2.2 presents the monthly mean time series of tropospheric NO2 vertical column densities for three different latitudinal bands over the period from 2007 to 2023. Although small discrepancies among three GOME-2 instruments are observed, primarily due to differences in the instrument characteristics, level-1 calibration, and sensor degradation, they generally exhibit good agreement and consistency during overlapping periods as well as in the overall temporal trends.

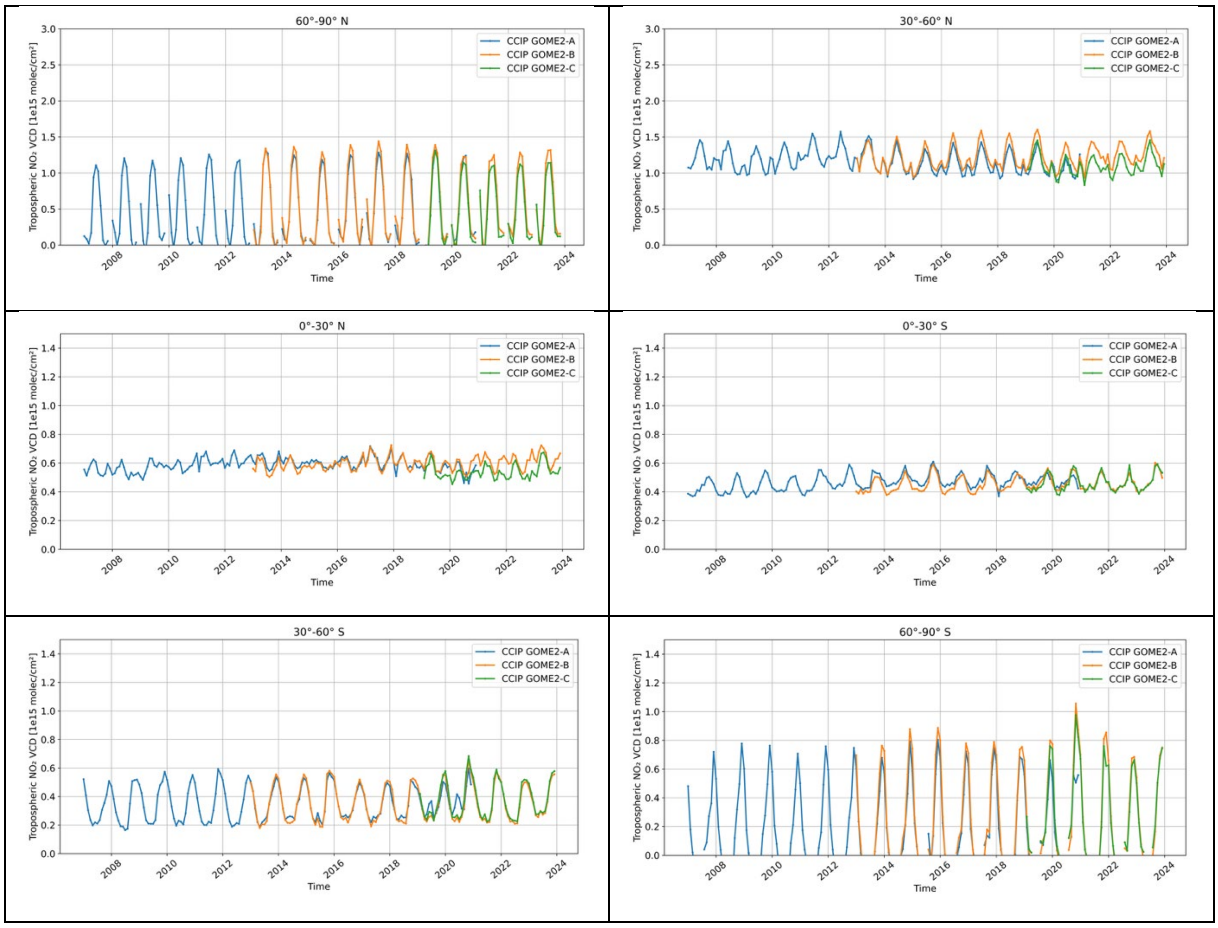


Figure 4.2.2 Time series of monthly zonal averages of tropospheric NO_2 columns from GOME-2A, GOME-2B and GOME-2C across different latitudinal bands from 2007 to 2023.

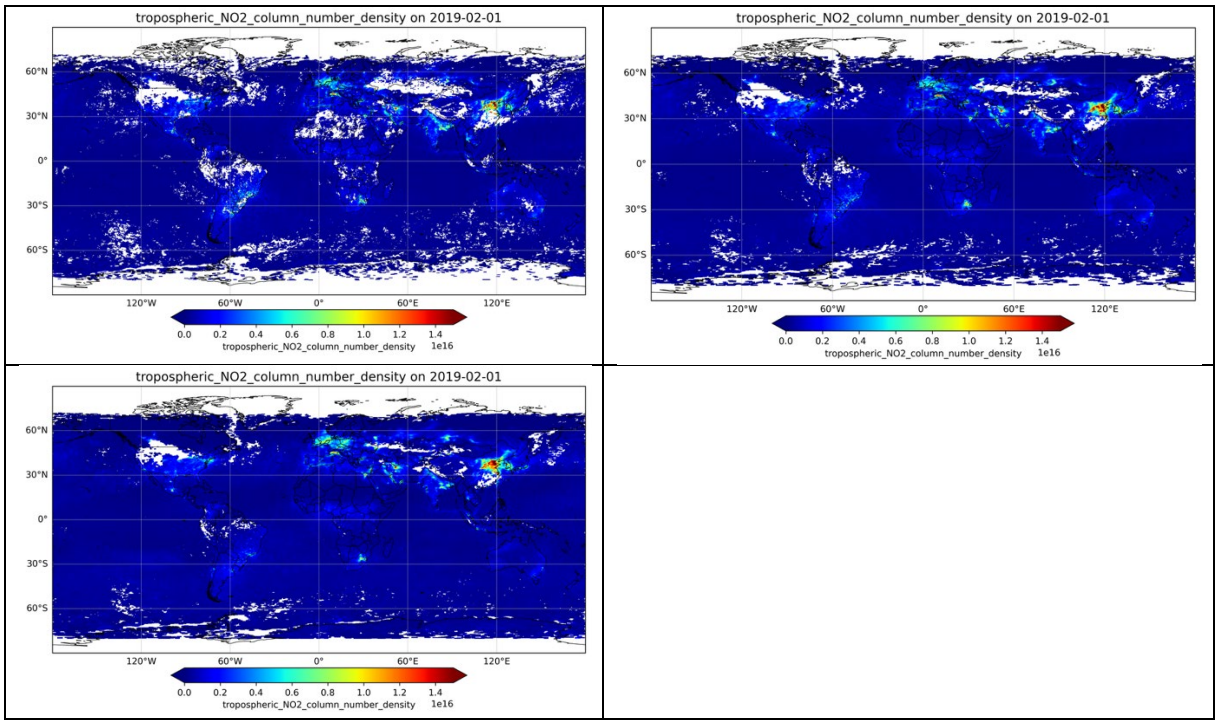


Figure 4.2.3 Monthly averaged tropospheric NO2 columns at 0.2°× 0.2° resolution from Level-3 products of GOME-2A, GOME-2B and GOME-2 for February 2019.

The global maps of monthly level-3 tropospheric NO_2 columns from GOME-2A, GOME-2B, and GOME-2C for February 2019 are presented in Fig. 4.2.3. Overall, the spatial distributions and values of tropospheric NO_2 columns from the three sensors show good agreement. However, GOME-2A shows more missing data and higher uncertainties compared to GOME-2B and GOME-2C. This is mainly attributed to less spatial averaging and instrument degradation. The swath width of GOME-2A was reduced from 1920 km to 960 km in July 2013, while the swath widths of GOME-2B and GOME-2C remain unchanged. Additionally, the noise levels of GOME-2A is higher than those of GOME-2B and GOME-2C due to instrument degradation.

For SCIAMACHY, which has a nadir pixel size of 30 x 60 km^2, comparable to GOME-2 (40 x 80 km^2), L3 products on 0.5° x 0.5° resolution grid show spatial distributions and values similar to those of L2 products. Therefore, 0.5° x 0.5° resolution was selected for SCIAMACHY L3 products, following the same gridding approach as for GOME-2 to ensure consistency. However, the difference in global coverage time between SCIAMACHY and GOME-2 should be noted. While SCIAMACHY takes 6 days to achieve global coverage, GOME-2 provides daily global measurements. This discrepancy in global coverage time leads to differences in data sampling for the monthly L3 products during their overlapping period, resulting in slight variations in the NO_2 climate data records between two instruments.

GOME overlaps operationally with SCIAMACHY for ~ 1 year (August 2002-June 2003), although there is a gap of 3.5 years between GOME and GOME-2. Both GOME and SCIAMACHY L2 datasets from QA4ECV v1.1 are derived using calibrated level-1 data and state-of-the art retrieval algorithms with consistent ancillary data inputs. However, since GOME has a relatively low spatial resolution of 320 x 40 km^2 compared to SCIAMACHY and GOME-2, its L3 spatial resolution should be carefully evaluated to ensure creation of consistent NO_2 climate datasets. To assess the impact of the differences in the ground pixel area between GOME and SCIAMACHY measurements, we compared the original GOME

and SCIAMACHY data with spatially down-sampled SCIAMACHY measurements, using only the data points collected by both instruments on the same day. The spatial resolution of GOME approximated using SCIAMACHY measurements corresponds to an average of approximately five adjacent pixels. We explicitly corrected for the difference in ground pixel size between GOME and SCIAMACHY by applying the appropriate ratio grid of correction factors.

To evaluate the consistency and long-term trends of tropospheric NO_2 columns from all morning sensors, including GOME, SCIAMACHY, and three GOME-2 instruments, we analyzed time series covering the entire measurement period for six regions of interest worldwide using L3 NO2 products (Fig. 4.2.4). The selected regions are eastern China (25-45°N, 105-125°E), the Pacific Ocean (30°S-30°N, 180-21°E), biomass burning regions in northern Africa (15°S-0°N, 10-35 °E), western Europe (35-50 °N, 10 °W- 20 °E), the eastern coast of the United States (35-40 °N, 275-290 °E), and the Middle East (28-38 °N, 34-60 °E). Overall, tropospheric NO_2 columns from all morning sensors L3-NO2 CDR show a good level of agreement in monthly averaged across most regions, confirming their suitability for long-term ternd analysis.

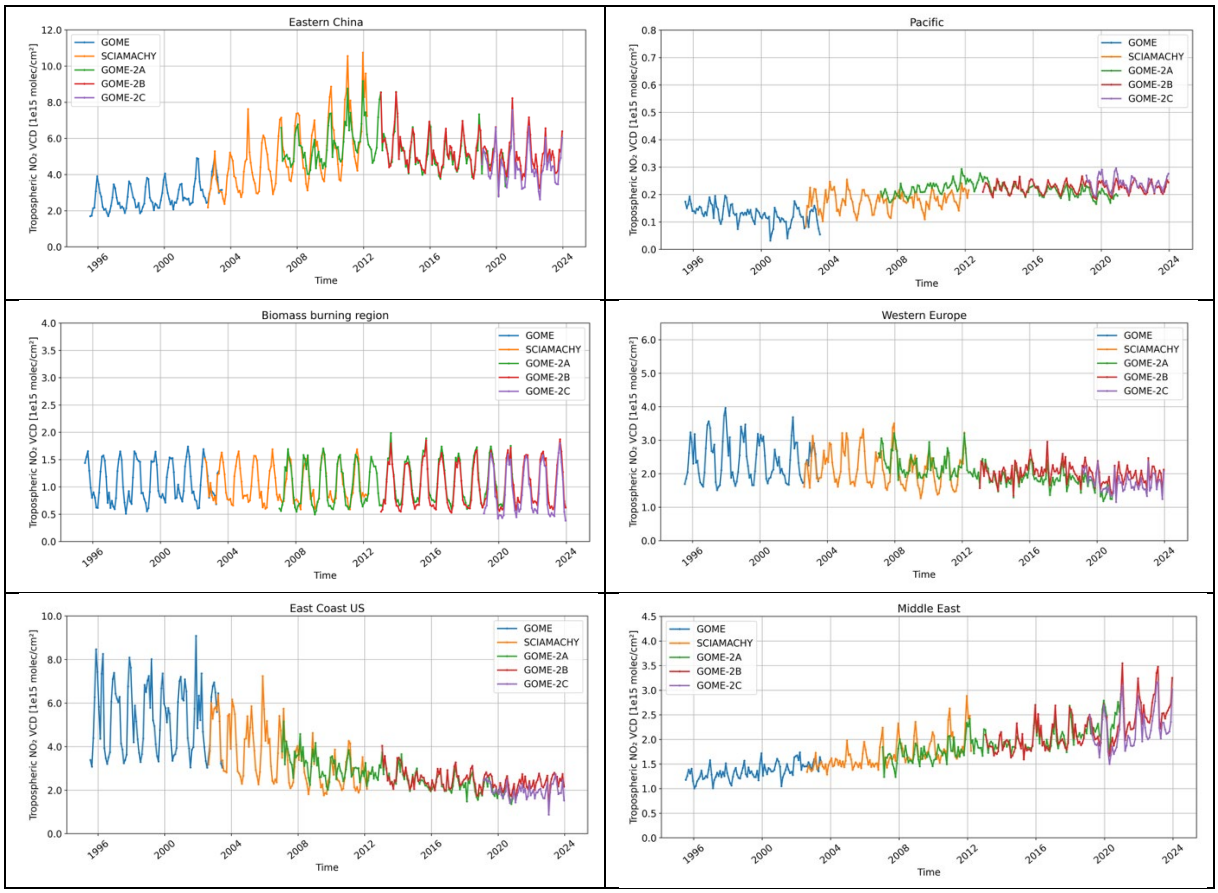


Figure 4.2.4 Time series of monthly average tropospheric NO2 VCDs from all morning sensors (GOME/ERS-2, SCIAMACHY/ENVISAT, and GOME-2/MetOp-ABC) L3 NO2 products, covering six source regions from 1995 to 2023.

4.2.3 L3 uncertainty budget estimate and data quality for GOME-2A

The uncertainty in the retrieved tropospheric NO_2 columns consists of two components: random and systematic errors. The random uncertainties are derived from the slant column errors divided by the air mass factors. For individual observations, random errors dominate the total uncertainties in pristine

and unpolluted regions and are largely affected by the instrument level-1 noise. The systematic uncertainties on tropospheric NO_2 columns are related to the main retrieval steps, including slant column retrievals, stratospheric column estimates, and tropospheric AMF calculations. These uncertainties are quantified in the L2 NO_2 products based on error propagation method (Boersma et al., 2004).

The slant column uncertainties for GOME-2 sensors are estimated using a statistical approach based on spatial variability in SCDs over a clean Pacific region, as described in Zara et al. (2018). These uncertainties range from 0.5 to 0.6 x 10^{15} molec.cm^{-2} (Table 3.1.1). The uncertainty in the stratospheric NO_2 column is 0.2 x 10^{15} molec. cm^{-2} (Table 3.1.2). The tropospheric AMF uncertainty, which is the typically dominant error source in polluted conditions, depends on several parameters, including surface albedo, the a priori NO_2 profile shape, cloud fraction and cloud pressure. The assumed uncertainties for each parameter are summarized in Table 3.1.3, and their contribution to the AMF uncertainty are calculated using the propagation equation (Eq. 3.1-3). A comprehensive uncertainty budget for the tropospheric NO_2 retrieval is provided in Table 3.1.4.

