Earth Syst. Sci. Data, 17, 4627–4650, 2025
https://doi.org/10.5194/essd-17-4627-2025

Open Access Earth System Science Data

TROPOMI Level 3 tropospheric NO_2 dataset with advanced uncertainty analysis from the ESA CCI+ ECV precursor project

Isolde Glissenaar[1,2], Klaas Folkert Boersma[1,3], Isidora Anglou[1], Pieter Rijsdijk[1,4,5], Tijl Verhoelst[6], Steven Compernolle[6], Gaia Pinardi[6], Jean-Christopher Lambert[6], Michel Van Roozendael[6], and Henk Eskes[1]

[1]Royal Netherlands Meteorological Institute (KNMI), Satellite Observations department, De Bilt, the Netherlands
[2]University of Tromsø The Arctic University of Norway, Department of Physics and Technology, Tromsø, Norway
[3]Wageningen University, Meteorology and Air Quality group, Wageningen, the Netherlands
[4]Vrije Universiteit, Department of Earth Sciences, Amsterdam, the Netherlands
[5]SRON Netherlands Institute for Space Research, Leiden, the Netherlands
[6]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

Correspondence: Klaas Folkert Boersma (folkert.boersma@knmi.nl)

Received: 28 December 2024 – Discussion started: 12 February 2025
Revised: 6 June 2025 – Accepted: 17 June 2025 – Published: 22 September 2025

Abstract. We introduce the new ESA Climate Change Initiative TROPOspheric Monitoring Instrument (TROPOMI) global monthly clear-sky Level 3 (L3) dataset of tropospheric nitrogen dioxide (NO_2) for May 2018 to December 2021. The dataset provides spatiotemporally averaged tropospheric NO_2 columns, associated averaging kernels, and L3 uncertainties at spatial resolutions of 0.2, 0.5, and 1.0° on a monthly timescale (https://doi.org/10.21944/CCI-NO2-TROPOMI-L3, KNMI, 2025). To improve our understanding of what fraction of the Level 2 (L2) uncertainty cancels when averaging over space or time (i.e. the random component of the L2 uncertainty) and what fraction persists despite the averaging (systematic component), we first determine spatial and temporal error correlations for all sources of uncertainty in the L2 retrieval. Spatial error correlations arise mainly from the stratosphere–troposphere correction and from coarse-gridded albedo climatologies used in the L2 air mass factor calculation and have been quantified in previous studies. We find the temporal error correlation in both the stratospheric uncertainty and the air mass factor uncertainty to be 30 %. Using these estimates, the L3 uncertainty budget has been established for every grid cell based on input L2 uncertainties and new methods to estimate spatial and temporal representativeness uncertainties and to propagate measurement uncertainties through space and time. The total relative uncertainty in the resulting L3 dataset is in the range of 15 %–20 % in polluted areas, which is significantly lower than in individual L2 orbit retrievals, and brings the tropospheric NO_2 data to within the Global Climate Observing System (GCOS) "goal" and "breakthrough" requirements. Validation of the tropospheric, stratospheric, and total columns confirms better correlation and reduced dispersion in the differences between satellite and ground-based reference data for the L3 data with respect to the underlying L2, albeit with a more pronounced negative bias in the tropospheric columns at pollution hot spots, most probably related to stronger spatial smearing.

Published by Copernicus Publications.

1 Introduction

Long-term monitoring of nitrogen oxides ($NO_x = NO + NO_2$) in the atmosphere is crucial for several reasons. Nitrogen oxides are harmful air pollutants, and long-term exposure is causally linked to chronic respiratory diseases and mortality in humans (Faustini et al., 2014; Fischer et al., 2015). When nitrogen dioxide (NO_2) is oxidised in the atmosphere, it forms nitric acid (HNO_3), which readily dissolves in airborne droplets and subsequently comes down as acid rain. Excessive deposition of HNO_3 has been shown to lead to adverse feedbacks on land and water ecosystems (Tan et al., 2020), and changes in N-deposition influence ecosystem carbon sinks, affecting the biosphere's capacity to capture carbon from the atmosphere in the long run (Liu et al., 2022). Moreover, the monitoring of NO_x concentrations helps to identify major sources of CO_2, because both species are co-emitted upon combustion of fossil fuels by vehicles, industrial activities, and power plants (e.g. Zhang et al., 2023). Finally, NO_x play a significant role in the formation of ozone smog (Zhang et al., 2021) and fine particulate matter (Zhang et al., 2015), and both these secondary pollutants further harm human health and the environment. The importance of nitrogen oxides and their indirect impact on methane, ozone, and aerosol is an important reason why tropospheric NO_2 columns along with other trace gases have been selected as the so-called essential climate variable (ECV) precursors for aerosols and ozone (Zemp et al., 2022).

To allow for policy development and air pollution and climate change assessments, inter-annual changes and trends in recent decades need to be calculated. For this, long-term, robust, sustainable, and scientifically sound climate data records (CDRs) are needed to provide trustworthy information on how, where, and to what extent nitrogen oxide concentrations are changing. The European Space Agency (ESA) Climate Change Initiative (CCI) set out to develop robust satellite data records of key components of the climate system, including nitrogen dioxide concentrations. Atmospheric monitoring of NO_x involves application of in situ measurement techniques from ground-based or airborne platforms, but these measurements are limited in their spatio-temporal coverage. Satellite remote sensing with UV/Vis sensors, by contrast, has provided global measurements of tropospheric NO_2 columns since the mid-1990s. The wide spatial coverage and continuity of satellite measurements make them fit for purpose for climate monitoring. Among the key satellite instruments used to monitor tropospheric NO_2 columns are OMI (Ozone Monitoring Instrument) on NASA's Aura satellite and TROPOMI (TROPOspheric Monitoring Instrument) on ESA's Sentinel-5 Precursor, which provide long-term, high-resolution datasets critical for understanding air pollution trends.

Level 2 (L2) retrieval algorithms to derive NO_2 columns from raw satellite measurements (Level 1 data) have received a lot of attention (e.g. Van Geffen et al., 2022a). A framework for the evaluation of quality assurance measures was created in the Quality Assurance for Essential Climate Variables (QA4ECV) project for multiple climate variables including NO_2 columns for the L2 products of a group of sensors (not including TROPOMI) (Nightingale et al., 2018), providing users with evidence-based confidence in the products and enabling judgement on the fitness for purpose for specific applications. In comparison, Level 3 (L3) data, spatially and/or temporally averaged products on a consistent grid derived from L2 data (e.g. Wei et al., 2022), have been less considered by the scientific community. However, L3 data are relevant for model evaluation (Visser et al., 2019; Eskes et al., 2024), data assimilation (Inness et al., 2019; Sekiya et al., 2022), and climate (trend) studies (e.g. Zara et al., 2021). The emphasis on L2 retrieval algorithms is understandable, as it reflects the need to ensure the foundational accuracy of the L2 data, which is the starting point for creating high-quality L3 products. However, the quality of L3 products not only depends on reliable input data, but also requires a good understanding of best methods for averaging both spatially and temporally and assessing the propagation of existing measurement uncertainties and the quantification of additional sampling uncertainties.

Including rigorous uncertainty information in L3 CDRs is important to support the application of the data (Merchant et al., 2017). This is necessary to avoid misinterpretation of artefacts arising from system limitations such as real geophysical changes or trends (e.g. Labzovskii et al., 2024), for modellers to get confidence in discriminating model–data discrepancies that unambiguously indicate model deficiencies from those where observational errors are significant, and to contribute appropriate weighting to different observations in data assimilation and reanalyses (Merchant et al., 2017). An accurate quantification of L3 uncertainties includes both the assessment of the magnitude of error sources, as well as a propagation of these uncertainties to the L3 data product, including a treatment of spatio-temporal error correlations between individual satellite observations and of aspects of spatio-temporal representativity.

Measurement uncertainties are often assumed to be either fully systematic or fully random in space and time when determining gridded datasets (e.g. Wenig et al., 2008; Chan et al., 2023). In reality, there will be spatial and temporal correlations between multiple error sources, depending on length scale and timescales. There are datasets where uncertainties are determined using partial error correlations, but until now the correlation coefficients used have been determined using expert opinion, introducing subjectivity (Miyazaki et al., 2012; Boersma et al., 2016, 2018). A more quantitative treatment of spatial error correlations and representativeness errors in case of incomplete coverage of a grid cell has been presented in Rijsdijk et al. (2025). Their analysis reveals a grid-size-dependent partially systematic error contribution from the air mass factor and a fully systematic error from the stratospheric column. A detailed analysis

of the effect of systematic sampling due to cloud cover resulted in a mathematically derived realistic spatial representation error. Rijsdijk et al.'s methods to reconstruct realistic uncertainties in spatial averages will be discussed in more detail in Sect. 3.2.2. In this paper we extend this spatial-averaging approach presented in Rijsdijk et al. (2025) to spatio-temporally averaged L3 products (gridded monthly means). This brings us to the following research questions for this study: (1) Can we improve our understanding of how uncertainties in satellite-derived tropospheric NO_2 columns are correlated in space and in time? (2) Can we use this understanding of correlations in uncertainties to better characterise how these uncertainties propagate into gridded, monthly mean uncertainty estimates? (3) What is the monthly mean uncertainty budget for TROPOMI NO_2 L3 data, and how do these L3 uncertainties vary in space and time? (4) To what extent does the validation with independent reference measurements help to assess the quality and fitness for purpose of the TROPOMI NO_2 L3 data?

We present the ESA CCI+ L3 TROPOMI atmospheric NO_2 dataset with a thorough assessment of the L3 uncertainty, combining measurement uncertainties, sampling uncertainties, and a proper assessment of local error correlations for the uncertainty propagation. This includes, for the first time, an empirical quantification of correlation coefficients for multiple error sources.

2 Instrument and dataset

2.1 TROPOMI instrument

The TROPOspheric Monitoring Instrument (TROPOMI) (Veefkind et al., 2012) provides data on tropospheric NO_2 (and many other trace gases) columns with daily global coverage at spatial resolution of $7 \times 3.5\,\mathrm{km}^2$ and, since 6 August 2019, of $5.5 \times 3.5\,\mathrm{km}^2$ at nadir. TROPOMI is aboard the European Space Agency (ESA) Sentinel-5 Precursor (S5P) satellite, which was launched on 13 October 2017 and has been providing nominal observations since May 2018. The near-polar sun-synchronous orbit provides afternoon observations with an Equator local overpass time of 13:30 and a nearly daily global coverage.

2.2 TROPOMI Level 2 tropospheric NO_2 columns and uncertainties

The starting point for the generation of L3 data is the L2 TROPOMI NO_2 tropospheric vertical columns on an orbital basis (Copernicus Sentinel-5P, 2021a). The aim of the ESA CCI+ project is to generate long-term climate data records, and therefore we use the TROPOMI PAL v2.3.1 L2 data, which are the most consistent with the OMI QA4ECV v1.1 product (Boersma et al., 2018). These OMI and TROPOMI data products are based on strongly consistent algorithms that use the same OMI surface albedo climatology (the so-called MINLER; Kleipool et al., 2008), which allows for better merging of the datasets and allows using data from 2004 to 2022. Still, the methods described in this paper are applicable on the operational dataset or any other version of the L2 data as well. TROPOMI retrievals with qa_value of > 0.75 were used, which corresponds to good-quality retrievals over (nearly) cloud-free scenes. Retrievals made in the descending part of the orbit are removed because high-latitude retrievals in the descending part of the orbit are not being used in the stratospheric correction procedure, leading to high-biased stratospheric NO_2 columns and, on average, negative tropospheric NO_2 columns for the descending part of the orbit (see Appendix A). A small bug with respect to the qa_value over snow and ice present in version 2.3.1 was corrected in this study (following the procedure set out in Sect. 5.1 in van Geffen et al., 2024).