For L3 uncertainty calculations, we keep the existing behaviour for total uncertainty. Averaged random (σ_r) and systematic (σ_s) uncertainties are provided for each grid cell. The total uncertainties (σ) can be obtained from the square root of the sum of squares of the random and systematic error components:

$$\sigma = \sqrt{\sigma_r^2 + \sigma_s^2} \quad (4.2\text{-}2)$$

Here, the systematic uncertainty components (σ_S) include slant column retrieval uncertainties (σ_{SC}) (e.g. spectral calibration, cross section spectra errors), estimated stratospheric column errors (σ_{strat}), and systematic errors in the tropospheric AMF (σ_{AMF}).

$$\sigma_S = \sqrt{(\sigma_{SC})^2 + (\sigma_{strat})^2 + (\sigma_{AMF})^2} \quad (4.2\text{-}3)$$

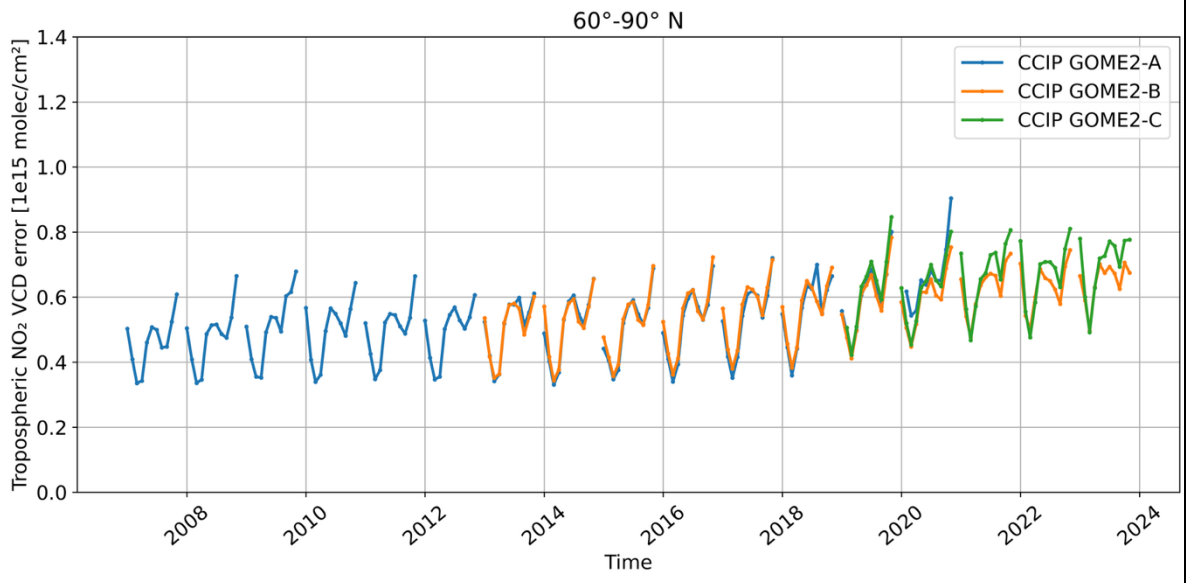
The uncertainty of gridded column variables (σ_g) can this be expressed as:

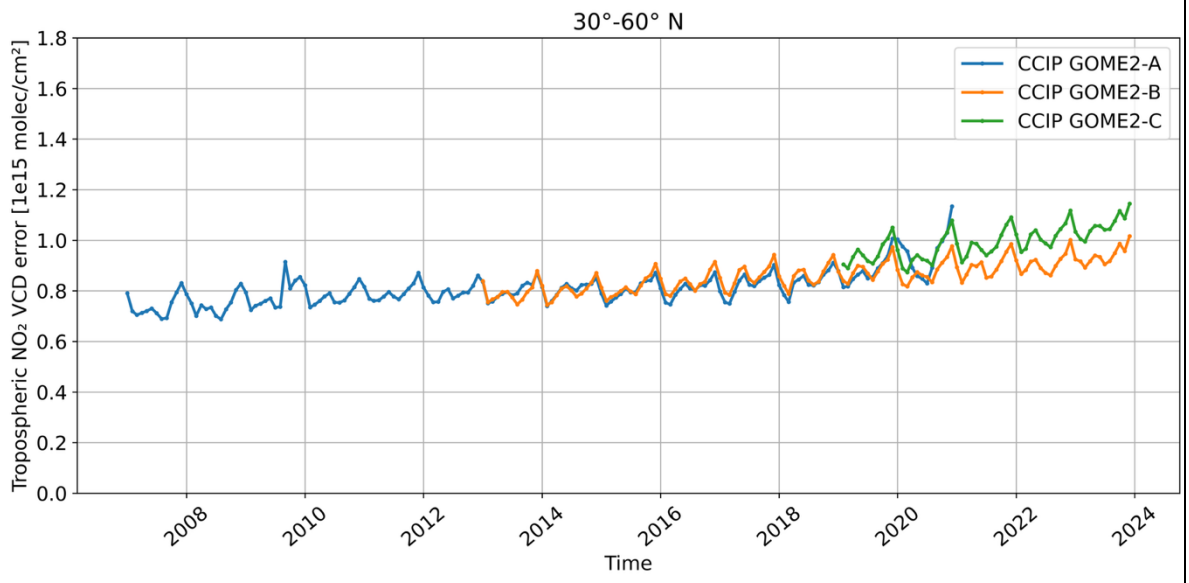
$$\sigma_g = \sqrt{\frac{\sum_{i=1}^{N_{eff}} \sigma_i^2 \times w_i^2}{\sum_{i=1}^{N_{eff}} w_i^2}} \quad (4.2\text{-}4)$$

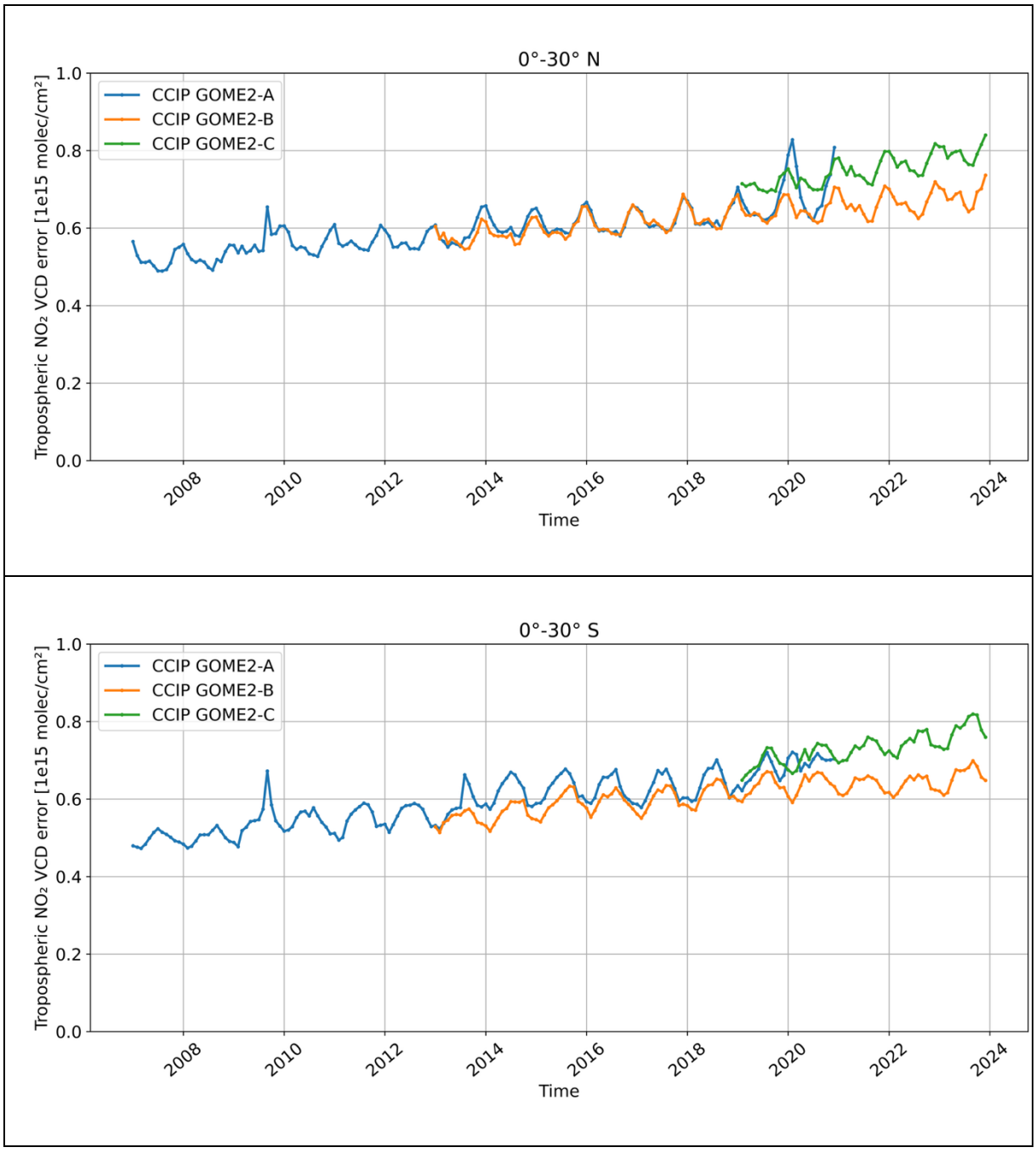
Where is the weighting factor representing the fractional coverage of the grid cell by the satellite pixel and is the effective number of observations within the grid cell, based on the number of available measurements from sensors within the L3 resolution. This approach for calculating uncertain ties is consistently applied for morning sensors.

Figure 4.2.5 shows the time series of zonal mean tropospheric NO_2 column uncertainties for different latitudinal bands from GOME-2 L3 NO_2 products. Overall, all three GOME-2 instruments exhibit an increasing trend in NO_2 column uncertainties over time, primarily driven by random errors from the propagation of measurement noise and increased spectral fitting errors. For GOME-2A, a relatively rapid increase in NO_2 uncertainties is observed until the second throughput test in September 2009. After the troughput test, GOME-2A NO_2 slant column uncertainties increase at a slower pace and remain relatively stable until 2019. However, from 2020 until the end of the mission, GOME-2A tropospheric NO_2 column uncertainties increase sharply, primarily due to a rise in random errors associated with the degradation of the level-1 data, along with a systematic contribution from degradation in the cloud products. Additionally, GOME-2C exhibits higher overall

NO_2 uncertainties during its initial operational phase compared to GOME-2A and GOME-2B, indicating higher random errors originating from relatively higher instrument noise as well as spectral fitting errors due to the use of a narrower spectrl fitting window.






0°-30° N
CCIP GOME2-A
CCIP GOME2-B
CCIP GOME2-C
Tropospheric NO₂ VCD error [1e15 molec/cm²]
1.0
0.8
0.6
0.4
0.2
0.0
2008
2010
2012
2014
2016
2018
2020
2022
2024
Time
0°-30° S
CCIP GOME2-A
CCIP GOME2-B
CCIP GOME2-C
Tropospheric NO₂ VCD error [1e15 molec/cm²]
1.0
0.8
0.6
0.4
0.2
0.0
2008
2010
2012
2014
2016
2018
2020
2022
2024
Time

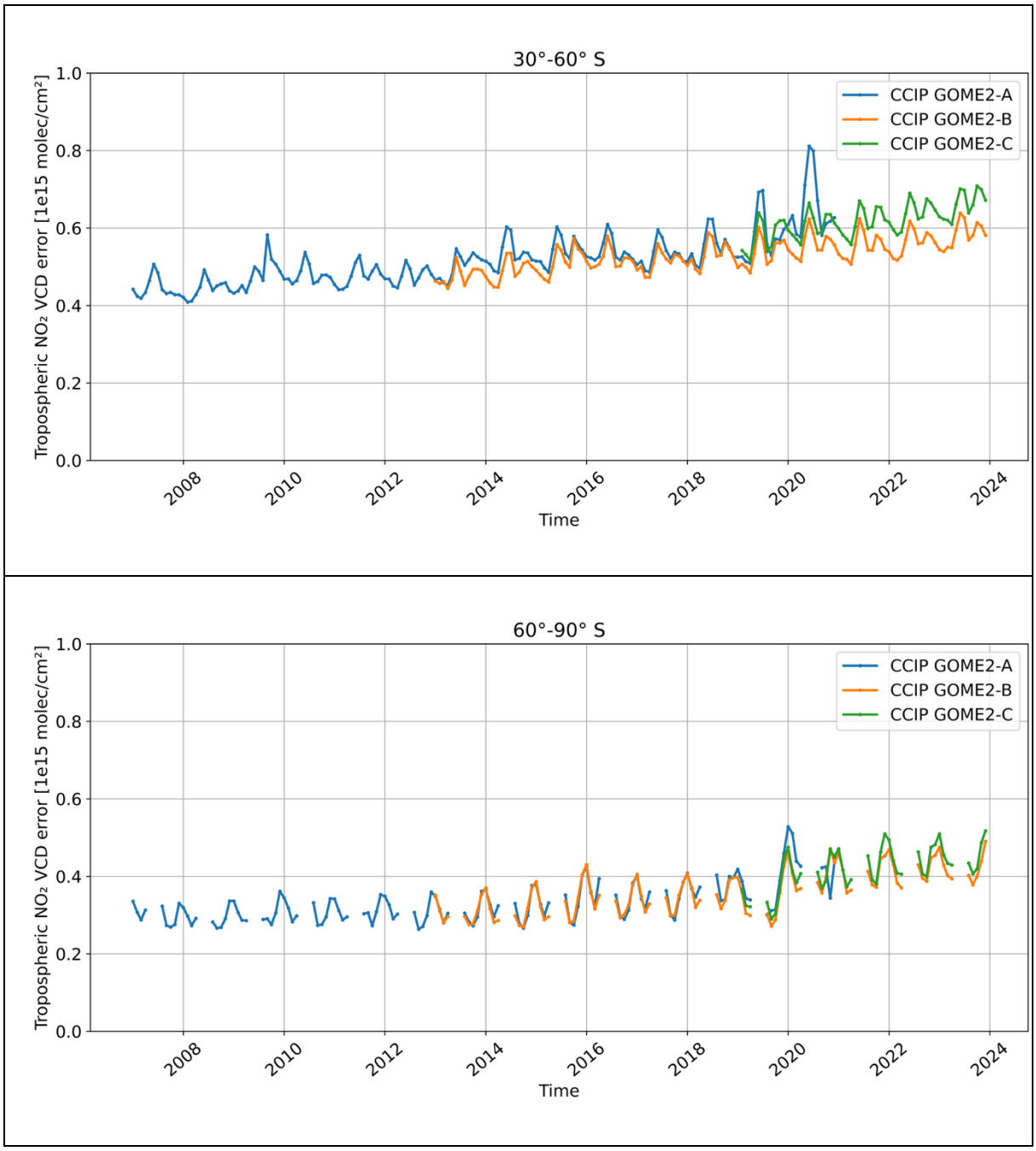


Figure 4.2.5 Time series of monthly zonal averages of tropospheric NO_2 column uncertainties (σ_{total}) from GOME-2A, GOME-2B, and GOME-2C across different latitudinal bands from 2007 to 2023.