The NO_2 retrieval procedure consists of three steps: the spectral fitting, the stratospheric correction, and the conversion of the tropospheric slant column density (SCD) into a tropospheric vertical column density using the air mass factor. Each of these three steps introduces potential errors and contributes to the overall uncertainty. The single tropospheric column uncertainty (σ_i) is quantified as in Boersma et al. (2004): $\sigma_i = \sqrt{\left(\sigma_{N_\mathrm{s}}\right)^2 + \left(\sigma_{N_\mathrm{s}^\mathrm{strat}}\right)^2 + \left(\sigma_{M^\mathrm{tr}}\right)^2}$, where the tropospheric column uncertainty sources are

$$\sigma_{N_\mathrm{s}} = \frac{\sigma'_{N_\mathrm{s}}}{M^\mathrm{tr}};\quad \sigma_{N_\mathrm{s}^\mathrm{strat}} = \frac{\sigma'_{N_\mathrm{s}^\mathrm{strat}}}{M^\mathrm{tr}};\quad \sigma_{M^\mathrm{tr}} = \frac{(N_\mathrm{s} - N_\mathrm{s}^\mathrm{strat}) \cdot \sigma'_{M^\mathrm{tr}}}{(M^\mathrm{tr})^2},$$

with N_s the slant column density; $N_\mathrm{s}^\mathrm{strat}$ the stratospheric slant column density; M^tr the tropospheric air mass factor; and σ'_{N_s}, $\sigma'_{N_\mathrm{s}^\mathrm{strat}}$, and σ'_{M^tr} their respective uncertainties.

The slant column density uncertainty σ'_{N_s} is estimated on a per-pixel basis during the spectral fitting and is obtained from the diagonal of the covariance matrix of the standard errors (Van Geffen et al., 2020) and has typical values of $\sim 0.6 \times 10^{15}$ molecules cm^{-2} (Van Geffen et al., 2020, 2022a). This value compares well with values obtained with an alternative method to estimate uncertainties in the NO_2 slant column densities (by statistical analysis of the distribution of SCDs corrected for viewing geometry over regions with very little variability in stratospheric NO_2, reported in Boersma et al., 2004, Zara et al., 2018, and, most recently, for TROPOMI, in Van Geffen et al., 2020).

TM5-MP is being used in the TROPOMI NO_2 retrieval algorithm by providing estimates of stratospheric NO_2 columns. TM5-MP is a global chemistry-transport model that simulates atmospheric trace gas distributions at $1^\circ \times 1^\circ$ resolution in the troposphere and in the stratosphere (Williams et al., 2017). The stratospheric slant column uncertainty $\sigma'_{N_\mathrm{s}^\mathrm{strat}}$ is based on a global statistical analysis of results from the data assimilation procedure used to separate the tropospheric and stratospheric columns. The data assimilation procedure uses the observed slant columns and TM5-MP 24 h forecast stratospheric NO_2 fields (after modelled

transport and chemistry). The difference between modelled forecasts and the actual observations (O–F) over unpolluted scenes is taken as an upper limit of the uncertainty in stratospheric NO_2 columns and has a statistical global-mean value of 0.2×10^{15} molecules cm^{-2} (Dirksen et al., 2011), which is applied in the L2 algorithm (van Geffen et al., 2022b). Here, we apply a more detailed latitude- and time-dependent L2 stratospheric uncertainty as derived by Rijsdijk et al. (2025), where the O–F is analysed over latitudinal and day-of-year bands, resulting in a look-up table. A recent comparison of stratospheric NO_2 columns obtained with data assimilation in TM5-MP and the STREAM approach (Beirle et al., 2016), an alternative method to separate the stratosphere and troposphere columns, for GOME-2A showed consistency to within 0.2×10^{15} molecules cm^{-2} (Richter et al., 2024). Similar results were found in a comparison of stratospheric NO_2 columns from the data assimilation method and STREAM for OMI of the QA4ECV dataset (Boersma et al., 2018; Compernolle et al., 2020) and are to be expected when applied to TROPOMI.

The air mass factor (AMF) uncertainty $\sigma'_{M^{\mathrm{tr}}}$ consists of uncertainty contributions from the cloud pressure, cloud fraction, surface albedo, and a priori NO_2 profile shape. The theoretical error propagation framework in Boersma et al. (2004) is used to estimate the overall AMF uncertainty. The overall tropospheric AMF uncertainties are estimated to be 30 %–50 % (Boersma et al., 2018) for individual retrievals.

The overall uncertainty for an individual retrieval therefore depends on details in the retrieval and is pixel specific. Over oceans and remote areas, with low tropospheric vertical columns, the relative overall uncertainty is typically more than 100 % and is dominated by uncertainty in the spectral fit ($\sigma_{N_{\mathrm{s}}}$) and the stratospheric slant column density ($\sigma_{N_{\mathrm{s}}^{\mathrm{strat}}}$) (van Geffen et al., 2022b). For more polluted regions over continental areas, the relative uncertainty reduces to 25 %–50 % and is dominated by uncertainty in the tropospheric air mass factor ($\sigma_{M^{\mathrm{tr}}}$) (van Geffen et al., 2022b). The overall uncertainty for individual TROPOMI tropospheric vertical column NO_2 retrievals ($N_{\mathrm{v}}^{\mathrm{trop}}$) is sometimes approximated as $\sigma_i \approx 0.5 \times 10^{15}$ molecules $cm^{-2} + [0.2 \text{ to } 0.5] \times N_{\mathrm{v}}^{\mathrm{trop}}$ (van Geffen et al., 2022b).

3 Methodology

3.1 L3 algorithm overview

Our starting point is individual L2 tropospheric vertical column retrievals with a retrieved column value x_i (molecules cm^{-2}) and the associated individual retrieval uncertainty σ_i (molecules cm^{-2}). To compute L3 spatial and temporal gridded means, we developed a two-step procedure, in which we

1. calculate spatially averaged (gridded) column values ($x_{\mathrm{o,t}}$) – where t stands for an instantaneous column value not averaged in time – along with their associated L2 measurement uncertainties (σ_{m}) and spatial representativeness uncertainty (σ_{rs}), combined into the spatial average uncertainty ($\sigma_{\mathrm{o,t}}$). Here we follow the spatial averaging approach and uncertainty assessment described in Rijsdijk et al. (2025).

2. provide temporal averaged estimates ($\bar{x}$) of the spatially averaged column values ($x_{\mathrm{o,t}}$) and their total uncertainty ($\bar{\sigma}_{\mathrm{total}}$), including associated temporal representativeness uncertainty ($\bar{\sigma}_{\mathrm{rt}}$) and measurement uncertainties ($\bar{\sigma}_{\mathrm{m}}$).

Figure 1 provides a schematic picture of the above procedure. After collecting the individual values of the retrieved column (x_i) and its uncertainty (σ_i), their spatially averaged (instantaneous) counterparts ($x_{\mathrm{o,t}}$, $\sigma_{\mathrm{o,t}}$) 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. This two-step method allows us to assess and apply different error correlation factors when propagating uncertainties spatially and temporally.

3.2 Spatial averaging

3.2.1 Averaging of variables

Step 1 concerns the spatial averaging of valid retrieved values of the individual pixels (x_i) with weights (w_i). Following Miyazaki et al. (2012), Boersma et al. (2016), Rijsdijk et al. (2025) the weight w_i of each measurement is taken equal to the spatial overlap area between the footprint of observation i and the L3 grid cell. This tiling approach results in the following estimate of the gridded L3 column:

$$x_{\mathrm{o,t}} = \frac{\sum_{i=1}^{N}(w_i x_i)}{\sum_{i=1}^{N} w_i}, \tag{1}$$

with N the number of valid L2 observations in the L3 cell.

The spatial average is determined on a (regular) grid on a per-orbit basis. This is possible for any given grid; we provide the dataset created here at spatial resolutions $0.2° \times 0.2°$, $0.5° \times 0.5°$, and $1° \times 1°$. The spatial average for a grid cell is only determined if the combined valid individual retrievals x_i cover at least 30 % of the grid cell, to avoid unrealistic uncertainty estimates in small data samples, following Rijsdijk et al. (2025).

3.2.2 Uncertainty estimate

To assess the overall uncertainty ($\sigma_{\mathrm{o,t}}$) in $x_{\mathrm{o,t}}$ of these spatial averages, we follow the procedure in Rijsdijk et al. (2025) and combine the propagated measurement uncertainties (σ_{m}) and the spatial representativeness uncertainty (σ_{rs}) in quadrature:

$$\sigma_{\mathrm{o,t}} = \sqrt{(\sigma_{\mathrm{m}})^2 + (\sigma_{\mathrm{rs}})^2}. \tag{2}$$

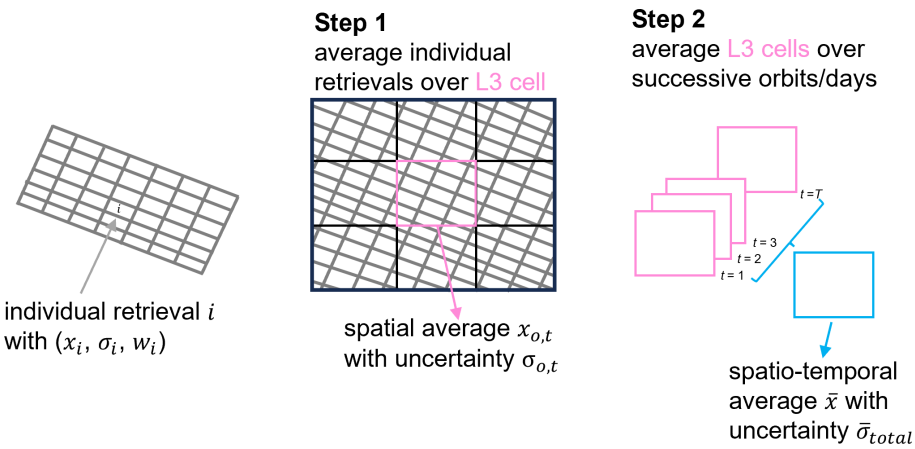

Figure 1. Schematic example of the procedure to calculate spatio-temporal (gridded) averages of retrieved columns and their associated uncertainties. The irregular grey rectangles indicate an ensemble of individual L2 retrievals in a satellite orbit; the black rectangle shows a regular grid to which the individual retrievals are averaged and gridded in Step 1. In Step 2, this procedure is repeated to obtain multiple spatial averages along with their associated uncertainties, which are then averaged over a period to produce a temporal mean and its associated L3 uncertainty.

This reflects that the uncertainty in the spatial average is composed of propagated L2 uncertainties and may also contain uncertainty from incomplete coverage of the grid cell.