The spatial distribution of tropospheric NO_2 column uncertainties for each GOME-2 sensor is also investigated, comparing the early operational period with the most recent year (Fig. 4.2.6). During the initial period, uncertainties are generally low but relatively higher over South America and polluted regions such as East Asia, western Europe, and the eastern United States. For polluted regions, tropospheric NO_2 column uncertainties largely consist of systematic error components derived by the tropospheric AMF calculation, which is dependent on combined effect of errors in surface albedo, a priori NO_2 profile, cloud fraction, and cloud pressure. On the other hand, elevated uncertainties over South America are strongly related to spectral fitting errors. This region is affected by the Southern

Atlantic Anomaly (SAA), where an anomaly in the Earth's magnetic field leads to enhanced radiation exposure of the satellite instruments (Richter et al., 2011). The influence of the SAA results in increased random errors due to elevated levels of measurement noise and consequent higher fitting errors.

Furthermore, a comparison between the initial and the most recent operational years reveals a global increase in tropospheric NO_2 column uncertainties across all GOME-2 sensors. This increase is particularly pronounced over the SAA region due to deterioration in spectral fitting quality with increasing propagation of measurement noise over time.

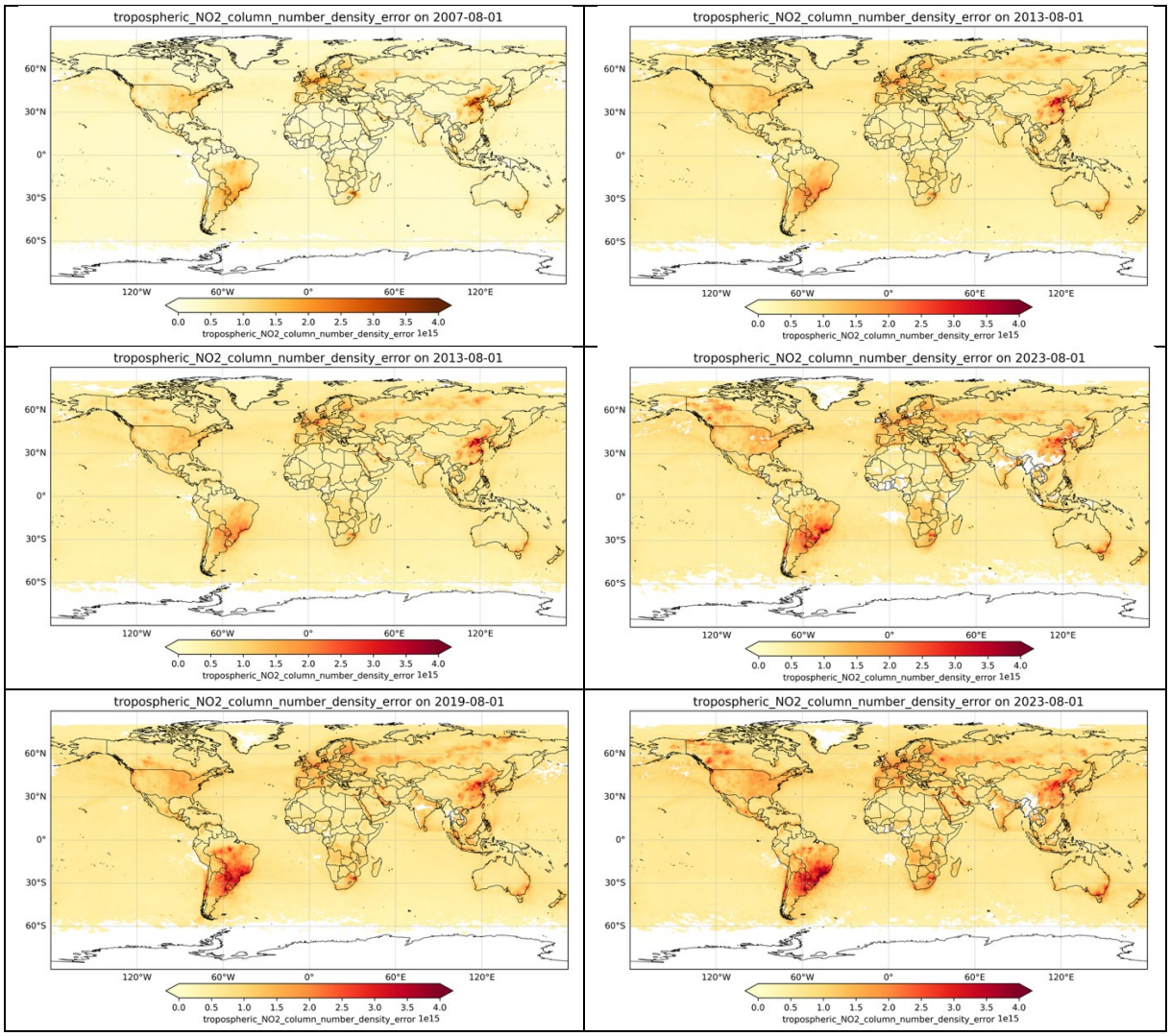


Figure 4.2.6 Monthly averaged tropospheric NO2 column uncertainty (σ_{total}) maps for GOME-2A (1st row), GOME-2B (2nd row) and GOME-2C (3rd row) in August of their respective initial operational year (left column) and recent year (right column).

Table 4.2.2 presents example values of monthly mean tropospheric NO_2 columns and their uncertainties in eastern Asia (30-40 °N, 110-120 °E) for January and July in 2007 and 2020 from GOME-2A L3 products. Tropospheric NO_2 columns are generally higher in 2007 compared to 2020, a year affected by reduced anthropogenic activity due to COVID-19 lockdowns. A clear seasonal pattern is also evident, with higher NO_2 levels observed in winter and lower values in summer. In the initial

operational year (2007), the total uncertainties in tropospheric NO_2 columns accounted for 25.&% (January) and 27.5% (July) of the column values. These relative uncertainties increased to 38.2% (January) and 33.7% (July) in 2020, primarily due to increased random errors from propagated measurement noise, as well as enhanced contributions from systematic uncertainties, including slant column retrieval errors.

Table 4.2.2 Example values for monthly mean tropospheric NO_2 columns (χ) and uncertainties (σ), including the relative uncertainty over eastern China for January and July in the initial ioperational year (2007) and a recent year (2020) from GOME-2A L3 products (units: 1015 molec. cm^{-2}).

	January 2007	July 2007	January 2020	July 2020
χ	11.84	7.83	7.99	5.36
σ	3.04 (25.7%)	2.15 (27.5%)	3.05 (38.2%)	1.81 (33.7%)

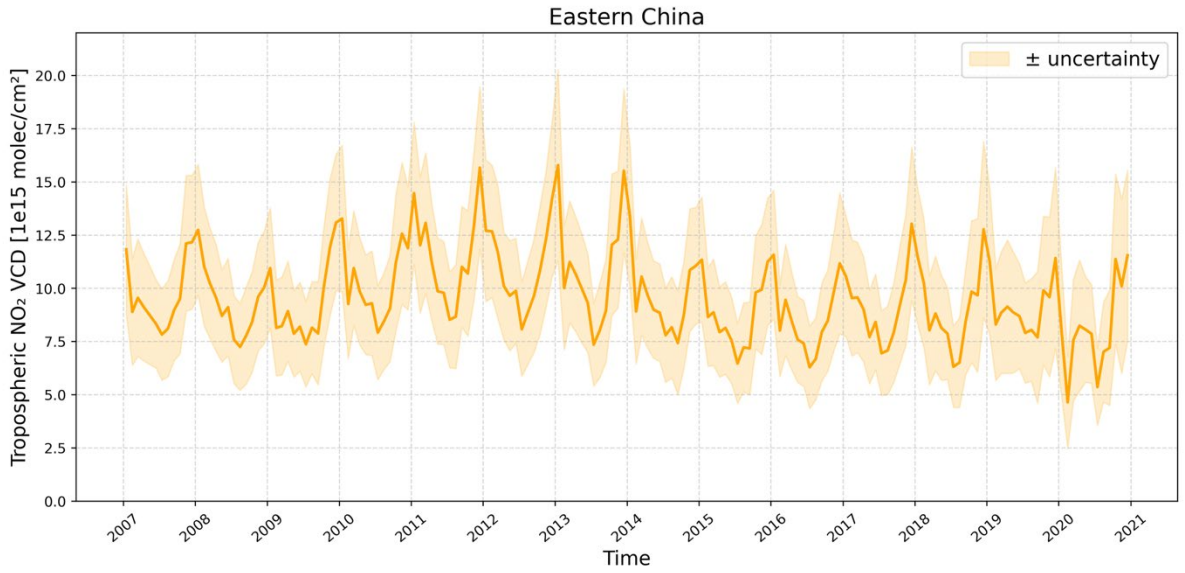


Figure 4.2.7 Time series of the monthly averaged tropospheric NO_2 column with total uncertainty (σ_{total}) over eastern China (region as indicated in Table 4.2.2) from GOME-2A L3 NO_2 products.

A quality flag is provided for the L3 file to filter cells with insufficient observations. It is recommended to use only cells with qa_L3 = 1. The qa_L3 flag is based on the effective number of observations within each grid cell after applying the L2 quality filtering. If the observation count is below the 5th percentile of the cell counts for the day/month, qa_L3 is set to 0.

4.3 Uncertainty assessment for formaldehyde (L3-HCHO)

4.3.1 Random and systematic uncertainties provided in the Level 2 data products

In the L2 HCHO data products, random and systematic uncertainties are provided separately for each individual observation. The uncertainty estimates are based on error propagation. See section 3.2 Formaldehyde (L2-HCHO) of Precursors_cci+_D2.2_E3UB_01_02 for a full description of uncertainties. The random uncertainties are the slant columns errors divided by the air mass factors (first term of equation (3.2-3)). For individual observations, random uncertainties dominate the total uncertainties. Systematic uncertainties are made of systematic uncertainty of the slant columns, air mass factors and background correction. Systematic uncertainties are dominated by uncertainties on model a priori profiles, surface albedo and cloud properties (see Table 3.2-4).

4.3.2 Averaging Procedure

The data gridding method for the HCHO L3 data products is described in the ATBD (4.2. Formaldehyde (L3-HCHO)). The core of the HCHO Level-3 gridding algorithm is based on the HARP atmospheric toolbox (https://atmospherictoolbox.org/harp/) and its spatial binning functionalities. We further develop on HARP to facilitate the gridding of daily and monthly data products from the different L2 products, and to include specific information and statistical estimates.

4.3.2.1 Data filtering

L2 pixels are filtered using the L2 qa_values recommended threshold of 0.5.

4.3.2.1 Spatial binning / gridding

Spatial binning grids the L2 data to a rectilinear latitude/longitude grid, using an area weighted average (see https://stcorp.github.io/harp/doc/html/algorithms/regridding.html). Per day/month, the selected L2 pixels are given a weight proportional to the fractional overlap of the given pixels with the grid cell. The weight determines the contribution of the source pixel *i* to the target grid cell *(j,k)*. With the calculated weights each variable is then regridded using:

$$x_t(j,k) = \frac{\sum_i w(i,j,k) x_s(i)}{\sum_i w(i,j,k)}$$

The effective number of overlapping L2 pixels per grid cell for a given day or month (N_{eff}) is provided in a *count* variable

4.3.3 L3 uncertainty estimate and data quality

4.3.3.1 Random and systematic uncertainties

Averaged random (σ_r) and systematic (σ_s) uncertainties are available for each grid cell. If a total uncertainty is desired, this can be obtained from the square root of the sum of squares of the different terms:

$$\sigma^2 = \sigma_r^2 + \sigma_s^2$$

Individual random uncertainty variables from L2 are averaged using the square of each value. The final value is given by:

$$\sigma_{r,t}(j,k) = \frac{\sqrt{\sum_i \left(w(i,j,k)\sigma_{r,s}(i)\right)^2}}{\sum_i w(i,j,k)}$$

Individual systematic uncertainty variables from L2 are averaged as correlated (i.e. using a regular average).

4.3.3.2 Standard deviation

In addition to uncertainty propagation, we also provide the standard deviation of the key variables within the grid cell, for one day or for one month. The standard deviations provide a first estimate of the spatial and temporal variations. The monthly averages are provided at a finer spatial resolution than the daily averages (0.125° / 0.25°) and are therefore dominated by temporal variability, while the daily averages give an estimate of the spatial variability.

4.3.3.3 L3 data quality flag

A quality flag is provided for the L3 file, with the aim to filter cells containing not enough observations. It is advised to use only qa_L3==1. The qa_L3 flag is based on the effective number of observations within the cell, after applying the L2 quality filtering. If the count N_{eff} is lower than the 5th percentile of the cell counts for the day/month, the qa_L3 is set to 0 (the grid cell is not covered enough).

4.3.1.1 Analysis of L3 uncertainty budget for different sensors

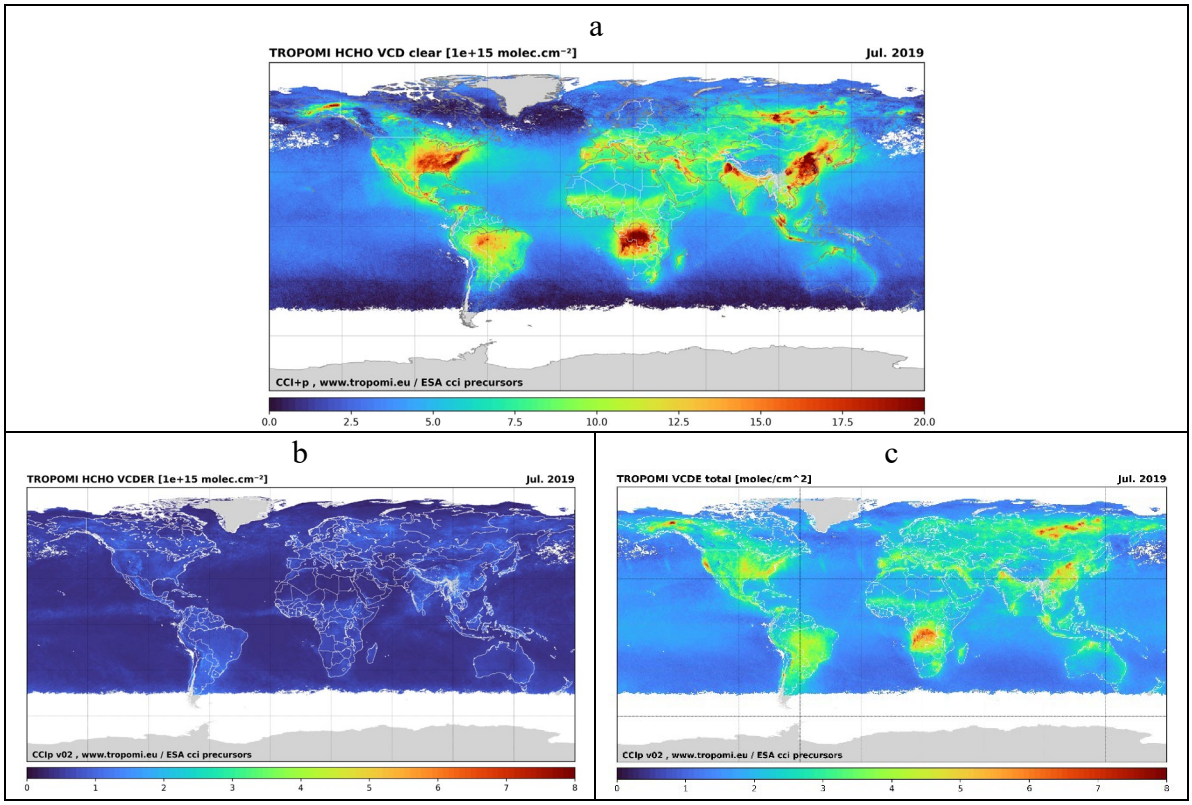
The random uncertainties of the slant columns dominate the other sources of uncertainties for individual measurements. Therefore, the largest contribution to the final L3 uncertainty will depend on the spatial and temporal resolution of grid, and on the instrument. For daily averages, it is dominated by the random uncertainties. For monthly averages, it is dominated by systematic uncertainties on AMFs.

There is no difference between the treatment of uncertainties between the sensors. But the random error is significantly lower for TROPOMI than for other instruments. The number of observations is also significantly larger. The systematic errors are much closer for all the instruments. These differences are illustrated in Figures 4.3.1 and 4.3.2, that compare L3 monthly gridded variables for TROPOMI and OMI (at the same resolution of 0.125°). The differences between the random uncertainty and the effective number of observations appear clearly.

Figure 4.3.3 presents the provided HCHO VCD random uncertainty over the Equatiorial Pacific Region, for the morning sensors (at the resolution of 0.25°) and the afternoon sensors (at the resolution of 0.125°), showing the impact of the different number of observations on the random uncertainties.

Figure 4.3.4 shows some examples of regional averages of HCHO VCD from OMI and TROPOMI L3 datasets. The represented total uncertainties have been calculated with the provided random and systematic uncertainties for each grid cell. For OMI and TROPOMI monthly averages, they are largely dominated by the systematic uncertainties.

TROPOMI



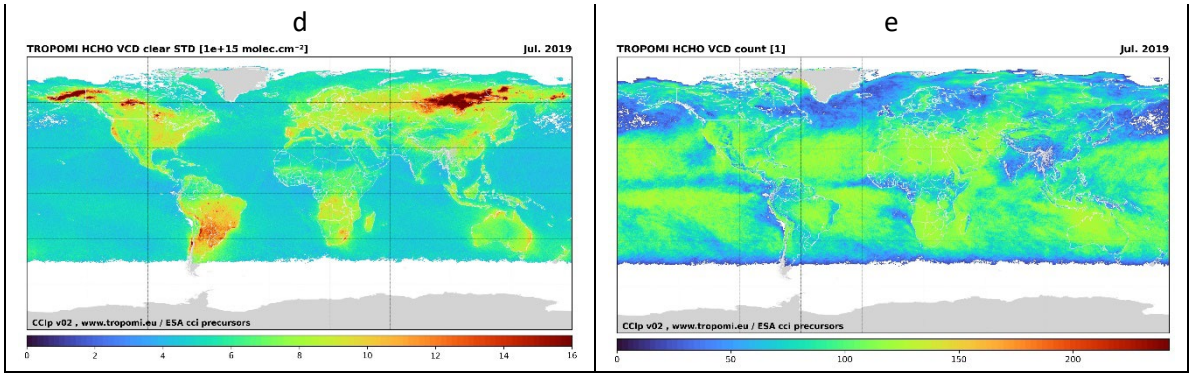


Figure 4.3.1: TROPOMI HCHO L3 gridded variables for July 2019: (a) HCHO VCD clear, (b) random uncertainty, (c) total uncertainty, (d) standard deviation, (e) count: effective number of observations included in the grid cells. For a monthly gridded file at the spatial resolution of 0.125°.

OMI

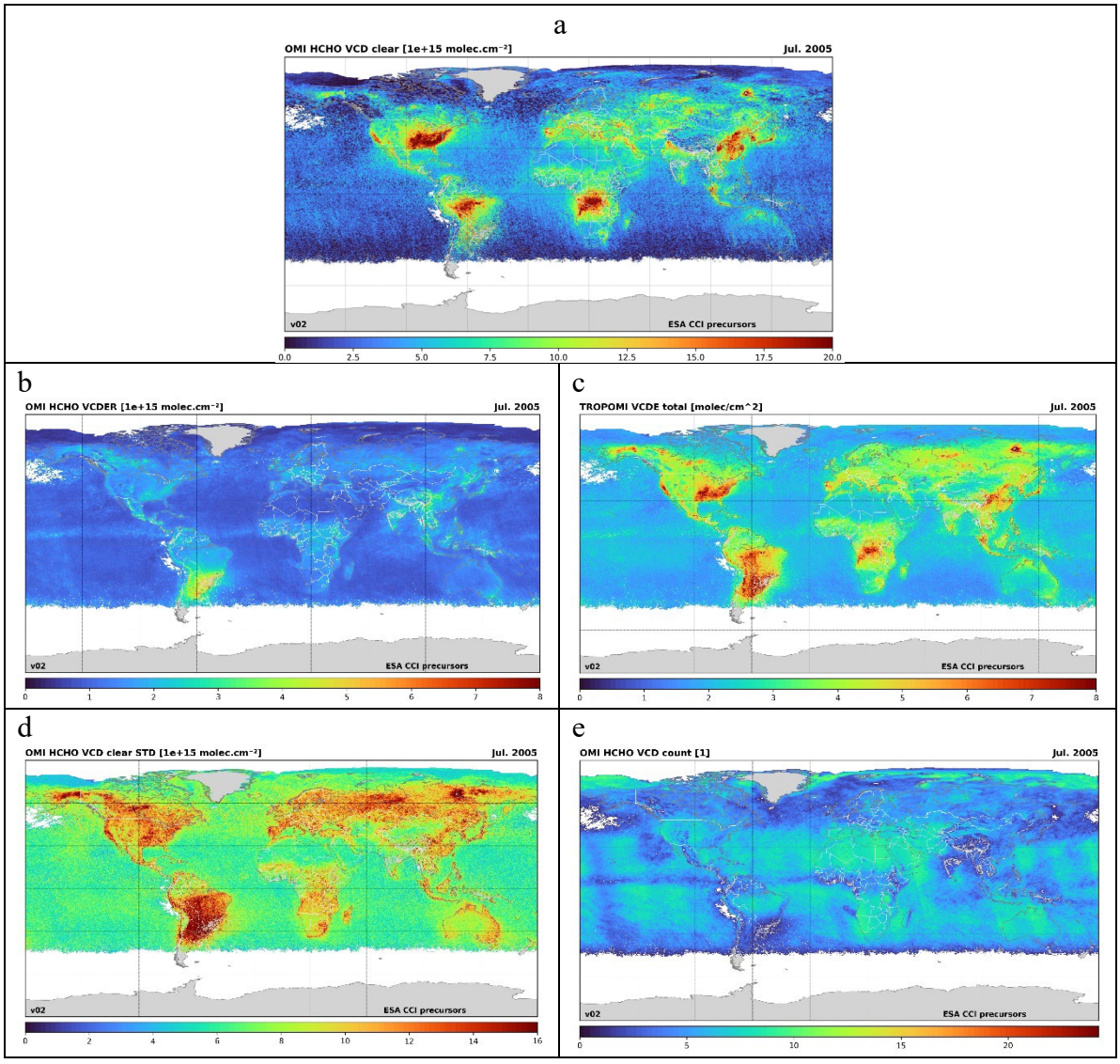


Figure 4.3.2: OMI HCHO L3 gridded variables for July 2005: (a) HCHO VCD clear, (b) random uncertainty, (c) total uncertainty, (d) standard deviation, (e) count: effective number of observations included in the grid cells. For a monthly gridded file at the spatial resolution of 0.125°.

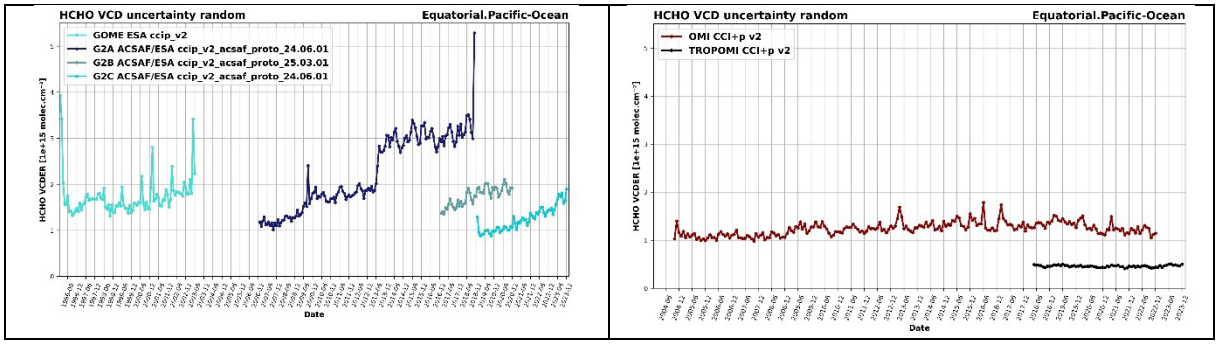
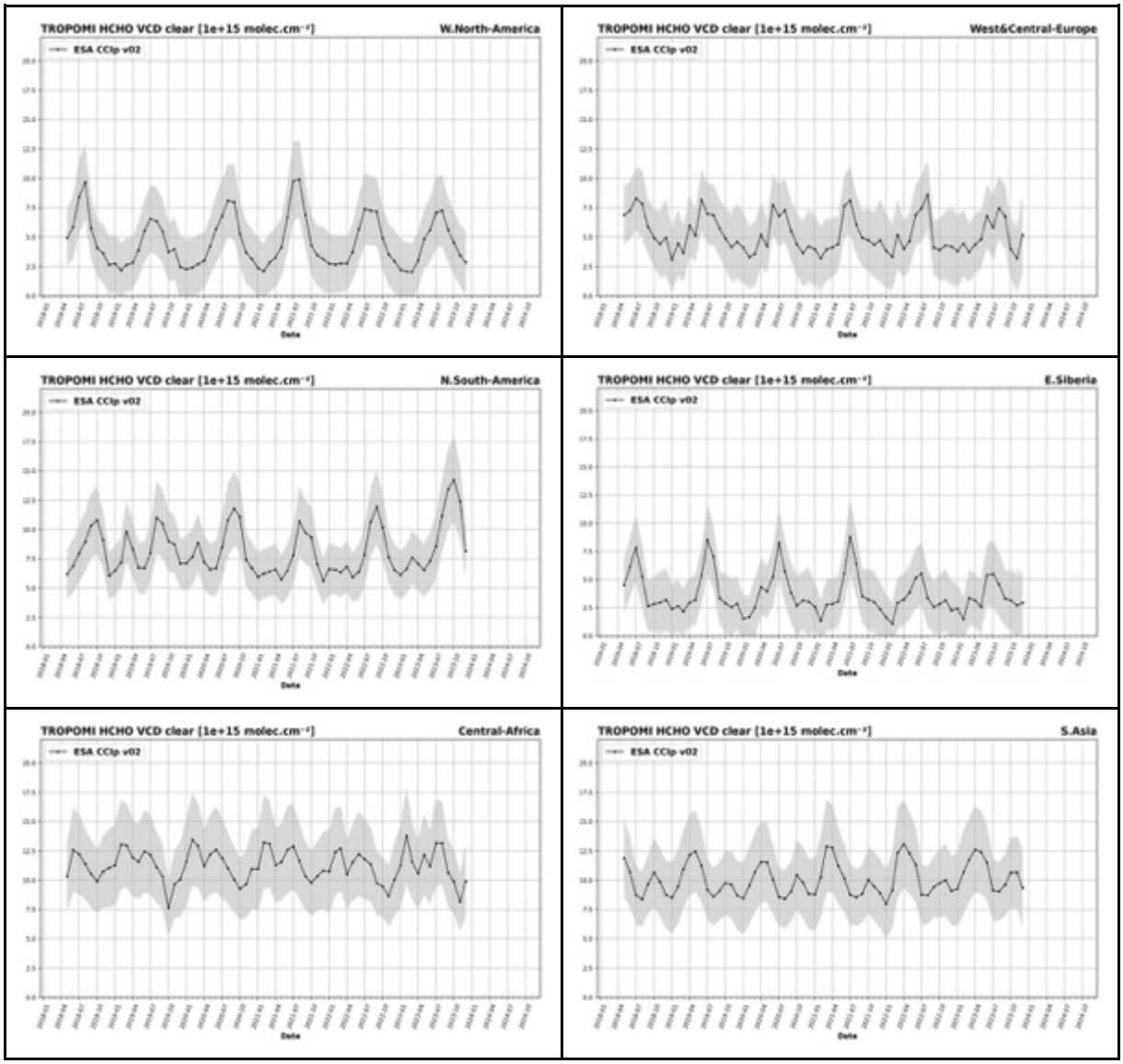


Figure 4.3.3: Example of the HCHO column random uncertainties over the Equatorial Pacific Region as provided in the morning and afternoon monthly L3 CDRs. For monthly gridded files at the spatial resolution of 0.25° (morning) and 0.125° (afternoon).



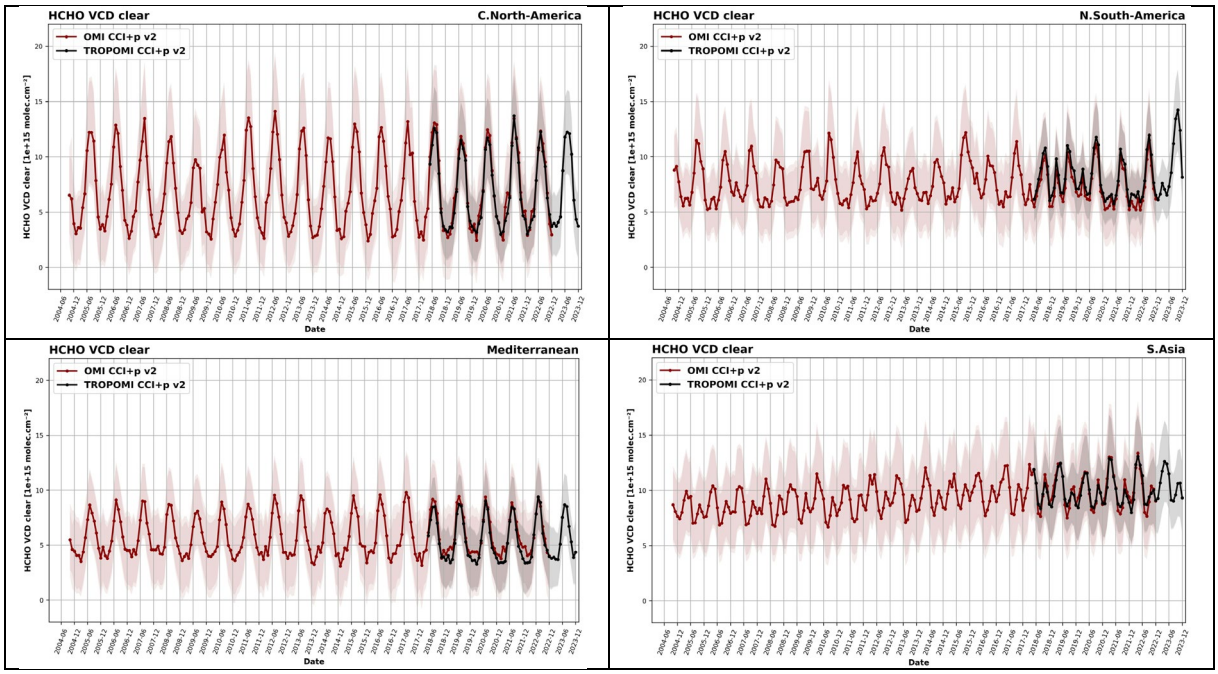


Figure 4.3.4: Regional averages of HCHO VCD calculated using the CCI+p L3 OMI and TROPOMI datasets (averages of cells provided in the monthly grids, at 0.125° resolution). The total uncertainties have been calculated with the provided random and systematic uncertainties for each grid cell. They are largely dominated by the systematic uncertainties.