Measurement uncertainty

The measurement uncertainty σ_{m} is the combined tropospheric vertical column uncertainty from error contributions ($\sigma_{\mathrm{m,c}}$) in the L2 retrieval, including the uncertainty due to the slant column density measurements ($\sigma_{N_{\mathrm{s}}}$), the uncertainty resulting from errors in the stratospheric column ($\sigma_{N_{\mathrm{s}}^{\mathrm{strat}}}$), and the uncertainty from the air mass factor ($\sigma_{M^{\mathrm{tr}}}$) (Boersma et al., 2004).

$$\sigma_{\mathrm{m}} = \sqrt{(\sigma_{N_{\mathrm{s}}})^2 + (\sigma_{N_{\mathrm{s}}^{\mathrm{strat}}})^2 + (\sigma_{M^{\mathrm{tr}}})^2} \tag{3}$$

Note that these are the contributions of each uncertainty source to the tropospheric vertical column uncertainty, not the source uncertainties themselves. Each of these ($\sigma_{N_{\mathrm{s}}}$, $\sigma_{N_{\mathrm{s}}^{\mathrm{strat}}}$, and $\sigma_{M^{\mathrm{tr}}}$) is represented by Eq. (4) (Sekiya et al., 2022; Rijsdijk et al., 2025), which considers that the random component of the individual uncertainties σ_i tends to cancel out when averaging over many observations (first term on the right-hand side), while a fraction ϕ_{c} of the individual uncertainties σ_i persists after averaging because they come from systematic, spatially correlated contributions to the uncertainties (second term):

$$\sigma_{\mathrm{m,c}} = \sqrt{(1-\phi_{\mathrm{c}})\frac{\sum_{i=1}^{N}(w_i^2\sigma_{i,\mathrm{c}}^2)}{\sum_{i=1}^{N}w_i^2} + \phi_{\mathrm{c}}\frac{(\sum_{i=1}^{N}(w_i\sigma_{i,\mathrm{c}}))^2}{\sum_{i=1}^{N}w_i^2}}. \tag{4}$$

The spatial error correlation factor ϕ_{c} in Eq. (4) quantifies the portion of the error that is systematic, and here we determine its value for each of the measurement error contributions c (slant column density, stratospheric column, and air mass factor) separately. We apply the spatial error correlations as determined by Rijsdijk et al. (2025).

The slant column error is largely uncorrelated over space as it is dominated by measurement noise (Van Geffen et al., 2020; Rijsdijk et al., 2025). There will be a systematic error component in the slant column density due to gaps in knowledge, such as offsets in absorption cross sections, inaccurate ring coefficients in the spectral fit, or the lack of a correction for vibrational Raman scattering (Richter et al., 2011; Zara et al., 2018; Rijsdijk et al., 2025). In Van Geffen et al. (2020) the slant column retrieval noise estimated over a Pacific sector was found to match the DOAS uncertainty estimate to within 15 % for clear-sky pixels, suggesting systematic retrieval errors are roughly a factor of 2 smaller than the random errors. However, the effect of these systematic errors is absorbed in the stratospheric column estimate (as discussed below). Therefore, the uncertainty in slant column density is assumed to be fully random ($\phi_{N_{\mathrm{s}}} = 0$) (Rijsdijk et al., 2025).

In the retrieval method the stratospheric column is determined by assimilating slant column superobservations in the chemical transport model TM5-MP (Williams et al., 2017). The resolution of the TM5-MP model is $1° \times 1°$, and the horizontal correlation length scale used in the assimilation is about 500 km. This is coarser than the spatial average grid sizes of 0.2 to 1.0° considered here. Therefore, it is assumed that the error in the stratospheric column is fully correlated in space with the L3 grid resolution ($\phi_{N_{\mathrm{s}}^{\mathrm{strat}}} = 1$). A spatially correlated error in the slant column would lead to an increased bias in the O–F, increasing the L2 stratospheric column uncertainty. The spatial mean of the stratospheric col-

umn uncertainty in Eq. (3) will thus include the contribution from the systematic slant column uncertainty.

The AMF uncertainty is caused by a combination of components that go into the calculation of the AMF. The a priori NO_2 profile is a large contribution to the AMF uncertainty but is shown to become irrelevant when the averaging kernel is used when comparing with three-dimensional model output (Eskes et al., 2003). Other major components of the AMF uncertainty are the uncertainties in surface albedo and cloud parameters (cloud fraction and cloud pressure). All three of these variables depend on the quality of the climatological surface albedo dataset. Rijsdijk et al. (2025) review the AMF uncertainty by comparing versions 2.3.1 and 2.4 of the NO_2 L2 datasets. The main difference between these two versions is that they apply different surface albedo climatologies. We argue that the differences between albedo estimates from these state-of-science climatologies are indicative of realistic albedo uncertainties. The albedo values directly affect the clear-sky AMF calculations but also drive the retrieved values for the cloud fraction and cloud pressure (see e.g. Riess et al., 2021). Differences between albedo values thus propagate to different values of cloud fraction and cloud pressure, which impact the cloudy-sky AMF values. Albedo differences thus impact the overall AMF calculation, and spatio-temporal characteristics in the albedo differences are expected to lead to spatio-temporal patterns in AMF differences, allowing us to analyse how the AMF uncertainty patterns are correlated in time and space. The effects of changing the surface albedo climatology on the results of the tropospheric NO_2 retrieval are expressed as differences between v2.3.1 and v2.4 tropospheric NO_2 columns. The error in the AMF is shown by Rijsdijk et al. (2025) to be partly correlated depending on the size of the grid cells ($\phi_{M^{\mathrm{tr}}} = e^{-d/l}$, where d is a distance between observations depending on the grid cell size and $l = 35$ km is a typical average correlation length over polluted regions). This spatial correlation is partly due to the low resolution of surface albedo datasets (0.5°) but also because surface modifying conditions are often spatially extensive. For example, droughts impact the surface albedo not just in one grid cell but typically over a larger area. Table 1 provides typical values of the spatial error correlation coefficients for TROPOMI spatial averages at different grid resolutions.

Table 1. Spatial error correlation sources and their random and systematic fractions for TROPOMI NO_2 tropospheric vertical column spatial averages at different spatial resolutions. For the calculation of the AMF coefficients, $\phi_{M^{\mathrm{tr}}} = e^{-d/l}$ has been used with $l = 35$ km the typical correlation length over polluted regions (Rijsdijk et al., 2025) and d the typical grid cell dimension (in km).

Spatial error correlation source	Random fraction (averages out)	Systematic fraction (persists)
Slant column density	$(1-\phi_{N_{\mathrm{s}}}) = 1.00$	$\phi_{N_{\mathrm{s}}} = 0.00^*$
Stratospheric correction	$(1-\phi_{N_{\mathrm{s}}^{\mathrm{str}}}) = 0.00$	$\phi_{N_{\mathrm{s}}^{\mathrm{str}}} = 1.00$
AMF (0.2° × 0.2°)	$(1-\phi_{M_{\mathrm{tr}}}) = 0.44$	$\phi_{M_{\mathrm{tr}}} = 0.56$
AMF (0.5° × 0.5°)	$(1-\phi_{M_{\mathrm{tr}}}) = 0.75$	$\phi_{M_{\mathrm{tr}}} = 0.25$
AMF (1.0° × 1.0°)	$(1-\phi_{M_{\mathrm{tr}}}) = 0.94$	$\phi_{M_{\mathrm{tr}}} = 0.06$

* Note: the systematic component of the slant column density uncertainty is included in the stratospheric uncertainty.

Spatial representativeness uncertainty

The spatial representativeness uncertainty σ_{rs} (molecules cm^{-2}) accounts for incomplete sampling (mainly due to cloud-covered pixels not being included in the spatial averaging) of the cell by the observations available to calculate $x_{\mathrm{o,t}}$ (Rijsdijk et al., 2025). 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 equal to the standard deviation of the tropospheric vertical columns within the cell: 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).

$$\sigma_{\mathrm{rs}} = f\sigma_{x_i}, \tag{5}$$

where σ_{x_i} is the standard deviation in retrievals within the grid cell. The unitless fraction f is calculated as the representativeness of the retrieved observations for the fully covered grid cell:

$$f = \frac{1}{\sqrt{N_{\mathrm{eff}}\alpha_{\mathrm{o,t}} + 1}}\sqrt{1-\alpha_{\mathrm{o,t}}}, \tag{6}$$

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

$$\alpha_{\mathrm{o,t}} = \frac{\sum_{i=1}^{N} w_i}{A}, \tag{7}$$

and N_{eff} is the number of effective observations, which is dependent on the number of available observations, the sensor, the trace gas, the L3 grid resolution, and whether a grid cell is sensitive to systematic sampling (e.g. due to a cloud field covering part of a grid cell) or not (Rijsdijk et al., 2025). The lower the value for N_{eff}, the more sensitive a region is to systematic sampling. N_{eff} has been determined empirically by Rijsdijk et al. (2025), and a summary of the methods used to quantify N_{eff} can be found in Appendix B.

3.3 Temporal averaging

3.3.1 Averaging of variables

The second step in the generation of the L3 dataset consists of temporal averaging, a method newly introduced in this study. We average the spatial-mean values (superobservations) $x_{\mathrm{o,t}}$ obtained in Step 1 over time (e.g. a period of

1 month), again in a weighted fashion to account for different degrees of representativity ($w_{o,t} = 1 - f$). f is high for grid cells with a large representativeness uncertainty and low for superobservations with almost no representativeness uncertainty. This value depends on coverage – higher coverage results in a lower representativeness error – as well as other factors such as sensitivity to systematic sampling. Taking a weighted average like this implies that superobservations with a low representativeness uncertainty obtain more weight in the temporal average than superobservations with a high representativeness uncertainty. This results in the following estimate of the spatially and temporally averaged L3 column $\bar{x}$:

$$\bar{x} = \frac{\sum_{t=1}^{T}(w_{o,t}x_{o,t})}{\sum_{t=1}^{T} w_{o,t}}, \tag{8}$$

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

3.3.2 Uncertainty estimate

Next, the uncertainty associated with $\bar{x}$ (i.e. the monthly mean L3 uncertainty) is determined. We calculate a total spatio-temporal averaged uncertainty which combines the propagated measurement uncertainty ($\bar{\sigma}_m$, including the spatial representativeness uncertainty) and a temporal representativeness uncertainty ($\bar{\sigma}_{rt}$) in quadrature:

$$\bar{\sigma}_{total} = \sqrt{(\bar{\sigma}_m)^2 + (\bar{\sigma}_{rt})^2}. \tag{9}$$

Propagated measurement uncertainty

The spatio-temporally averaged measurement uncertainty $\bar{\sigma}_m$ is the combined uncertainty of the measurement error contributions $\bar{\sigma}_{m,c}$ including the spatial representativeness uncertainty and a priori profile uncertainty:

$$\bar{\sigma}_m = \sqrt{(\bar{\sigma}_{N_s})^2 + (\bar{\sigma}_{N_s^{strat}})^2 + (\bar{\sigma}_{M^{tr}})^2 + (\bar{\sigma}_{rs})^2 + (\bar{\sigma}_{a\ priori})^2}. \tag{10}$$

The propagated slant column density uncertainty ($\bar{\sigma}_{N_s}$), stratospheric column uncertainty ($\bar{\sigma}_{N_s^{strat}}$), air mass factor uncertainty ($\bar{\sigma}_{M^{tr}}$), and spatial representativeness uncertainty ($\bar{\sigma}_{rs}$) are each determined using

$$\bar{\sigma}_{m,c} = \sqrt{(1-\tau_c)\frac{\sum_{t=1}^{T}(w_{o,t}{}^2\sigma_{m,c_t}{}^2)}{(\sum_{t=1}^{T} w_{o,t})^2} + \tau_c\frac{(\sum_{t=1}^{T}(w_{o,t}\sigma_{m,c_t}))^2}{(\sum_{t=1}^{T} w_{o,t})^2}}, \tag{11}$$

where τ_c is the temporal correlation factor of the uncertainty contributions c (not the same as the spatial correlation factor applied in Eq. 4). The last term in Eq. (10) ($\bar{\sigma}_{a\ priori}$) represents the contribution from the uncertainty in the a priori profile shapes and is approximated as 10 % of the tropospheric AMF. 10 % of the AMF is deemed appropriate for the a priori profile uncertainty, as the spatial resolution of the spatial means ($0.2° \times 0.2°$ to $1° \times 1°$) is low compared to the TROPOMI pixel resolution. Earlier studies (gridding OMI to 0.5°) also used an estimate of 10 % (Boersma et al., 2018). $\bar{\sigma}_{a\ priori}$ 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 τ is determined for each of the contributions to measurement uncertainty (slant column density, stratospheric, and AMF) separately. We do so by evaluating to what extent discrepancies in the stratospheric NO_2 column and in the AMF calculation vanish over time.