4.3.4 Recommendations regarding the use of HCHO L3 data

In a general sense, users should be aware that the HCHO L3 data are filtered to keep only valid observations, and therefore the cloudy scenes are not included. Therefore the monthly averages not necessarily representative for all-sky monthly means (that would be obtained from CTM simulations).

HCHO L3 data users:

We recommend using only daily or monthly mean L3 data with qa_L3 == 1 values. These represent monthly means that have been calculated with a sufficient number of valid days of data. The main estimates of the L3 uncertainties are the "tropospheric_HCHO_column_number_density_uncertainty_random" and "tropospheric_HCHO_column_number_density_uncertainty_systematic" that should be combined to obtain the total uncertainty.

For users comparing L3 monthly means with model monthly mean NO_2 data, we recommend using "tropospheric_HCHO_column_number_density_uncertainty_systematic_kernel", which is slightly lower because the contribution from the HCHO profile errors has been removed (See also the PUG for more recommandations).

4.4 Uncertainty assessment for sulphur dioxide (L3-SO2)

4.4.1 Uncertainty provided in the Level 2 data products

In the L2 SO_2 data products, random and systematic uncertainties are provided separately for each individual observation. The uncertainty estimates are based on error propagation. See section 3.3 Sulfur dioxide (L2-SO2) for a full description of uncertainties.

4.4.2 Averaging procedure

The data gridding method for the SO_2 L3 data products is described in the ATBD (4.3. Sulfur dioxide (L3-SO2)). The core of the SO_2 Level-3 gridding algorithm is based on the HARP atmospheric toolbox (https://atmospherictoolbox.org/harp/) and its spatial binning functionalities. We further develop on HARP to facilitate the gridding of daily and monthly data products from the different L2 products, and to include specific information and statistical estimates.

4.4.2.1 Data filtering

L2 pixels are filtered using the L2 qa_values recommended threshold of 0.5. One exception is for total_SO2_column_number_density_box. This variable gathers the SO_2 vertical columns for box profiles at 1, 7, 15 km and is mostly useful to study volcanic plumes. To avoid masking useful data, the vertical columns for box profiles are filtered only for large SZA (>70°).

4.4.2.1 Spatial binning / gridding

Spatial binning grids the L2 data to a rectilinear latitude/longitude grid, using an area weighted average (see https://stcorp.github.io/harp/doc/html/algorithms/regridding.html). Per day/month, the selected L2 pixels are given a weight proportional to the fractional overlap of the given pixels with the grid cell. The weight determines the contribution of the source pixel *i* to the target grid cell *(j,k)*. With the calculated weights each variable is then regridded using:

$$x_t(j,k) = \frac{\sum_i w(i,j,k) x_s(i)}{\sum_i w(i,j,k)}$$

The effective number of overlapping L2 pixels per grid cell for a given day/month is stored in a *count* variable.

4.4.3 L3 uncertainty estimate and data quality

4.4.3.1 Random and systematic uncertainties

Averaged random (σ_r) and systematic (σ_s) uncertainties are available for each grid cell. If a total uncertainty is desired, this can be obtained from the square root of the sum of squares of the different terms:

$$\sigma^2 = \sigma_r^2 \quad + \sigma_s^2$$

Random uncertainty variables are averaged using the square of each value. The final value is given by:

$$\sigma_{r,t}(j,k) = \frac{\sqrt{\sum_i \left(w(i,j,k)\sigma_{r,s}(i)\right)^2}}{\sum_i w(i,j,k)}$$

Systematic uncertainty variables are averaged as correlated (i.e. using a regular average).

4.4.3.1 Standard deviation

In addition to uncertainty propagation, we also provide the standard deviation of the key variables within the grid cell, for one day or for one month.

4.4.3.1 L3 data quality flag

A quality flag is provided for the L3 file, with the aim to filter cells containing not enough observations. It is advised to use only qa_L3==1. The qa_L3 flag is based on the effective number of observations within the cell, after applying the L2 quality filtering. If the count is lower than the 5th percentile of the cell counts for the day/month, the qa_L3 is set to 0.

4.4.4 Examples

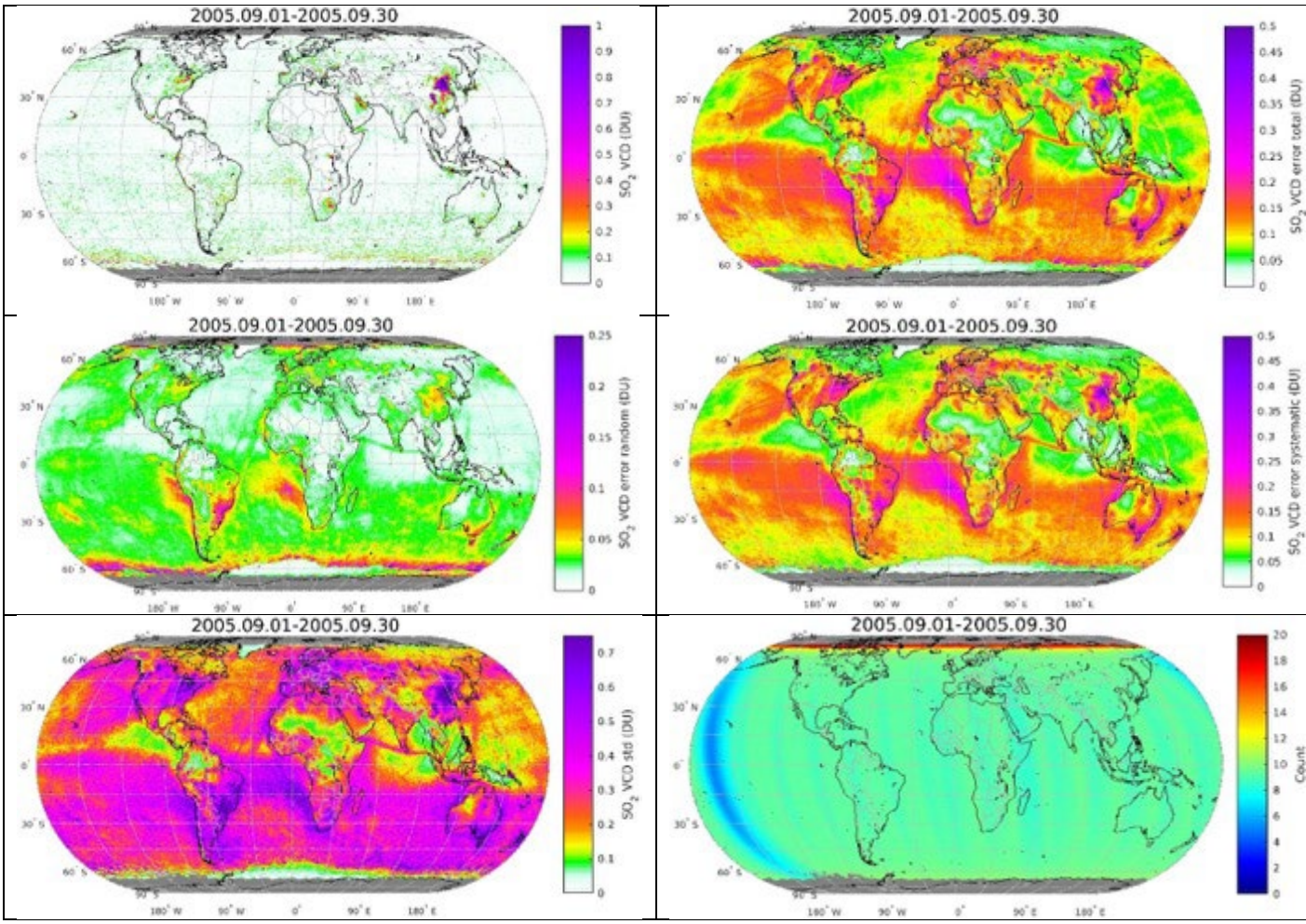


Figure 4.4.1: OMI SO2 L3 gridded variables for September 2005: (a) SO_2 VCD, (b) SO_2 VCD total uncertainty, (c) SO_2 VCD random uncertainty, (d) SO_2 VCD systematic uncertainty, (e) SO_2 VCD standard deviation, (f) SO_2 VCD count: effective number of observations included in the grid cell.

4.5 Uncertainty assessment for glyoxal (L3-CHOCHO)

4.5.1 Procedure

Estimated uncertainties for the L3 glyoxal product follow directly from the estimated uncertainties of the L2 data, and the L3 averaging or gridding procedure. As described in the ATBD (4.4 Glyoxal), L2 data is filtered using the recommended QA value threshold of 0.5. The L3 gridded tropospheric glyoxal vertical column density for a given grid cell and time period is the average of all L2 observations for that time period which pass the QA value filter, where each observation is weighed according to the overlap area of the satellite pixel and the L3 grid cell. These calculations are carried out using the HARP software package's routines for spatial and temporal averaging (bin_spatial and bin) (see https://stcorp.github.io/harp/doc/html/algorithms/regridding.html). The estimated random and systematic uncertainty for the L3 vertical column is then calculated from the L2 estimated uncertainties according to the standard rules of uncertainty propagation. Here, L2 random uncertainties from different observations are assumed to be uncorrelated, while systematic uncertainties are assumed to be fully correlated.

4.5.2 L3 uncertainty estimate and data quality

4.5.2.1 Random and systematic uncertainties

Averaged random (σ_r) and systematic (σ_s) uncertainties are available for each L3 grid cell. The total uncertainty for a grid cell can be obtained from the square root of the sum of squares of the different terms:

$$\sigma^2 = \sigma_r^2 + \sigma_s^2.$$

The random uncertainty for a L3 pixel is calculated from the contributing pixels' random uncertainties $\sigma_{i,r}$ using the equation

$$\sigma_r = \frac{\sqrt{\sum_i \left(w_i \sigma_{i,r}\right)^2}}{\sum_i w_i},$$

where the weights w_i are given by the overlap area of the L3 pixel and L2 pixel i. The systematic uncertainty for a L3 pixel is given by the weighted mean of the contributing L2 systematic uncertainties:

$$\sigma_s = \frac{\sum_i w_i \sigma_{i,s}}{\sum_i w_i}.$$

4.5.2.2 Standard Deviation

In addition to the propagated L2 uncertainties, we also provide the standard deviation of the key variables within the grid cell, for one day or for one month.

4.5.2.3 L3 data quality flag

The L3 file contains a quality flag, which can be used to filter out grid cells which contain too few observations. The L3 quality flag is based on the effective number of observations, which is the sum of the ratio of the overlap area w_i and pixel area A_i for the contributing L2 pixels:

$$N_{\text{eff}} = \sum_i \frac{w_i}{A_i}$$

If a L3 grid cell's effective L2 pixel count is lower than the 5th percentile of the counts for that day/month, the qa_L3 flag is set to 0.

4.5.2.4 Examples

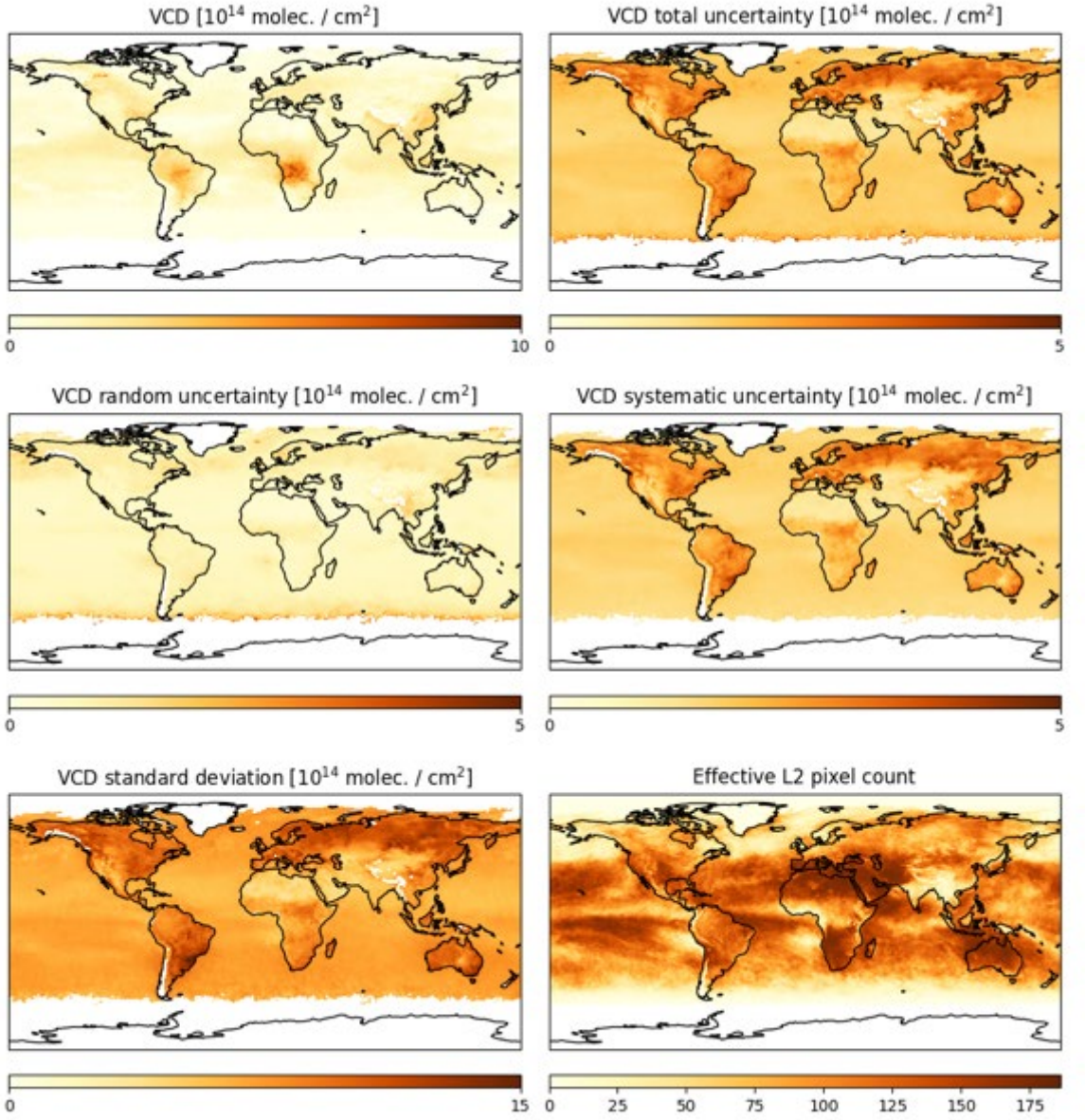


Figure 4.5.1 Tropomi glyoxal L3 gridded variables for August 2023: vertical column density, random, systematic and total uncertainties, standard deviation within a grid cell, and number of contributing L2 pixels per grid cell.

4.6 Uncertainty assessment for carbon monoxide (L3-CO)

4.6.1 Procedure

We describe here how we determine the monthly Climate Data Record (CDR) Level 3 Carbon Monoxide (CO) merged product: total column and uncertainty grids. This new merged product incorporates data from a IASI non-public intermediate multiplatform dataset as well as MOPITT V9T data, processed with a weighted averaging approach. During cycle 1, we generated the IASI non-public intermediate multiplatform dataset averaging cloud free data from Metop-A, Metop-B, and Metop-C. Related uncertainties have been described in detail in Section 3.5.

To generate this new CCI+ L3 merged product, we used IASI and MOPITT monthly L3 data as input. Our approach is different from the other species presented in this document as we do not use L2 data to generate our L3 product but we use two different L3 products, that include CO total column grids as well as uncertainty grids, both for daytime and nighttime separately.

CDR L3 CO merged grids are at 1° x 1° resolution, like MOPITT and IASI L3 CO monthly data.

Here, we present the different steps as well as the formula to generate total column and uncertainty grids for the merged product. Other grids (surface altitude and flag for data source identification) are also provided and described in the ATBD.

4.6.1.1 Calculation of the total column grids

4.6.1.1.1 Calculation of MOPITT total column/prior ratio and weight grids

To account for the strength of CO signature in the MOPITT measurement, the ratio total column/prior is first calculated. This ratio indicates how much the retrieved MOPITT value deviates from its prior estimate. The ratio is compared to predefined thresholds (e.g., 0.9 and 1.1). If the retrieved total column is 10% larger or smaller than the prior, the weight assigned to MOPITT data is increased from 0.5 to 0.65. Otherwise, MOPITT is assigned a neutral weight of 0.5. The IASI weight is then determined as the difference between 1 and the MOPITT weight. Total weight sums to 1. The ratio and weight grids are calculated separately for daytime and nighttime observations.

4.6.1.1.2 Weighted averaging of CO total columns

The merged CO total column for each grid point is calculated as the weighted average of IASI and MOPITT total columns, using the weights previously calculated:

$$X_{merged} = \omega_{MOPITT}.X_{MOPITT} + \omega_{IASI}.X_{IASI}$$

Where :

X_{MOPITT} and X_{IASI} are the CO total column monthly grids used as inputs

ω_{MOPITT} and ω_{IASI} are the weights associated with each grid, and ω_{MOPITT} + ω_{IASI}=1.

4.6.1.2 Uncertainty grids calculation corresponding to the merged CO product

The uncertainty for the merged CO column is calculated using a weighted root-mean-square approach. The formula for the propagated uncertainty in the weighted mean is:

$$\sigma_{merged} = \sqrt{(\omega_{MOPITT}{}^2 . \sigma_{MOPITT}{}^2 + \omega_{IASI}{}^2 . \sigma_{IASI}{}^2)}$$

Where :

- σ_{merged} is the uncertainty associated with the weighted mean X_{merged}.

- σ_{MOPITT} is the uncertainty grid corresponding to the MOPITT CO total column grid. It correspond to the total uncertainty of the MOPITT CO total column grid, obtained by summing the two dimensions of the « RetrievedCOTotalColumnDiagnostics » variable in the monthly L3 V9T HDF5 files available on the NASA Earthdata portal (Worden et al., 2022).

- σ_{IASI} is the uncertainty grid corresponding to the IASI intermediate (non public) multiplatform CO total column grid.

- ω_{MOPITT} and ω_{IASI} are the weights associated with each grid.

σ_{IASI} is given by :

$$\sigma_{IASI}(i,j) = \sqrt{\frac{1}{\sum_l \frac{1}{CO_total_column_uncertainty_l^2}}}$$

for each grid cell (*i,j*) (*i=1,…,360;j=1,…180*), with CO_total_column_uncertainty the total retrieval uncertainty given by FORLI-CO. In Rodgers' formalism, the term 'error' is used synonymously with 'uncertainty'. Hence, users are provided with a 'Total Error' associated with the retrieved CO total column for each pixel when using IASI FORLI-CO Level 2 data (see Section 3.5). We refer to the ATBD for the description of the handling of missing data.

4.6.2 Analysis of the L3 uncertainty budget

Figure 4.5.1 provides an example for the month of July 2008, daytime, Metop-A. We present total uncertainties for MOPITT v9T (top), IASI (middle) and the newly merged CCI+ CO product (bottom).

On a global scale, mean uncertainty for daytime merged product is 1.1 10^{17} molecules/cm^2 with a standard deviation of 0.7 10^{17} molecules/cm^2. For nighttime, the mean uncertainty is slightly larger: 1.3 10^{17} molecules/cm^2 [std: 0.7 10^{17} molecules/cm^2]. CO uncertainty time series for the merged product, as well as for IASI and MOPITT V9T are shown in Fig. 4.5.2 for three different latitude bands.

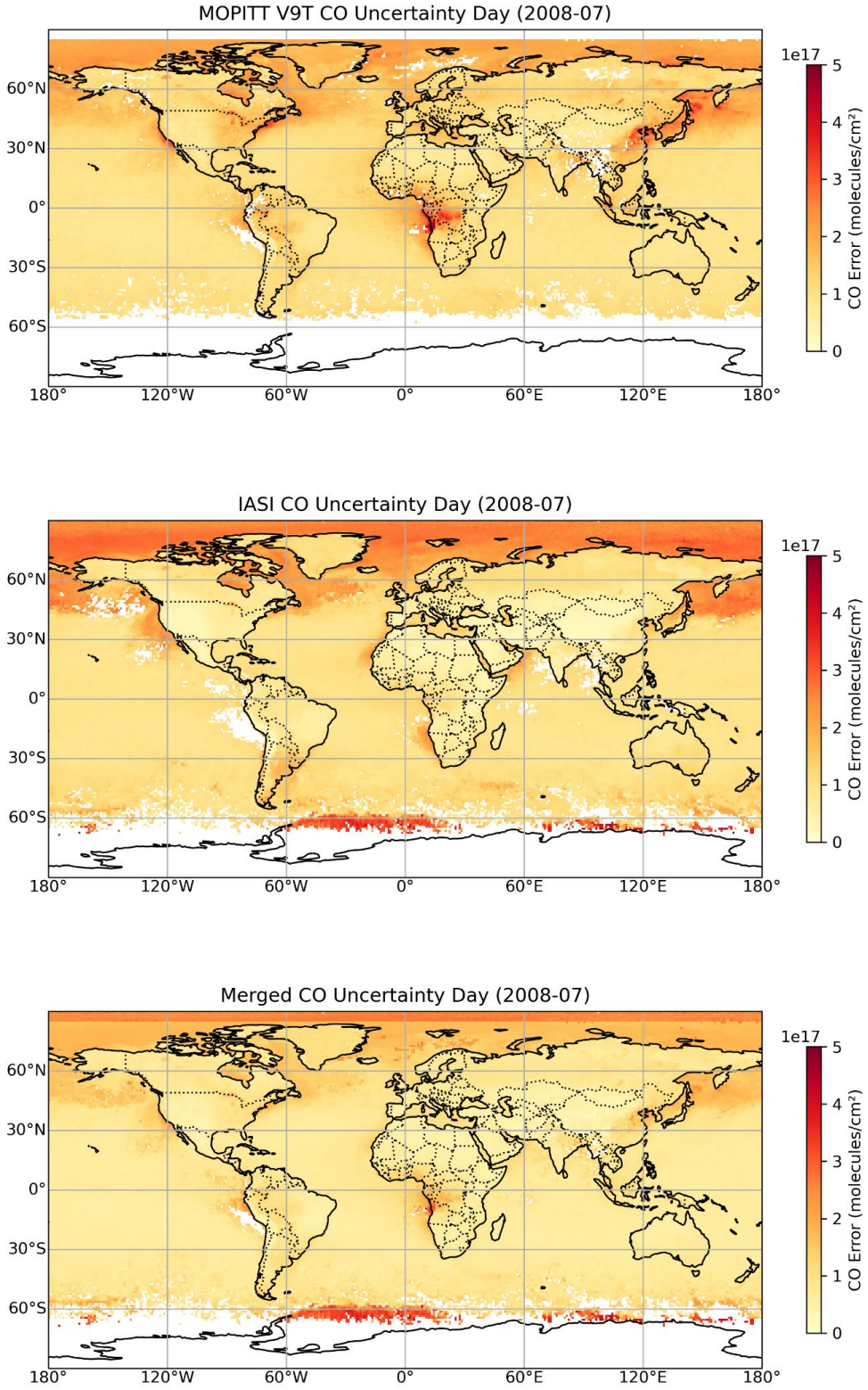


Figure 4.6.1 MOPITT, IASI and CCI+ MERGED CO uncertainty grids for July 2008 (daytime). White areas are cell grids with no data.

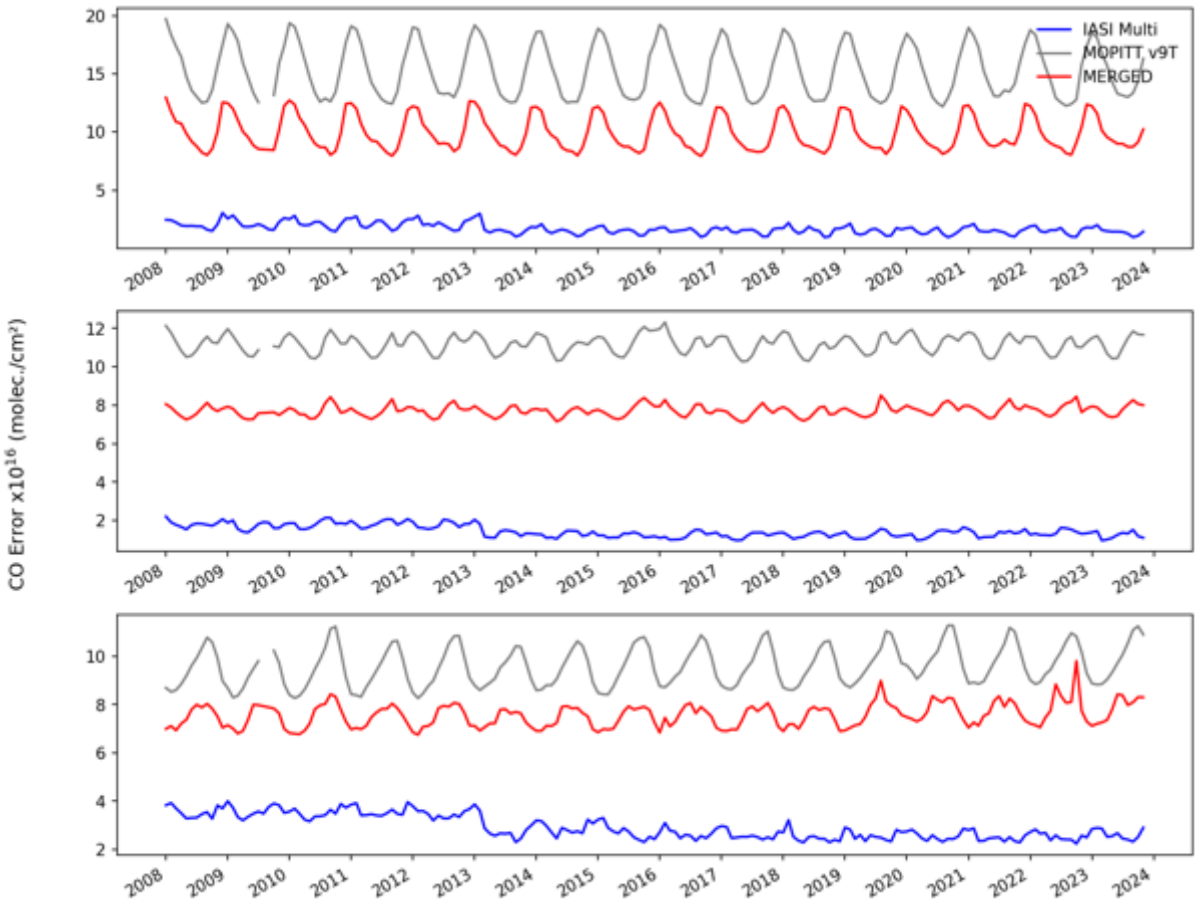


Figure 4.6.2 CO uncertainty time series for three latitude band ([20°N-60°N] top, [20°S-20°N] middle and [60°S-20°S] bottom), for daytime measurements. The newly merged CO CCI+ product is represented in red, IASI in blue and MOPITT in gray.

4.7 Uncertainty assessment for ammonia (L3-NH3)

4.7.1 Uncertainty on the Level 2

The uncertainty on the NH_3 Level 2 product is detailed in Clarisse et al. (2023) and is reproduced verbatim here below for those elements that are relevant for the L3.

In version 4 of the ANNI retrieval framework, the uncertainty on the measurement is estimated via propagation of uncertainty of the input parameters following the general formula

$$\sigma_X^2 = J^T S_p J$$

with $\mathbf{S}_p$ the error covariance matrix of the input parameters (with covariances $S_{p,ij} = \sigma_{p_i p_j}$) and $\boldsymbol{J}$ the Jacobian of the retrieval, with components $\frac{\partial X}{\partial p_j}$.

The input parameters include the skin temperature, the surface pressure, the HRI, the surface emissivity, the zenith angle, the width and the peak of the Gaussian vertical NH_3 profile, the temperature profile (15 levels) and the water vapour profile (7 levels). After preliminary analysis, it was concluded that only the correlations between the uncertainties in the temperature profile cannot be neglected. We therefore employ a block diagonal covariance matrix, block diagonal for the elements pertaining to the temperature profile, and diagonal for all other input parameters. Random and systematic errors are calculated separately.

As most input parameters come without an uncertainty budget, let alone covariances, we made best effort estimates of the co(variance) based on the limited information that is available. For now, the same (co)variances were used for the near-real time as for the reanalysed NH_3 product. It is also important to note that the systematic uncertainties of the input parameters vary according to the time and space scales that are considered (Boersma et al., 2004). Temperature profiles, for example, may be more biased monthly than annually. Here, we estimate systematic uncertainties with a typical L3 gridded data product in mind, i.e. for spatial scales of the order of 1° latitude and longitude or less, and for time periods of the order of 1 month or less.

The (co)variances, summarized in Table 4.7.1 were determined as follows:

HRI. By definition, the random uncertainty on the HRI equals one. We estimate a systematic uncertainty of 0.1 due to potential and residual interferences (e.g. surface emissivity, other species). To account for potential biases in the spectroscopy and generalized error covariance matrix, we add to this an additional 10 % on the calculated HRI value.

Skin temperature. Random and systematic uncertainties were set to 1.5 and 0.5 K respectively. These values are in line with the difference between the IASI L2 skin temperature product and the dedicated neural network used for the reanalysis product of ANNI.

Emissivity. For emissivity, which originates from the monthly climatology of Zou et al. (2013) an uncertainty of 0.01 and 0.005 was assumed for respectively the random and systematic components.