Firstly, the temporal error 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 it is shown to become irrelevant when the averaging kernel is used when comparing with three-dimensional model output (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 error correlation between the daily observations.

Assuming the AMF error is fully correlated in time would result in an overestimation of the AMF uncertainty in Level 3 products, as in reality the AMF error is only partly systematic and the random part averages out with many observations. Here, we empirically evaluate the systematic character of the AMF uncertainty on different temporal scales, by determining the degree of temporal correlation between AMF errors.

We estimate the temporal error correlation required to determine the AMF uncertainty in spatio-temporal averaged L3 data by comparing versions 2.3.1 and 2.4 of the L2 retrieval, as was done for the spatial error correlation (Sect. 3.2.2). These versions apply different climatological surface albedo datasets, with v2.3.1 applying the albedo derived from OMI and GOME-2 and v2.4 applying albedo derived from TROPOMI spectra (Tilstra et al., 2024). 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. Because the change of albedo climatology is the largest change between the two L2 versions, the difference in tropospheric NO_2 between the versions should be indicative of uncertainty in tropospheric NO_2 resulting from the climatological surface albedo.

We determine the difference in tropospheric NO_2 column in $0.2° \times 0.2°$ superobservations between v2.3.1 and v2.4 L2 data for each day in the period 1 January–31 March and the period 1 June–31 August 2019. For each combination of 2 d within the same period, we determine the Pearson correlation coefficient in tropospheric NO_2 for polluted areas (tropospheric vertical column from both versions >

1.8×10^{15} molecules cm^{-2}) (Fig. 2a):

$$r_{\Delta t} = \frac{\sum_{i=1}^{I}(\Delta N_{t_1 i} - \overline{\Delta N_{t_1}})(\Delta N_{t_2 i} - \overline{\Delta N_{t_2}})}{\sqrt{\sum_{i=1}^{I}(\Delta N_{t_1 i} - \overline{\Delta N_{t_1}})^2 \sum_{i=1}^{I}(\Delta N_{t_2 i} - \overline{\Delta N_{t_2}})^2}}, \quad (12)$$

where $r_{\Delta t}$ is the Pearson correlation coefficient between tropospheric NO_2 differences due to albedo differences and retrieval discrepancies on any 2 d in the time period (blue points in Fig. 2a), $\Delta t = t_2 - t_1$ the time difference, $\Delta N_{t_1} = N_{t_1}^{v2.4} - N_{t_1}^{v2.3.1}$ and $\Delta N_{t_2} = N_{t_2}^{v2.4} - N_{t_2}^{v2.3.1}$ the difference in tropospheric NO_2 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 difference" between daily superobservations, we determine the mean correlation ($\overline{r_{\Delta t}}$; black points in Fig. 2a). Then we take the mean of the $\overline{r_{\Delta t}}$ for the time differences 1–30 d. This results in $\tau = 0.29$ for 1 January to 31 March and $\tau = 0.30$ for 1 June to 31 August, suggesting that tropospheric vertical column errors only partially persist in time. The temporal error correlation in the AMF is set to be $\tau_{M^{tr}} = 0.30$ for both the NH winter and summer period. When analysing longer time differences than a month, the correlation disappears. The method was repeated for $1° \times 1°$ superobservations, resulting in similar results, giving us confidence that the temporal AMF error correlation coefficient is not grid-size-dependent.

Next, we estimated temporal correlation factors for the stratospheric uncertainty. 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 investigating the difference between the forecasted column and the observed column (O–F) over unpolluted regions; 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 2 d within the same month, we determine the Pearson correlation coefficient in stratosphere O–F for clean areas (tropospheric vertical column $< 0.6 \times 10^{15}$ molecules cm^{-2}) (Fig. 2b). This is again done with Eq. (12) but now with $\Delta N_{t_1} = N_{t_1}^{observation} - N_{t_1}^{forecast}$ and $\Delta N_{t_2} = N_{t_2}^{observation} - N_{t_2}^{forecast}$ the difference in stratospheric NO_2 between observation and forecast for the first day and second day respectively. For each "time difference" between daily O–F maps, we determine the mean correlation ($\overline{r_{\Delta t}}$). Then we take the mean of $\overline{r_{\Delta t}}$ for the time differences 1–30 d. This results in $\tau = 0.30$ for the NH winter period of 1 January to 31 March. The same method is carried out for the NH summer period 1 June–31 August, which showed a slightly lower correlation coefficient of $\tau = 0.21$. We take the higher value of $\tau_{N_s^{strat}} = 0.30$ 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 altogether for periods longer than 3 months.

The uncertainty in the slant column density is assumed to be fully uncorrelated in time ($\tau_{N_s} = 0$), following the same arguments as in the spatial uncertainty correlation (see Sect. 3.2.2).

The temporal correlation coefficient of the spatial representativeness uncertainty is set at fully random ($\tau_{rs} = 0$). The representativity can be correlated through time due to the effect of persistent cloud cover, but this will be assessed separately in the temporal representativeness uncertainty.

The monthly mean L3 NO_2 columns are only representative when sufficient observations are available for calculating the means. The L3 qa_value (qa_L3 in the final data product – Table C1) is set to equal 1 when the data are representative as a monthly mean and equal 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 (qa_L3) is set at 0, and the L3 NO_2 columns are advised not to be used for further analyses, as these would not be a good representation of the given month.

Temporal representativeness uncertainty

The temporal representativeness uncertainty ($\bar{\sigma}_{rt}$) can be interpreted 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 true monthly mean is obtained with at least one observation (note that the TROPOMI orbits overlap at higher latitudes, where multiple observations per day are possible) each day. A finite population correction is applied to the method of calculating the standard error to correct for this (Bondy and Zlot, 1976; Isserlis, 1918; Rijsdijk et al., 2025), resulting in the standard error decreasing to zero when observations are available for every day of the month. This results in the following theoretical formula for the temporal representativeness uncertainty:

$$\bar{\sigma}_{rt} = \frac{\sigma_{x_{o,t}}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}, \quad (13)$$

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 $\sigma_{x_{o,t}}$ is determined using

$$\sigma_{x_{o,t}} = \sqrt{\frac{\sum_{t=1}^{T}(w_{o,t}(x_{o,t} - \bar{x})^2)}{(T-1)\sum_{t=1}^{T} w_{o,t}}}, \quad (14)$$

with T the total number of valid spatial averages in the period over which the averaging is performed, and $w_{o,t} = 1 - f$.

We apply the method also used in Rijsdijk et al. (2025) and Appendix B to assess whether Eq. (14) is suitable for assessing the temporal representativeness uncertainty. We start

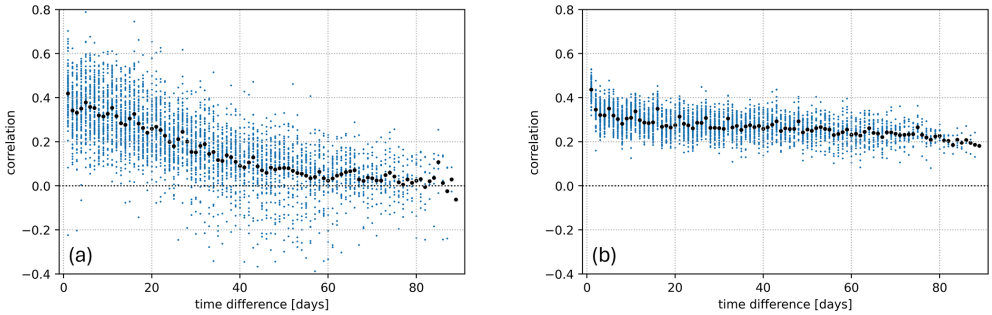

Figure 2. Temporal correlation in **(a)** difference of tropospheric NO_2 column for v2.3.1 and v2.4 for polluted areas (AMF uncertainty) and **(b)** observation forecast for unpolluted areas (stratospheric column uncertainty) between each combination of 2 d within the same month for the period 1 January–31 March 2019. Larger black points are the mean of correlations with the same time difference.

Table 2. Temporal error correlation sources and their random and systematic fractions for TROPOMI NO_2 tropospheric vertical column spatio-temporal averages.

Temporal error correlation source	Random fraction (averages out)	Systematic fraction (persists)
Slant column density	$(1-\tau_{N_s}) = 1.00$	$\tau_{N_s} = 0.00^*$
Stratospheric correction	$(1-\tau_{N_s^{str}}) = 0.70$	$\tau_{N_s^{str}} = 0.30$
AMF	$(1-\tau_{N_{M_{tr}}}) = 0.70$	$\tau_{N_{M_{tr}}} = 0.30$

* Note: the systematic component of the slant column density uncertainty is included in the stratospheric uncertainty.

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 representativeness uncertainty can be quantified by comparing the mean of this completely covered grid cell to the mean of several samples (a subset of days) taken from that grid cell. We start by sampling a single day, which we use to estimate the temporal 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 representativeness uncertainty f_t and temporal coverage of the grid cell (Fig. 3a). This process is then repeated for 100 grid cells in polluted areas (Fig. 3b). We only take polluted grid cells into account (mean NO_2 tropospheric vertical column of $\geq 2\times10^{15}$ molecules cm^{-2}), as these are expected to have the highest temporal variability and thus are more likely to be sensitive to a sampling bias. The mean of the experiments (green) is almost identical to the theoretical solution of Eq. (13) (blue), showing that the formula is a suitable method for quantifying the temporal representativeness uncertainty for random sampling. Similar results are found for June 2019.

As with the spatial representativeness (cloud masking one part of the grid cell), there might be a sensitivity to systematic sampling temporally. This could for example be the case if large weather systems gave 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 $\sim 50\,\%$ 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 non-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 look into 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 1 random day and then repeatedly add adjacent days). Figure 3d shows that systematic sampling produces more variability of grid cells around the theoretical solution (blue line, Eq. 13) than with random sampling (Fig. 3b) but that the mean of the experiments (green line) is close to the theoretical solution. Comparison with the spatial representativeness curve in Fig. B1b shows that the experimental curves in Fig. 3d are closer to the theoretical solution. The fit from Eq. (6) 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 system-