Temperature profile. Variances were set based on validation results of the IASI level 2 (Eumetsat, 2021) systematic uncertainties of 1 K for the surface level and 0.5 K for the other levels; random uncertainties of 2 K for the surface level and 1 K for the other levels for land observations, and 1 K for the surface level and 0.5 K for the other levels for ocean observations. Covariance matrices were then built by appropriate scaling of correlation matrices. These were built based on a statistical analysis of the differences between collocated ERA5 and IASI L2

profiles. Correlation coefficients were set to 0.5 between neighbouring levels and 0.25 between levels that are two levels apart. Above 10 km, no strong correlations were observed, and the covariance was therefore assumed to be diagonal for these levels.

Water vapour profiles. Relying again on the IASI level 2 validation report (Eumetsat, 2021) random uncertainties were set to 10 % below 3 km and 20 % above. Systematic uncertainties were set to half these numbers.

Surface pressure. A random and systematic uncertainty of 500 and 250 Pa was used.

NH_3 profiles. The uncertainties related to the NH_3 profile arise from uncertainties associated with the width and the peak of the Gaussian shaped vertical profile. Random and systematic uncertainties of 200 and 100 m were used for both parameters. Given the short lifetime of NH_3 in the atmosphere, these are likely of the right order of magnitude. To obtain better estimates in the future, a thorough analysis using in situ measurements or modelled profiles would be desirable.

Table 4.7.1 Estimated random and systematic uncertainties of the input parameters.

Component	Random σ_r	Systematic σ_s
HRI	1	0.1 + 10 %
Surface temperature (K)	1.5	0.5
Emissivity	0.01	0.005
Temperature profile, land (K)	1–2	0.5 – 1
Temperature profile, sea (K)	0.5–1	0.5 – 1
Surface pressure (Pa)	500 Pa	250 Pa
Water vapour profile	10 %– 20 %	5 %–10 %
NH3 profile peak altitude (m)	200	100
NH3 profile width (m)	200	100

4.7.2 Averaging Procedure

Using a subgrid of 0.025° × 0.025°, the fractional overlap of the native IASI footprint on ground with the 0.2° × 0.2° L3 grid is calculated (see Figure 4.6.1). The code for this is publicly available at https://squares.ulb.be/NH3-IASI.html, see also Clarisse et al. (2019). The L3 averages for a single grid cell are calculated as

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$$

with x_i all the observations that intersect with the given grid cell, and w_i the associated fractional overlap. This formula is used for all L3 variables, except for the average random uncertainty, which is calculated as

$$\sigma_{r\bar{X}}^2 = \frac{\sum_{i=1}^{n} \sigma_{rx_i}^2 w_i^2}{(\sum_{i=1}^{n} w_i)^2}$$

Note that the systematic uncertainty is calculated as normal. Averaged random (σ_r) and systematic (σ_s) uncertainties are available for each grid cell. If a total uncertainty is desired, this can be obtained from the square root of the sum of squares of the different terms:

$$\sigma^2 = \sigma_r^2 + \sigma_s^2$$

No attempt was made to estimate the uncertainty associated with the application of the postfilter (removing cloudy observations and observations with very low sensitivity to NH_3).

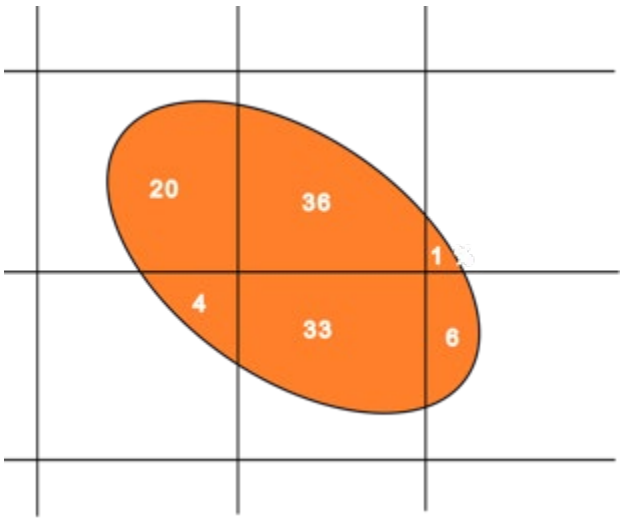


Figure 4.7.1 Illustration of the weight calculation of an IASI observation (orange ellipse), onto a regular grid.

4.7.3 Examples

Figures 4.7.2, 4.7.3 and 4.7.4 provide an example for the month of June 2023, morning overpass, Metop-B. Figure 4.6.2 shows the NH_3 total column together with the NH_3 total uncertainty, both in absolute and relative numbers. The distributions were filtered so that only grid cells with at least one observation are shown. Figure 4.6.3 shows the random and systematic components separately. Finally, figure 4.6.4 provides statistics on the total uncertainty for all columns and for columns above 2.5 10^{16} molec/cm². Most L3 monthly values have an estimated total uncertainty between 1 and 5 10^{15} molec/cm², which corresponds to a relative error of 10-50 % for enhanced columns.

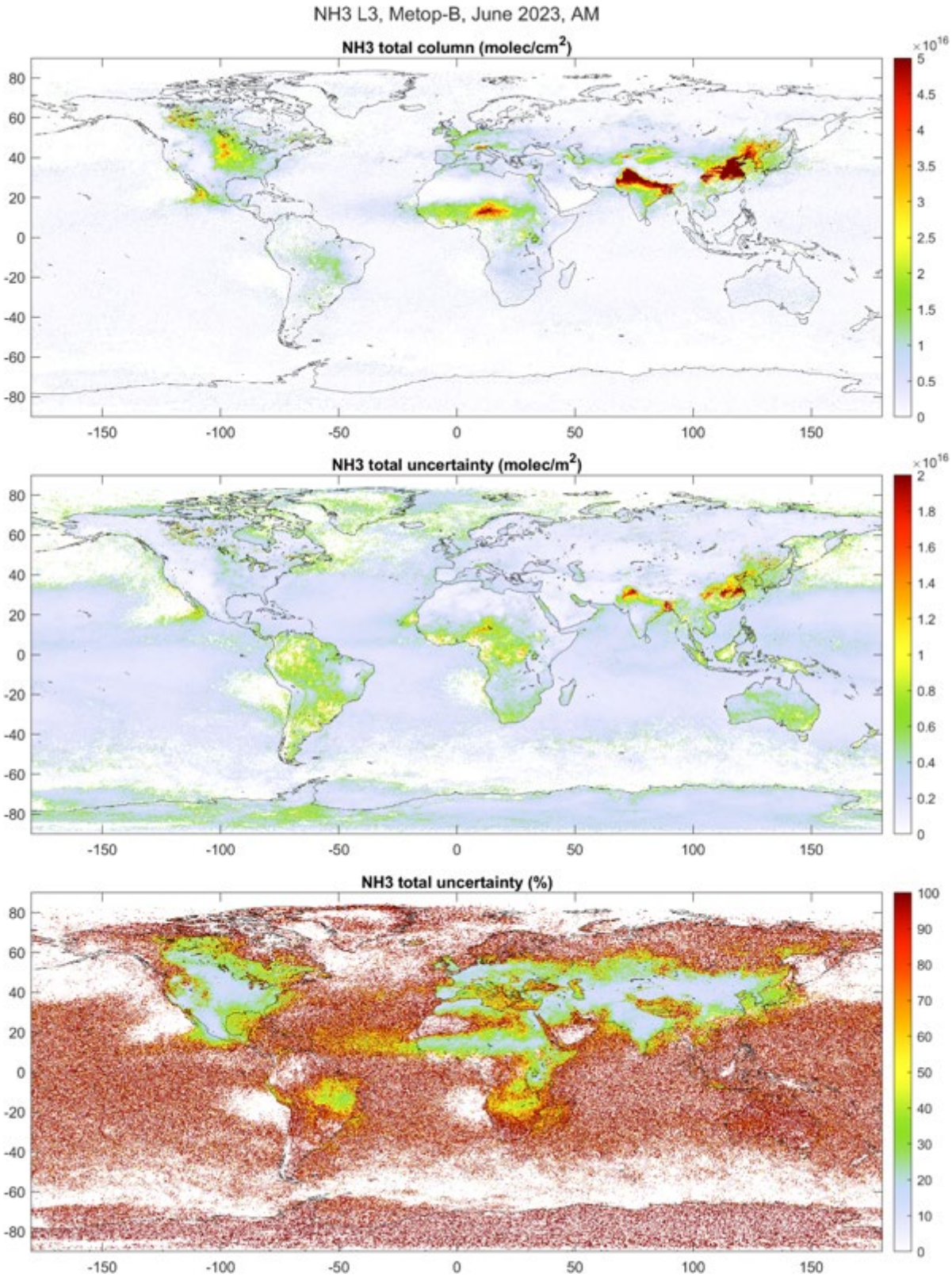


Figure 4.7.2 Example NH_3 L3 distribution (top), total uncertainty in absolute numbers (middle) and total uncertainty in percent (bottom).

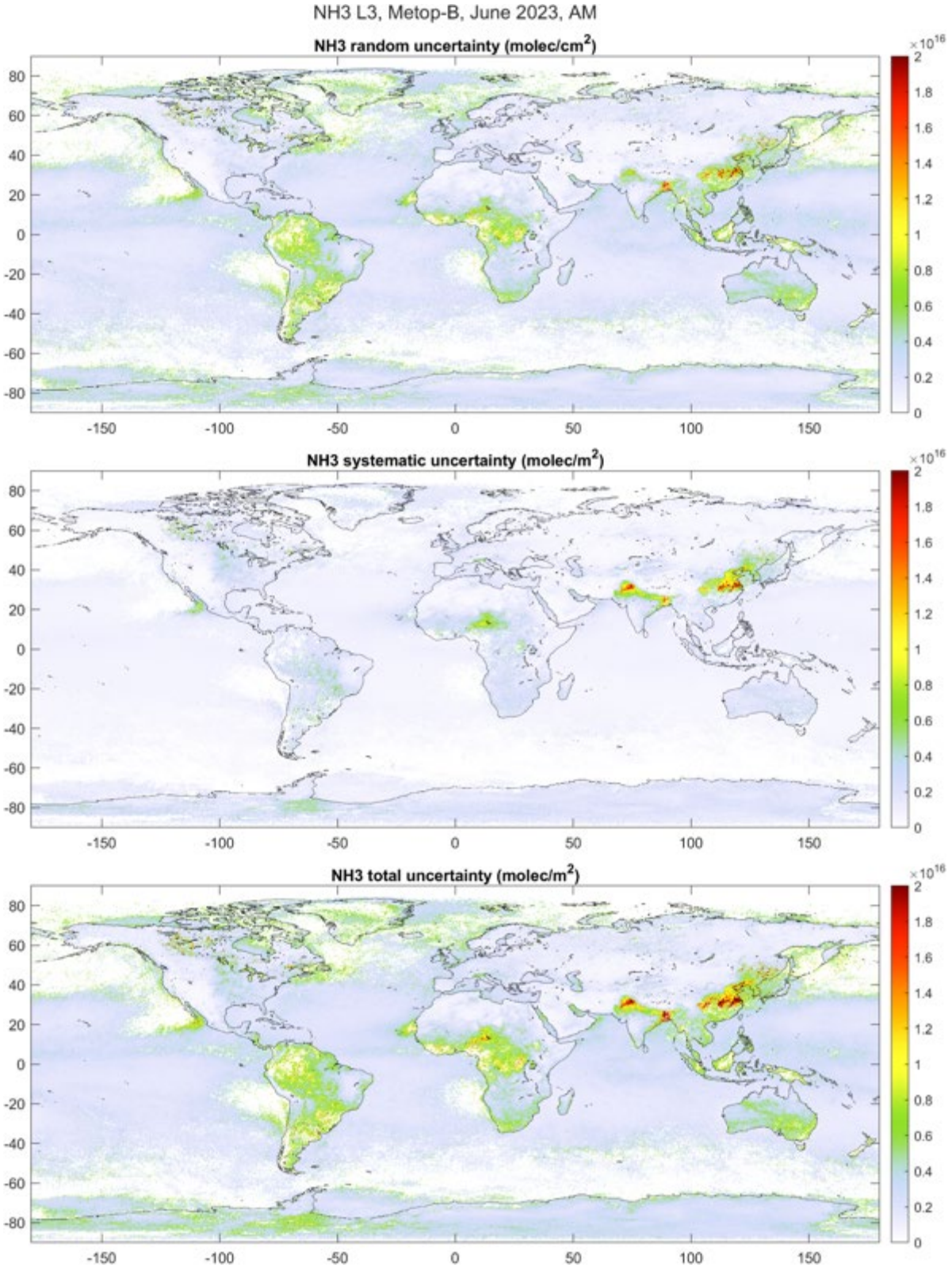


Figure 4.7.3 Example random (top), systematic (middle) and total uncertainty (bottom) for June 2023.

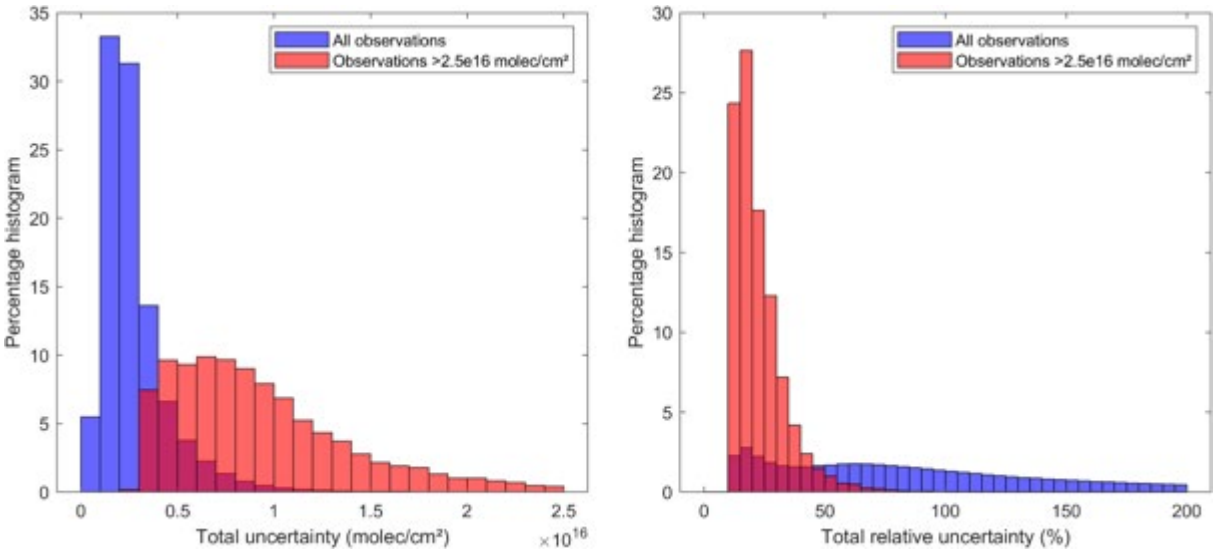


Figure 4.7.4 Statistics of the total NH_3 retrieval uncertainty for June 2023 (AM) from IASI onboard Metop-B. The left panel shows the uncertainty in absolute numbers, the right panel in relative numbers.

4.7.4 General recommendations

To users and the validation team: please see the corresponding section in the Precursors CCI+ Product User Guide (PUG).

In a general sense, users should be aware that trace gas L3 data are clear-sky sampled, and therefore not necessarily representative for all-sky monthly means (that would be obtained from CTM simulations).