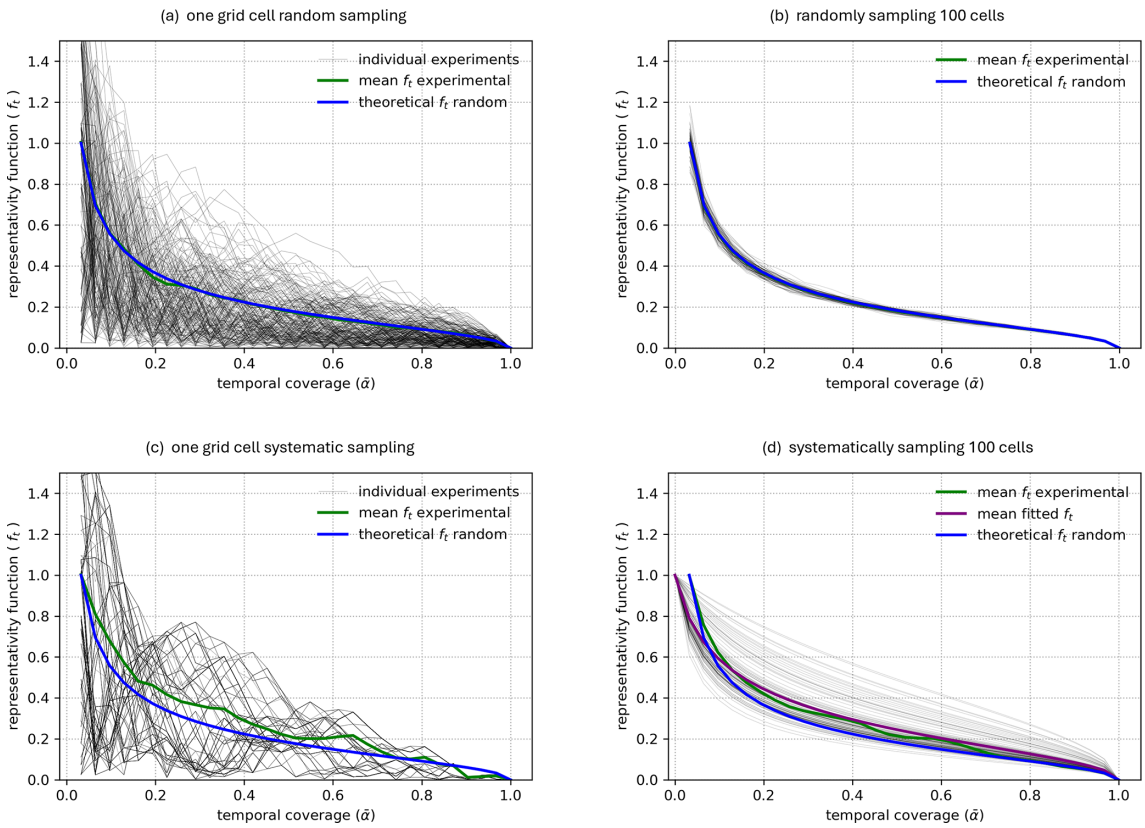

Figure 3. Results of repeatedly sampling grid cells to calculate $\bar{\sigma}_{rt}$. **(a)** Repeated random sampling of 31 d of a single grid cell with values $x_{o,t}$. Daily observations in the grid cell are randomly sampled 200 times. The thin grey lines represent the difference between the sampled 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 Eq. (13). **(b)** Results of randomly sampling 100 grid cells. The grey lines are the green line from **(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. Observations in all panels are from January 2019 and have a mean tropospheric vertical column of $\geq 2 \times 10^{15}$ molecules cm^{-2}. Only grid cells with observation available every day are sampled.

atic sampling, and the uncertainty can be assessed with the theoretical solution suggested in Eq. (13).

It should be noted that the sampling method can only be applied to grid cells that provide coverage for all days of the month. This is essentially only true for arid areas with almost no cloud cover. The question arises of how representative these regions are. We filter for relatively polluted regions ($\geq 2 \times 10^{15}$ molecules 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 dry season. We assume these polluted areas are representative of polluted areas in other parts of the world, and the theoretical solution can be applied globally.

4 TROPOMI NO_2 Level 3 dataset

4.1 Dataset

The here presented ESA CCI+ TROPOMI Level 3 dataset is available for the period May 2018 to December 2021 at a monthly resolution. The dataset is available on a global regular grid at different spatial resolutions: $0.2° \times 0.2°$, $0.5° \times 0.5°$, and $1° \times 1°$. The datasets are organised into a user-friendly and self-describing netCDF-4 format, following CF metadata conventions. The dataset represents the NO_2 columns at satellite overpass time ($\sim$ 13:30 LT) under mostly clear-sky conditions.

The dataset contains both the tropospheric and stratospheric vertical column density (see Appendix C for all variables). The dataset contains two estimates of the total uncertainty for the tropospheric vertical column $\bar{\sigma}_{total}$: one including the a priori uncertainty and one version excluding the a priori uncertainty, which can be used when applying the averaging kernel. A spatio-temporal average of the tropospheric averaging kernel from the L2 dataset is also available in the L3 dataset. For assessing the L3 data quality, the L3 qa_value, count variable, and average cloud radiance fraction are also available.

This dataset can be used for, for example, temporal analysis, emission monitoring, data assimilation and model vali-

dation, and atmospheric chemistry studies. It should be noted that this dataset is a clear-sky average, and model comparisons should only be done on cloud-free sampled model data, as previously recommended in Boersma et al. (2016). The cloud-free sampling of data causes biases with the true monthly mean in multiple ways. Firstly, photochemical reactions cause NO_2 to break down faster in clear-sky conditions (Miller et al., 1956), so the true monthly mean of tropospheric columns (when including vertical columns in overcast conditions) is higher by up to 25 % (Geddes et al., 2012; Boersma et al., 2016). Secondly, incident cloudiness over more or less polluted areas would cause uneven sampling, causing a bias either way. Lastly, cloudiness could be connected to atmospheric transport conditions that may bring clean or polluted conditions. The uncertainty assessment here represents the uncertainty of the L3 product to the clear-sky average and does not include an uncertainty to the true monthly mean.

4.2 Analysis of dataset

The results of the spatio-temporal average monthly mean and uncertainties for NO_2 for January and June 2019 are shown in Fig. 4, with local values given in Table 3. Missing data are mainly due to polar night at high latitudes and filtering of low-quality L2 data, due to, for example, consistent cloud or snow cover. All results presented in this section are from the $0.2° \times 0.2°$ spatial resolution dataset.

In June the tropospheric NO_2 columns are high in Africa due to wildfires and biomass burning in the dry season. In the urbanised areas in the Northern Hemisphere, tropospheric NO_2 columns have higher values during the winter season, reflecting longer lifetimes (e.g. Shah et al., 2020) and higher emissions due to heating, and thus we see higher values over Amsterdam and Beijing in January than in June (Table 3, Fig. 6). The opposite is true for urbanised areas in the Southern Hemisphere, for example Rio de Janeiro (Table 3, Fig. 6). The total uncertainty in winter when not applying the averaging kernel attributes to 21 % and 17 % of the total tropospheric vertical column for Amsterdam and Beijing respectively. These percentages are 19 % and 14 % respectively when applying the averaging kernel. Due to the averaging of random errors, the L3 uncertainty is lower than the uncertainty in L2 orbits, despite the introduction of representativeness errors. The average relative uncertainty in valid L2 pixels in January 2019 in Amsterdam is 52 %, compared to the 21 % in the L3 dataset. In Beijing the average relative L2 uncertainty is 28 %, compared to the 17 % in L3 in January 2019.

The separate uncertainty sources are shown in Fig. 5 for January 2019. Over unpolluted areas the largest source of uncertainty comes from the estimation of the stratospheric column concentration (Fig. 5b). In polluted regions the largest error source is the AMF uncertainty (Fig. 5c). Because the slant column density uncertainty is assumed random over both space and time, the uncertainty averages out over a large number of observations and is very small in the L3 dataset (Fig. 5a), but it should be noted that the systematic component of the slant column density uncertainty is included in the stratospheric uncertainty estimate. The spatial and temporal representativeness uncertainty is largest in Europe and East Asia, where the standard deviation of the tropospheric vertical column is largest but is a minor component in the total uncertainty budget.

Time series of the tropospheric NO_2 column for four polluted locations are shown in Fig. 6. The available time series is too short to show long-term changes. The time series in Beijing and Amsterdam show features linked to the pandemic lockdowns (2020–2021) (Bauwens et al., 2020) and indicate an overall reduction of tropospheric NO_2 columns over the 5 years, although Amsterdam displays a large variability and the decrease may not be significant.

The Global Climate Observing System (GCOS) states requirements for observational datasets of the ECVs, including the precursors for aerosol and ozone variable NO_2 tropospheric vertical column (World Meteorological Organization (WMO) et al., 2022a). We examine the uncertainties in the NO_2 tropospheric vertical column against the GCOS-required measurement uncertainty (Fig. 7), which is formulated as a threshold, breakthrough, and goal value (World Meteorological Organization (WMO) et al., 2022b). The threshold requirement, the minimum requirement to be met to ensure that data are useful, requires the relative uncertainty to be lower than 100 % and the absolute uncertainty to be less than 5×10^{15} molecules cm^{-2}. The breakthrough requirement represents a significant improvement and requires the relative uncertainty to be lower than 40 % and the absolute uncertainty to be less than 2×10^{15} molecules cm^{-2}. Lastly, the goal is reached when the relative uncertainty is lower than 20 % and the absolute uncertainty is less than 1×10^{15} molecules cm^{-2} (World Meteorological Organization (WMO) et al., 2022b). Due to the averaging of uncorrelated uncertainties in creating the L3 dataset, the GCOS requirements are met more frequently in L3 than in L2 (Fig. 7), suggesting L3 data are useful for climate monitoring. In June 2019 over Beijing the absolute uncertainty (as well as relative uncertainty) drops from an average of 2.81×10^{15} molecules cm^{-2} (27.9 %) in L2, with most pixels reaching the threshold requirement, to 1.72×10^{15} molecules cm^{-2} (15.5 %) in L3, well within the breakthrough requirement (Fig. 7c). Over Amsterdam the absolute uncertainty (as well as relative uncertainty) drops from an average of 1.44×10^{15} molecules cm^{-2} (32.9 %) in L2, with most pixels reaching the breakthrough requirement, to 0.78×10^{15} molecules cm^{-2} (16.9 %) in L3, within the goal requirement (Fig. 7d).

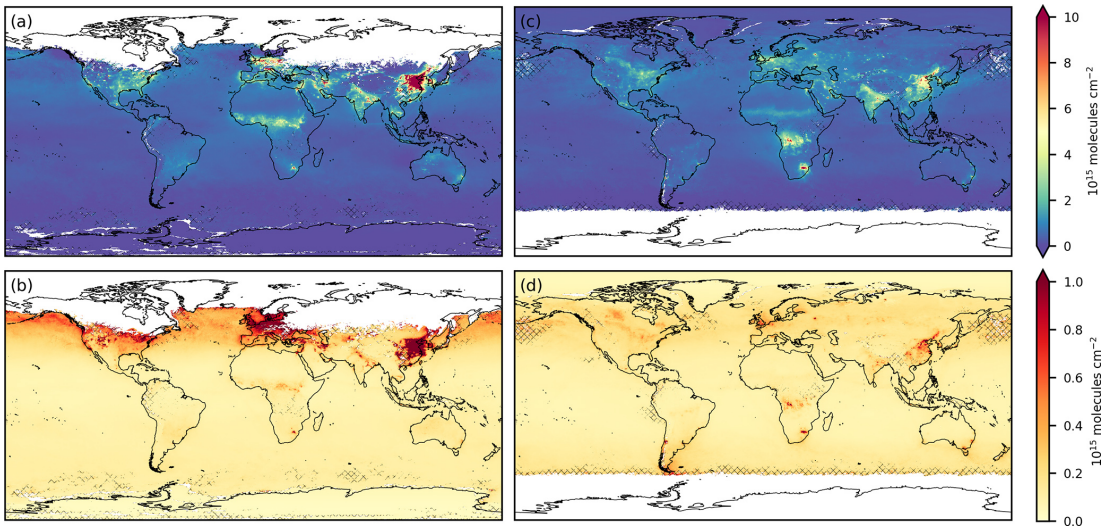

Figure 4. Example of results of the monthly **(a)** tropospheric NO_2 column with its **(b)** total uncertainty ($\bar{\sigma}_{total}$) for January 2019 and the same **(c, d)** for June 2019. Monthly means are calculated from individual L2 retrievals with good quality (qa_value = 0).

Table 3. Example values for monthly mean tropospheric NO_2 vertical column and uncertainties (with percentage of mean tropospheric vertical column) for a few selected locations for January and June 2019 (units: 10^{15} molecules cm^{-2}). The values for Amsterdam, Beijing, and Rio de Janeiro were taken from one grid cell in the city centre location (52.3° N, 4.9° E for Amsterdam; 39.9° N, 116.3° E for Beijing; and 22.9° S, 43.1° W for Rio de Janeiro). 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° W 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	Beijing	Rio de Janeiro	Africa biomass burning
$\bar{x}$	Jan	5.45	32.28	2.87	7.20
	Jun	4.58	11.04	4.54	14.63
$\bar{\sigma}_{total}$	Jan	1.15 (21 %)	5.50 (17 %)	0.38 (13 %)	1.03 (14 %)
	Jun	0.78 (17 %)	1.72 (16 %)	0.59 (13 %)	2.10 (14 %)

5 Validation

As independent validation, the ESA CCI+ TROPOMI Level 3 NO_2 dataset was compared to ground-based remote sensing measurements using the same reference data and methodologies used for the validation of the underlying Level-2 data reported in Verhoelst et al. (2021), in Van Geffen et al. (2022a), and in the S5P ATM-MPC quarterly Routine Operations Consolidated Validation Reports (ROCVR, available at https://mpc-vdaf.tropomi.eu/, last access: 30 August 2025). However, the monthly gridded nature of the L3 data does imply a need for some adaptations in the comparison methodology. The specific aspects and validation results are reported below, per (sub-)column. The L3 TROPOMI data were all filtered using the L3 qa_value (qa_L3).

5.1 Stratospheric slant column

The L3 TROPOMI stratospheric NO_2 columns were compared to the consolidated LATMOS_v3 sunset SAOZ measurements (Pommereau and Goutail, 1988) obtained at 10 sites covering mostly clean sites distributed globally from the southern high latitudes up to the northern high latitudes. These twilight zenith-sky measurements (Solomon et al., 1987) have an estimated uncertainty of about 10 %–14 % (Yela et al., 2017; Bognar et al., 2019), and they were adjusted to the average TROPOMI overpass time of the monthly averages (represented in the file as `eff_frac_day`) using a model-based photochemical adjustment (Hendrick et al., 2004). For optimal spatial co-location, we compared the SAOZ measurements to the TROPOMI 1.0° × 1.0° L3 grid cell covering the centre of the SAOZ observation. This procedure accounts for the large horizontal smoothing and offset in the SAOZ measurement sensitivity towards the setting sun (Lambert et al., 1997; Verhoelst et al., 2015). An illustration of such a comparison, at the Observatoire de Haute Provence in France, is shown in Fig. 8. This comparison shows excellent agreement, within the uncertainties of each product.

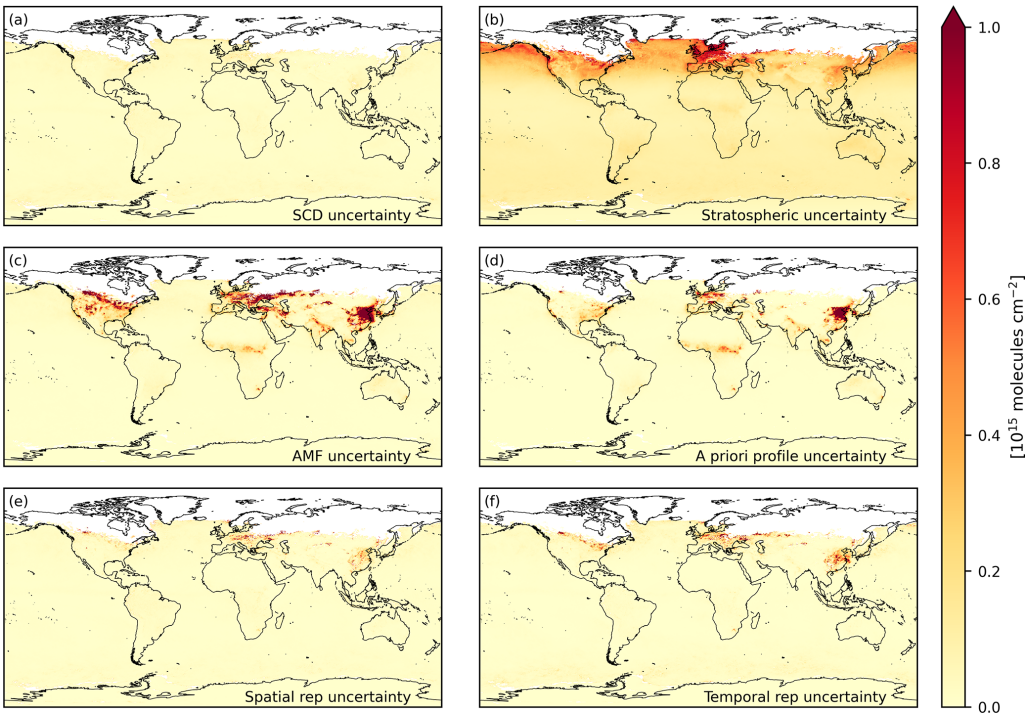

Figure 5. All components of tropospheric vertical column density uncertainty using the discussed spatial and temporal correlations for January 2019 (total uncertainty in Fig. 4b).

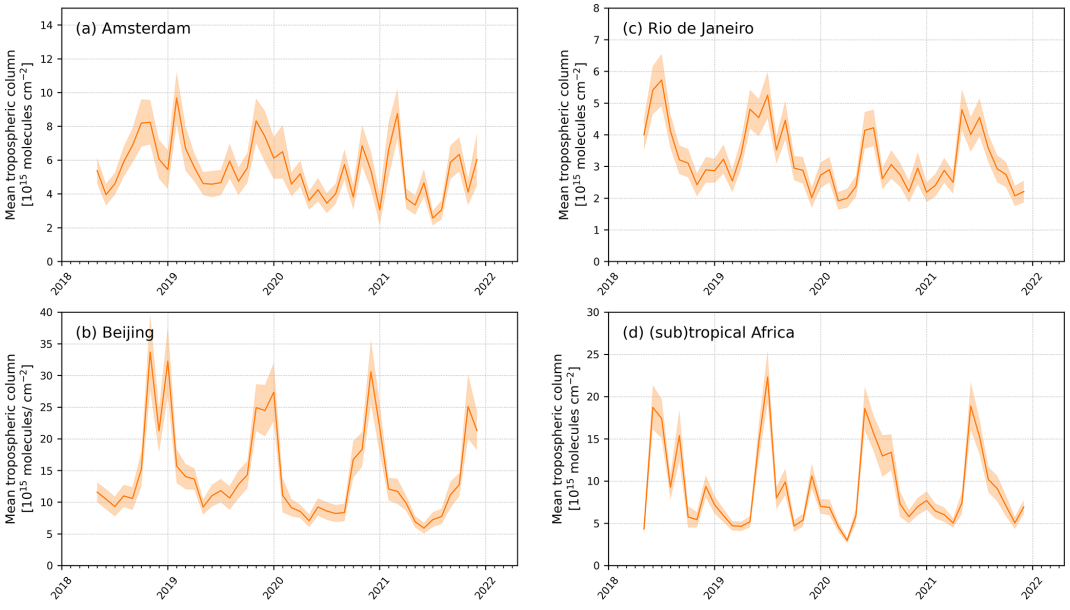

Figure 6. Time series of the monthly mean tropospheric vertical column NO_2 with the total uncertainty ($\bar{\sigma}_{\text{total}}$) for the same locations as in Table 3. The values in the time series represent a single L3 grid cell value, and the associated total uncertainty is given as the shaded area.

The network-wide results are summarised in Fig. 9. Overall, these stratospheric column comparisons yield results very similar to those for the underlying L2 (Verhoelst et al., 2021), with a virtually insignificant network-mean bias and a dispersion of typically around 0.2×10^{15} molecules cm^{-2}. The somewhat larger bias and dispersion over Paris is probably related to tropospheric contamination in the SAOZ measurements.

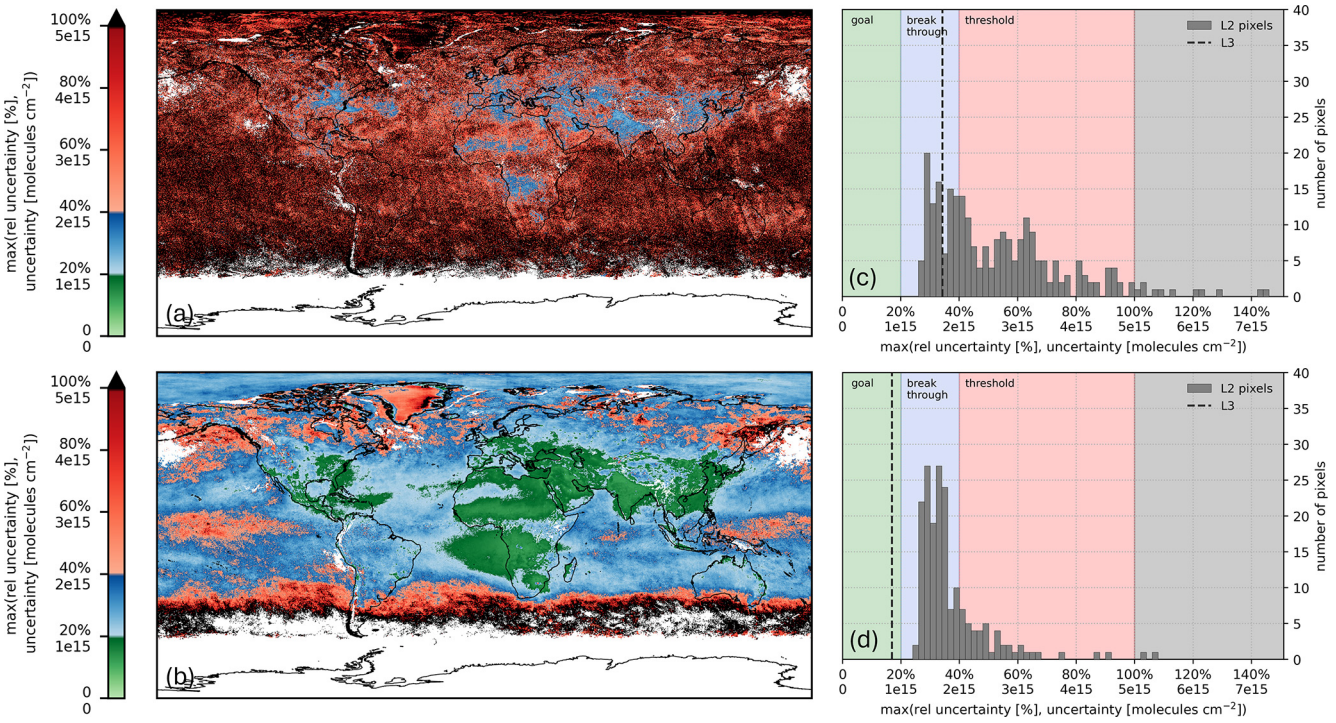

Figure 7. June 2019 **(a)** L2 and **(b)** L3 dataset as reviewed by the GCOS requirements for measurement uncertainty. Black regions do not fulfil the GCOS requirements, red regions fulfil "threshold" requirements, blue regions fulfil "breakthrough" requirements, and green regions fulfil "goal" requirements. The histograms show the L2 pixel uncertainties that fall into a single 0.2° × 0.2° L3 grid cell for **(c)** Beijing and **(d)** Amsterdam and the grid cell L3 total uncertainty (vertical line).