Recommendation to the validation team is to as much as possible sample the validation data as the satellite data, i.e. both data sets should be co-sampling in space and time (i.e. on clear-sky days). Regarding the spatial mismatch, the validation team should be aware of pixel smearing of the trace gas over a grid cell vs. the more local representation of ground-based NO_2.

Recommendation to ESA:
Future technical work to improve the product quality or the completeness of the uncertainty characterization could focus on

- better understanding of trace gas vertical profiles.
- better constraining the systematic uncertainties in the product (cfr Table 4.7.1).
- characterizing potential biases in the product due to the application of the postfilter (clouds, observations with limited sensitivity to NH_3).

A better understanding of vertical trace gas profiles can be achieved by making available new in situ measured vertical profiles. Aircraft trace gas profiles, sampled around the overpass time of the satellite, allow the testing and improvement of the CTM simulated profile shapes, but also validation of the retrieved tropospheric columns. Traditionally, these aircraft experiments are led by the US, but European initiatives, for instance by making available IAGOS aircraft profiles to the validation community would be helpful.

Further efforts are needed to address remaining discontinuities, which are primarily duet o differences in spatial resolution, sampling frequency, measurement precision, and the use of different auxiliary inputs to algorithms from different sensors.

5 References

Alvarado, L. M. A., Richter, A., Vrekoussis, M., Wittrock, F., Hilboll, A., Schreier, S. F. and Burrows, J. P.: An improved glyoxal retrieval from OMI measurements, Atmos. Meas. Tech., 7(12), 4133–4150, doi:10.5194/amt-7-4133-2014, 2014.

Anglou, I., Intercomparison of OMI NO2 slant columns and O2-O2 cloud parameters retrieved from Collection 3 and Collection 4 level-1 data, in preparation, 2024.

Astoreca, R., Hurtmans, D., Coheur, P.-F., Fast Optimal Retrieval on Layers for IASI, Algorithm Theoretical Basis Document (ATBD), 2014, available at: https://acsaf.org/docs/atbd/Algorithm_Theoretical_Basis_Document_IASI_CO_Feb_2014.pdf

Beirle, S., Hörmann, C., Jöckel, P., Liu, S., Penning de Vries, M., Pozzer, A., Sihler, H., Valks, P., and Wagner, T.: The STRatospheric Estimation Algorithm from Mainz (STREAM): estimating stratospheric NO_2 from nadir-viewing satellites by weighted convolution, Atmos. Meas. Tech., 9, 2753–2779, https://doi.org/10.5194/amt-9-2753-2016, 2016.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res., 109, D04311, doi: 10.1029/2003JD003962, 2004.

Boersma, K. F., Jacob, D. J., Eskes, H. J., Pinder, R. W., Wang, J., and van der A, R. J.: Intercomparison of SCIAMACHY and OMI tropospheric NO_2 columns: Observing the diurnal evolution of chemistry and emissions from space, *J. Geophys. Res.*, 113, D16S26, doi:10.1029/2007JD008816, 2008.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905-1928, https://doi.org/10.5194/amt-4-1905-2011, 2011.

Boersma, K. F., Vinken, G. C. M., and Eskes, H. J.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV–Vis tropospheric column retrievals, Geosci. Model Dev., 9, 875-898, doi:10.5194/gmd-9-875-2016, 2016.

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., van Geffen, J. H. G. M., Zara, M., Peters, E., Van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., De Rudder, A., Irie, H., Pinardi, G., Lambert, J.-C., and Compernolle, S.: Improving algorithms and uncertainty estimates for satellite NO_2 retrievals: Results from the Quality Assurance for Essential Climate Variables (QA4ECV) project, Atmos. Meas. Tech., 11, 6651-6678, https://doi.org/10.5194/amt-11-6651-2018, 2018.

Bondy, W. H. and Zlot, W.: The Standard Error of the Mean and the Difference between Means for Finite Populations, The American Statistician, 30, 96–97, https://doi.org/10.1080/00031305.1976.10479149, 1976.

Castellanos, P., Boersma, K. F., Torres, O., and de Haan, J. F.: OMI tropospheric NO2 air mass factors over South America: effects of biomass burning aerosols, Atmos. Meas. Tech., 8, 3831-3849, doi:10.5194/amt-8-3831-2015, 2015.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J. and Chance, K.: Glyoxal retrieval from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 7(11), 3891–3907, doi:10.5194/amt-7-3891-2014, 2014.

Clarisse, L., Van Damme, M., Clerbaux, C., and Coheur, P.-F.: Tracking down global NH3 point sources with wind-adjusted superresolution, Atmos. Meas. Tech., 12, 5457–5473, https://doi.org/10.5194/amt-12-5457-2019, 2019.

Clarisse, L., Franco, B., Van Damme, M., Di Gioacchino, T., Hadji-Lazaro, J., Whitburn, S., Noppen, L., Hurtmans, D., Clerbaux, C., and Coheur, P.: The IASI NH3 version 4 product: averaging kernels and improved consistency, Atmos. Meas. Tech., 16, 5009–5028, https://doi.org/10.5194/amt-16-5009-2023, 2023.

Danckaert, T., Fayt, C., van Roozendael, M., De Smedt, I., Letocard, V., Merlaud, A. and Pinardi, G.: QDOAS software user manual. [online] Available from: http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf, 2017.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 4947–4963, https://doi.org/10.5194/acp-8-4947-2008, 2008.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395-2426, https://doi.org/10.5194/amt-11-2395-2018, 2018.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561–12593, https://doi.org/10.5194/acp-21-12561-2021, 2021.

Dirksen, R., Boersma, K., Eskes, H., Ionov, D., Bucsela, E., Levelt, P., and Kelder, H.: Evaluation of stratospheric NO2 retrieved from the Ozone Monitoring Instrument: Intercomparison, diurnal cycle, and trending, Journal of Geophysical Research Atmospheres, 116, https://doi.org/10.1029/2010JD014943, 2011.

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., and Veefkind, P.: Comparing Sentinel-5P TROPOMI NO_2 column observations with the CAMS regional air quality ensemble, Geosci. Model Dev., 16, 509–534, https://doi.org/10.5194/gmd-16-509-2023, 2023.

Eskes, H. and Boersma, K.: Averaging kernels for DOAS total-column satellite retrievals, Atmospheric Chemistry and Physics, 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, 2003.

Eskes, H. and Eichmann, K.-U.: S5P Mission Performance Centre Nitrogen Dioxide [L2 NO2] Readme, Tech. Rep. S5P-MPC-KNMI-PRF-NO2, KNMI, 2023.

Eskes, H., Velthoven, P., Valks, P., and Kelder, H.: Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model, Quarterly Journal of the Royal Meteorological Society, 129, 1663—–1681, https://doi.org/10.1256/qj.02.14, 2003.

EUMETSAT: Validation, Tech. Rep. EUM/OPS/DOC/19/1091375, EUMETSAT, https://doi.org/10.15770/EUM_SEC_CLM_0027, 2021

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445-473, https://doi.org/10.5194/gmd-3-445-2010, 2010.

Hurtmans, D., Coheur, P., Wespes, C., Clarisse, L., Scharf, O., Clerbaux, C., Hadji-Lazaro, J., George, M. and Turquety, S.: FORLI radiative transfer and retrieval code for IASI. J. Quant. Spectrosc. Radiat. Transfer, 113, 1391-1408, https://doi.org/10.1016/j.jqsrt.2012.02.036, 2012.

Isserlis, L.: On the Value of a Mean as Calculated from a Sample, Journal of the Royal Statistical Society, 81, 75–81, URL http://www.jstor.org/stable/2340569, 1918.

Ku, H.: Notes on the use of propagation of error formulas, Journal of Research of the National Bureau of Standards, Section C: Engineering and Instrumentation, 70C, 263, https://doi.org/10.6028/jres.070c.025, 1966.

Lerot, C., Stavrakou, T., De Smedt, I., Müller, J. F., Van Roozendael, M., M̈uller, J. F. and Van Roozendael, M.: Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model,

Atmos. Chem. Phys., 10(24), 12059–12072, doi:10.5194/acp-10-12059-2010, 2010.

Lerot , C., Stavrakou, T., Van Roozendael, M., Alvarado, L.M.A. and Richter, A.: GLYoxal Retrievals from TROPOMI (GLYRETRO) – ATBD, 2021, available at https://data-portal.s5p-pal.com/product-docs/chocho/CHOCHO_ATBD_S5P+I_BIRA_v3.1.pdf

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J. T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J. and Krol, M.: Structural uncertainty in air mass factor calculation for NO2 and HCHO satellite retrievals, Atmos. Meas. Tech., 10(3), 759–782, doi:10.5194/amt-10-759-2017, 2017.

Liu, M., Lin, J., Boersma, K. F., Pinardi, G., Wang, Y., Chimot, J., Wagner, T., Xie, P., Eskes, H., Van Roozendael, M., Hendrick, F., Wang, P., Wang, T., Yan, Y., Chen, L., and Ni, R.: Improved aerosol correction for OMI tropospheric NO_2 retrieval over East Asia: constraint from CALIOP aerosol vertical profile, Atmos. Meas. Tech., 12, 1-21, https://doi.org/10.5194/amt-12-1-2019, 2019.

Liu, S., Valks, P., Pinardi, G., De Smedt, I., Yu, H., Beirle, S., and Richter, A.: An improved total and tropospheric NO2 column retrieval for GOME-2, Atmos. Meas. Tech., 12, 1029–1057, https://doi.org/10.5194/amt-12-1029-2019, 2019.

Liu, S., Valks, P., Pinardi, G., Xu, J., Argyrouli, A., Lutz, R., Tilstra, L. G., Huijnen, V., Hendrick, F., and Van Roozendael, M.: An improved air mass factor calculation for nitrogen dioxide measurements from the Global Ozone Monitoring Experiment-2 (GOME-2), Atmos. Meas. Tech., 13, 755–787, https://doi.org/10.5194/amt-13-755-2020, 2020.

Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofi eva, V. F., and Wagner, W.: Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data, 9, 511–527, https://doi.org/10.5194/essd-9-511-2017, 2017.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NOx emission estimates derived from an assimilation of OMI tropospheric NO2 columns, Atmospheric Chemistry and Physics, 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012.

Oldeman, A. M.: Effect of including an intensity offset in the DOAS NO2 retrieval of TROPOMI, Internship report, R-1944-SE, Eindhoven University of Technology, Eindhoven, The Netherlands, 2018.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice* (Vol. 2). World scientific.

Riess, T. C. V. W., Boersma, K. F., van Vliet, J., Peters, W., Sneep, M., Eskes, H., and van Geffen, J.: Improved monitoring of shipping NO_2 with TROPOMI: decreasing NO_x emissions in European seas during the COVID-19 pandemic, Atmos. Meas. Tech., 15, 1415–1438, https://doi.org/10.5194/amt-15-1415-2022, 2022.

Rijsdijk, P., Eskes, H., Dingemans, A., Boersma, F., Sekiya, T., Miyazaki, K., and Houweling, S.: Quantifying uncertainties of satellite NO2 superobservations for data assimilation and model evaluation, EGUsphere, 2024, 1–42, https://doi.org/10.5194/egusphere-2024-632, 2024.

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A. C., Stebel, K., and Witek, M. L.: A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing, Atmos. Meas. Tech., 13, 373–404, https://doi.org/10.5194/amt-13-373-2020, 2020.

Sekiya, T., Miyazaki, K., Ogochi, K., Sudo, K., Takigawa, M., Eskes, H., and Boersma, K. F.: Impacts of Horizontal Resolution on Global Data Assimilation of Satellite Measurements for Tropospheric Chemistry Analysis, Journal of Advances in Modeling Earth Systems, 13, e2020MS002 180, https://doi.org/10.1029/2020MS002180, 2021.

Sekiya, T., Miyazaki, K., Eskes, H., Sudo, K., Takigawa, M., and Kanaya, Y.: A comparison of the impact of TROPOMI and OMI tropospheric NO2 on global chemical data assimilation, Atmospheric Measurement Techniques, 15, 1703–1728, https://doi.org/10.5194/amt-15-1703-2022, 2022.

Seo, S., Valks, P., Boersma, K. F., De Smedt, I., Pinardi, G., Liu, S., Richter, A., and Loyola, D.: Development and evaluation of a long-term NO_2 data recored from GOME-2 observations, in preparation, 2023.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, Atmos. Chem. Phys., 9, 1037–1060, https://doi.org/10.5194/acp-9-1037-2009, 2009.

Tellinghuisen, J.: Statistical Error Propagation, The Journal of Physical Chemistry A, 105, 3917–3921, https://doi.org/10.1021/jp003484u, 2001.

Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedergnana, M., Loyola, D. and Van Roozendael, M.: Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis, Atmos. Meas. Tech., 10(1), 119–153, doi:10.5194/amt-10-119-2017, 2017.

Theys, N., Fioletov, V., Li, C., De Smedt, I., Lerot, C., McLinden, C., Krotkov, N., Griffin, D., Clarisse, L., Hedelt, P., Loyola, D., Wagner, T., Kumar, V., Innes, A., Ribas, R., Hendrick, F., Vlietinck, J., Brenot, H., and Van Roozendael, M.: A sulfur dioxide Covariance-Based Retrieval Algorithm (COBRA): application to TROPOMI reveals new emission sources, Atmos. Chem. Phys., 21, 16727–16744, https://doi.org/10.5194/acp-21-16727-2021, 2021.

van Geffen, J., Boersma, K. F., Eskes, H., Sneep, M., ter Linden, M., Zara, M., and Veefkind, J. P.: S5P TROPOMI NO_2 slant column retrieval: method, stability, uncertainties and comparisons with OMI, Atmos. Meas. Tech., 13, 1315–1335, https://doi.org/10.5194/amt-13-1315-2020, 2020.

van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., ter Linden, M., Ludewig, A., Boersma, K. F., and Veefkind, J. P.: Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data, Atmos. Meas. Tech., 15, 2037–2060, https://doi.org/10.5194/amt-15-2037-2022, 2022.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P. and Winkler, H.: TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, Atmos. Meas. Tech., 13(7), 3751–3767, doi:10.5194/amt-13-3751-2020, 2020.

Visser, A. J., Boersma, K. F., Ganzeveld, L. N., and Krol, M. C.: European NO_x emissions in WRF-Chem derived from OMI: impacts on summertime surface ozone, Atmos. Chem. Phys., 19, 11821–11841, https://doi.org/10.5194/acp-19-11821-2019, 2019.

Wald, A. and Wolfowitz, J.: An exact test for randomness in the non-parametric case based on serial correlation, Ann. Math. Statist., 14, 378–388, https://doi.org/10.1214/aoms/1177731358, 1943.

Worden et al., MOPITT Version 9 Product User's Guide, https://www2.acom.ucar.edu/sites/default/files/documents/v9_users_guide_20220203.pdf, 2022.

Zara, M., Boersma, K. F., De Smedt, I., Richter, A., Peters, E., Van Geffen, J. H. G. M., Beirle, S., Wagner, T., Van Roozendael, M., Marchenko, S., Lamsal, L. N., and Eskes, H. J.: Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty characterization, and trends, Atmos. Meas. Tech., 11, 4033-4058, https://doi.org/10.5194/amt-11-4033-2018, 2018.

Zhou, D. K., Larar, A. M., and Liu, X.: MetOp-A/IASI Observed Continental Thermal IR Emissivity Variations, IEEE J. Sel. Top. Appl., 6, 1156–1162, https://doi.org/10.1109/JSTARS.2013.2238892, 2013.