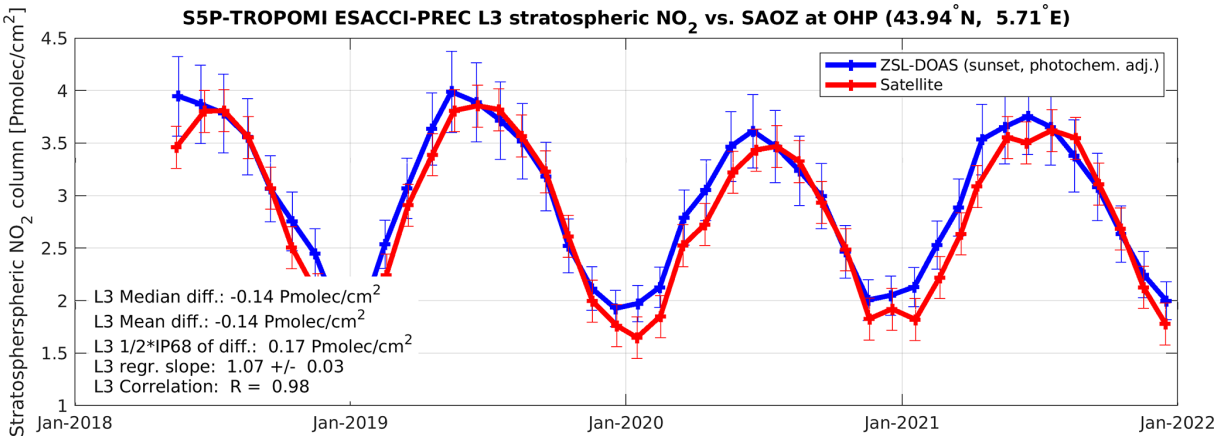

Figure 8. Time series of co-located L3 TROPOMI and photochemically adjusted sunset SAOZ stratospheric NO_2 column measurements at the Observatoire de Haute Provence (France).

5.2 Tropospheric vertical column

The L3 TROPOMI tropospheric NO_2 vertical columns were compared to the MAX-DOAS tropospheric column data (Hönninger and Platt, 2002) collected from various sources and harmonised (in terms of file format) through the NIDFORVAL project for the operational S5P L2 validation. Total uncertainty estimates on these tropospheric vertical column density measurements are of the order of 7 %–17 % in polluted conditions, including both random (around 3 % to 10 %, depending on the instrument) and systematic (11 % to 14 %) contributions (e.g. Hendrick et al., 2014). MAX-DOAS data obtained within 30 min of the underlying L2 S5P data were compared to 0.2° × 0.2° L3 grid cells covering the station location. To ensure good temporal representativeness, only those sites at which at least a full year of comparisons could be made were retained. This yielded eight sites, cov-

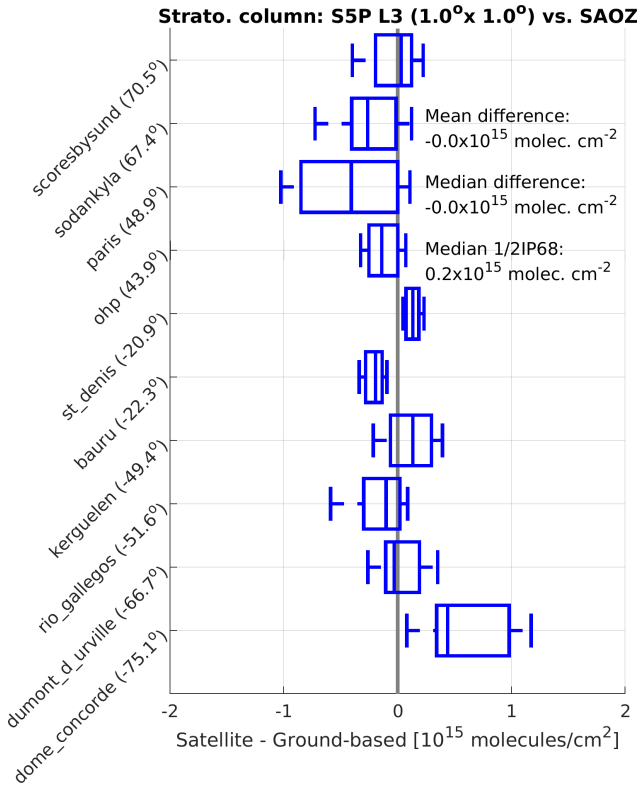

Figure 9. Box-and-whisker plots summarising from pole to pole the bias and spread of the difference between the L3 TROPOMI and ground-based SAOZ data (2018–2022). The median difference is represented by a vertical solid line inside the box that marks the 25 % and 75 % quantiles. The whiskers cover the 9 %–91 % range of the differences. Values between brackets in the labels denote the latitude of the station.

ering moderately to severely polluted conditions. The most polluted case, Xianghe in China (e.g. Pinardi et al., 2020), is analysed in Fig. 10, revealing a slightly better correlation and smaller dispersion of the differences than observed in the validation of the underlying L2 data. This is most likely due to the reduced measurement and atmospheric noise in the monthly averages. The mean and median difference are however more negatively biased, which can be attributed the poorer spatial resolution of the L3 dataset compared to the constituent L2 data. This is especially notable in the winter months, which due to pollution have high spatio-temporal variability in the L2 orbits and is smeared out in the averaging for the L3 data. There are also unresolved issues in wintertime MAX-DOAS retrievals done under very low viewing zenith angles in situations with thin boundary layers and high NO_2 concentrations, which could contribute to the bias (Cifuentes et al., 2025).

The network-wide results for the MAX-DOAS comparisons are visualised in Fig. 11. These confirm the relatively small dispersion of the differences, typically about 12 %, which is well within the combined prognostic uncertainty budget (quadratic sum of 20 % uncertainty on the TROPOMI L3 data and probably at least 10 % on the monthly averaged MAX-DOAS data, depending on how one propagates the systematic and random components of the MAX-DOAS uncertainty). Also confirmed is the strong negative bias in TROPOMI NO_2, to be understood as a combination of the L2 negative bias and additional systematic differences related to the spatial smearing in the L3 data. Unexpected is the more pronounced negative bias at the relatively clean sites, which is opposite to the behaviour observed for the underlying L2 data. This result is not confirmed by the total column comparisons described in Sect. 5.3 and may be a case of small-number statistics or peculiarities at these individual sites.

5.3 Total column

The TROPOMI L3 total NO_2 columns, calculated as the sum of the tropospheric and stratospheric vertical columns provided in the data files, were compared to Pandora direct-sun measurements (v1.8) from the Pandonia Global Network (PGN, https://pandonia-global-network.org/, last access: 30 August 2025). These measurements have a random error uncertainty of about 0.27×10^{15} molecules cm^{-2} and a systematic error uncertainty of 2.7×10^{15} molecules cm^{-2} (Herman et al., 2009). All Pandora measurements satisfying the recommended PGN quality filtering and obtained within 30 min from the satellite effective fractional day (`eff_frac_day` in the data files) were averaged, and only the $0.2° \times 0.2°$ L3 grid cell covering the instrument location was used. To ensure good temporal representativeness, only those sites at which at least a full year of comparisons could be made were retained. This yielded eight sites, covering both unpolluted rural conditions and severe pollution, e.g. in the megacity of Mexico City. Figure 12 demonstrates the agreement at this most polluted site for both the L3 product and the current operational L2 product, which is the v2.4 full mission reprocessed (RPRO) dataset for this period of S5P measurements.

As for the tropospheric vertical column density comparisons, the correlation between TROPOMI L3 and Pandora data is excellent ($r \sim 0.9$) for these comparisons in a highly polluted environment and even better than at L2. Again, as for the tropospheric column comparisons, the NO_2 underestimation already seen in L2 is more pronounced for the L3 NO_2 product with a median difference of about −39 % (versus −17 % for L2) and a regression slope of 0.54 (versus 0.62 for L2). For a pollution hot spot such as Mexico City, this is most likely due to the coarse L3 product resolution (0.2°), which tends to spatially smear out gradients compared to the pixel-resolution L2 data (0.05°). On the other hand, thanks to the temporal averaging, the dispersion of the differences between L3 and PANDORA NO_2 is significantly reduced, from nearly 20 % down to about 10 %. The network-wide results are summarised in Fig. 13. The tendency that TROPOMI NO_2 shows a more pronounced underestimation

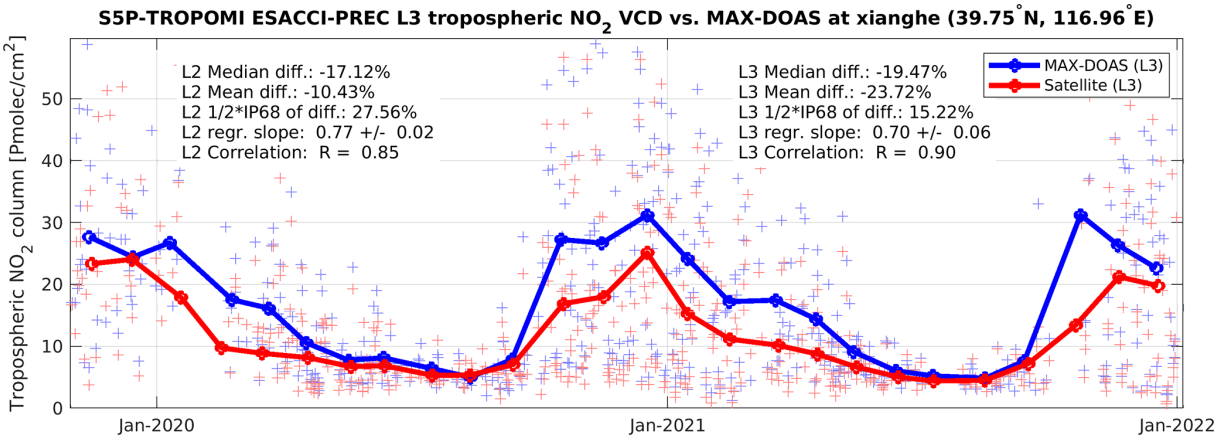

Figure 10. Time series of co-located TROPOMI and MAX-DOAS tropospheric NO_2 vertical column density (VCD) measurements at Xianghe (China). Light markers represent the co-located L2 data from the current operational processor (v2.4 RPRO, as used for ROCVR #24); the solid lines represent the L3 data.

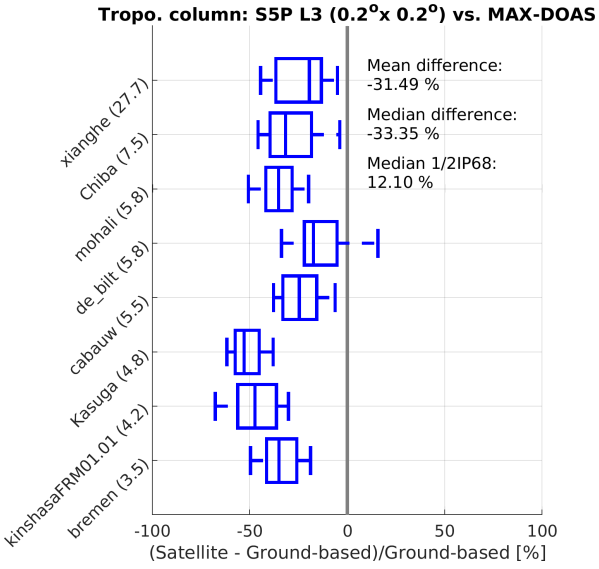

Figure 11. Similar to Fig. 9 but representing the agreement between the S5P-TROPOMI and MAX-DOAS tropospheric NO_2 column measurements, limited to those sites for which a full year of comparisons is available. Sites are ordered by mean tropospheric NO_2 vertical column density: cleaner sites at the bottom, more polluted sites at the top.

for larger total NO_2 column values is in line with the L2 validation results, which show virtually no underestimation (or even a very slight overestimation) at the cleanest sites, where the NO_2 column is dominated by the stratospheric contribution, and an underestimation up to 20 %–30 % at the most polluted sites (Verhoelst et al., 2021, and updates in the ROCVR). These total column results therefore do not confirm the more pronounced negative biases observed in the MAX-DOAS comparisons at sites with only moderate pollution (when compared to more polluted sites).

6 Code and data availability

Data described in this article can be accessed at https://doi.org/10.21944/CCI-NO2-TROPOMI-L3 (KNMI, 2025). The generated L3 dataset will be made available on the ESA Climate Change Initiative Open Data portal. The software to create the spatial average is available at https://doi.org/10.5281/zenodo.10726644 (Rijsdijk and Eskes, 2024). Software to generate the temporal–spatial mean resulting in the L3 dataset is available at https://doi.org/10.5281/zenodo.14505524 (Glissenaar, 2024). The TROPOMI L2 v2.3.1 NO_2 dataset, which is the input for the generated dataset, is not the operational version and thus no longer publicly available. A sample of the data can be shared upon request. TROPOMI NO_2 L2 v2.4 is available on the Copernicus Data Space Ecosystem (https://doi.org/10.5270/S5P-9bnp8q8, Copernicus Sentinel-5P, 2021b). The reference data used for the ground-based validation are available from the NDACC Data Host Facility at https://ndacc.larc.nasa.gov/ (NDACC, 2025) (SAOZ and selected MAX-DOAS data) and from both the PGN website (https://www.pandonia-global-network.org/, Pandonia Global Network (PGN), 2025) and ESA's Validation Data Center (ESA Validation Data Center (EVDC), 2025, https://evdc.esa.int/) for the Pandora data.

7 Conclusions

We developed a comprehensive Level 3 (L3) dataset of gridded and averaged tropospheric NO_2 column retrievals (v2.3.1) from the TROPOMI sensor, covering the period from May 2018 to December 2021. The dataset is available at multiple spatial resolutions ($0.2° \times 0.2°$, $0.5° \times 0.5°$, and $1° \times 1°$) on a monthly timescale. Using a tiling (superobservation) approach, we ensured that (1) valid NO_2 retrievals

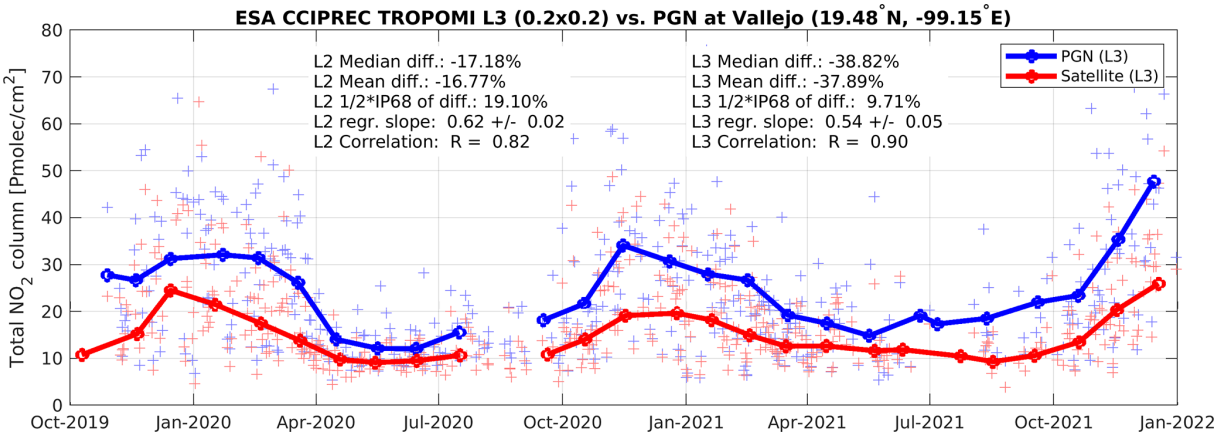

Figure 12. Time series of co-located TROPOMI and PGN total NO_2 column measurements at Vallejo (Mexico City). Light markers represent the co-located L2 data from the current operational processor (v2.4 RPRO, as available on the ATM-MPC VDAF AVS); the solid lines represent the L3 data.

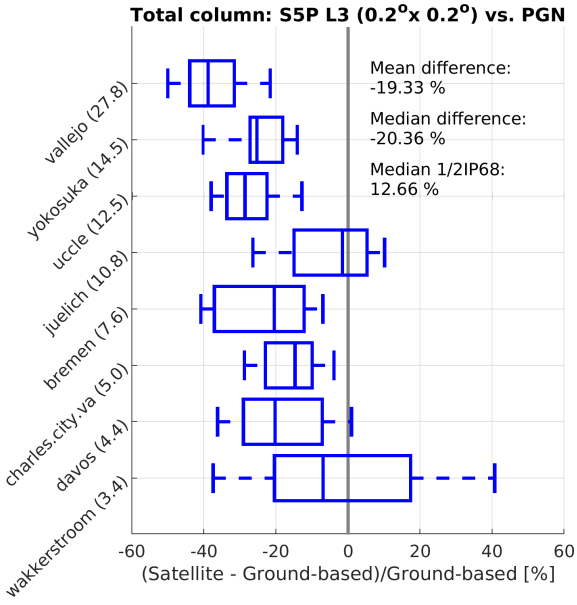

Figure 13. Similar to Fig. 9 but representing the agreement between the S5P-TROPOMI and PGN total NO_2 column measurements, limited to those sites for which a full year of comparisons is available. Sites are ordered by mean total NO_2 vertical column density: cleaner sites at the bottom, more polluted sites at the top.

are gridded according to pixel area, and (2) temporal averages are calculated as weighted means, with weights reflecting the representativeness of superobservations for specific days. This L3 dataset is applicable for clear-sky or low-cloud-fraction conditions at TROPOMI's overpass time of 13:30 LT.

Realistic uncertainties were derived by propagating L2 retrieval uncertainties. To address spatial and temporal error correlations, we considered errors in the stratosphere–troposphere separation and air mass factor calculations, which are partly spatially correlated. Representativeness uncertainty, stemming from incomplete coverage of grid cells by valid L2 retrievals (e.g. due to cloud cover), was found to be more significant in polluted regions. The combined L3 uncertainty accounts for both measurement uncertainties (including spatial representativeness errors) and temporal representativeness uncertainties. An analysis showed that 30 % of retrieval uncertainties persist over a month due to error correlations in the stratosphere–troposphere separation and air mass factor calculations.

Over polluted areas, the L3 dataset showed reduced uncertainties compared to averaged L2 retrievals, demonstrating the effectiveness of averaging large numbers of observations. For example, while monthly average L2 uncertainties are 30 %–50 % over Beijing and Amsterdam, they drop to 20 % or less in the L3 dataset, meeting the GCOS "breakthrough" and even "goal" requirements. Validation against ground-based PANDORA measurements revealed a consistent temporal correlation but with a low bias of 20 %, partly attributed to more pronounced spatial smearing in the L3 product.

This monthly mean L3 TROPOMI tropospheric NO_2 dataset offers a coherent and much-reduced (in size) data record, making it suitable for atmospheric chemistry studies, for evaluating atmospheric models, and for analysing spatiotemporal NO_2 trends. The methods presented here can be replicated when creating L3 datasets for other atmospheric gases and other Earth observation L3 datasets. Future work will include applying the presented methods on NO_2 retrievals from OMI and combine TROPOMI and OMI observations to create a decades-long consistent CDR.

Appendix A: L2 observations in the descending node

During polar summer, the polar regions are marked by 24 h daylight. In the Northern Hemisphere, this means that the polar region on the "backside" of the Earth is experiencing daylight from May to August, and observations are made in the descending node of the satellite. These observations in the descending node are included in the L2 NO_2 product and not quality-flagged in the qa_value before version 2.7.

However, we find that the observations in the descending node of the orbit show structurally different retrieval results for the tropospheric vertical NO_2 column in the Arctic region (north of 60° latitude) than the observations from the ascending node in June (Fig. A1a). While retrievals in the ascending part of the orbit show a median value of 0.2×10 molecules cm^{-2}, retrievals in the descending part show a long tail of negative tropospheric NO_2 values, and the mode of the probability density curve is negative. This difference between results from the ascending and descending part of the orbits is less obvious in the Antarctic region in December (Fig. A1b).

A TROPOMI validation report (Lambert et al., 2023) has shown that TROPOMI underestimates tropospheric NO_2 columns compared to ground-based Pandora instruments at the high-latitude locations of Ny-Ålesund and Eureka by about 15 %. The discrepancy between tropospheric NO_2 column observations from the descending node and ascending node could play a role in this underestimation.

The discrepancy between retrievals done in the descending and ascending node is caused because the descending node observations are not included in the data assimilation in TM5-MP, and thus the stratospheric columns are more uncertain and often overestimated (Fig. A1c). Based on these results, we recommend to not use retrievals from the descending part of the orbit. From TROPOMI L2 version 2.7 onwards, the qa_value is adjusted to include flags for the descending orbital observations.

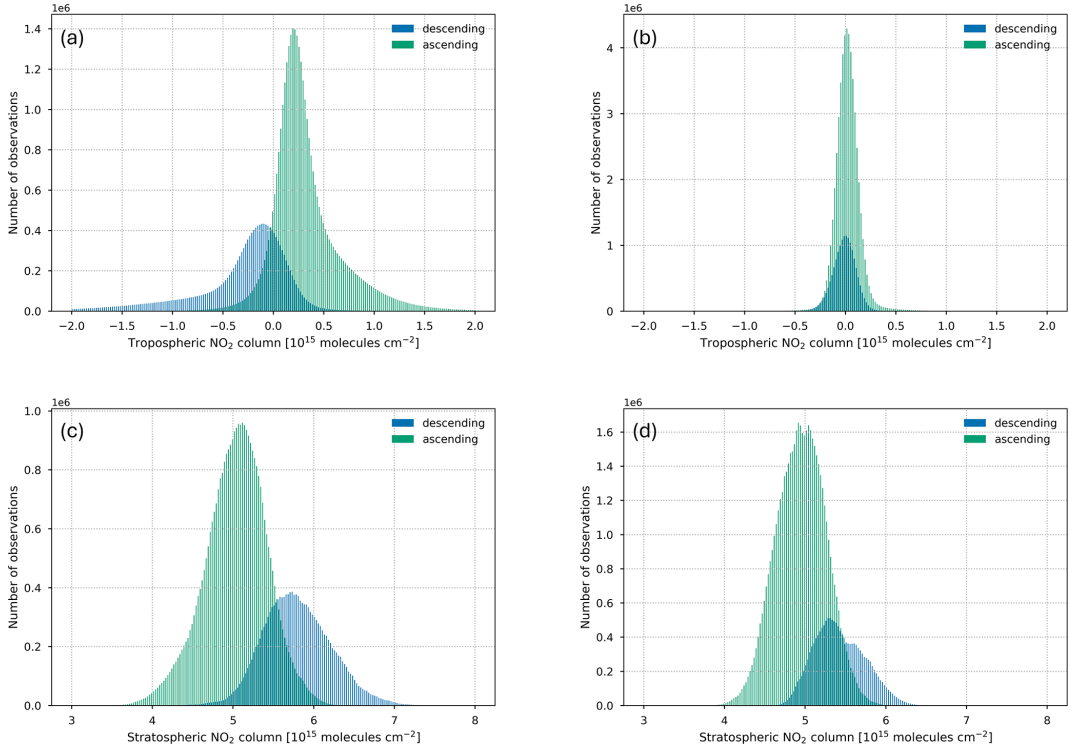

Figure A1. NO_2 vertical column observations (qa_value > 0.75) in ascending and descending mode for the troposphere **(a)** and stratosphere **(c)** north of 60° latitude from orbits in June 2019 and for the troposphere **(b)** and stratosphere **(d)** south of −60° latitude from orbits in December 2019.

Appendix B: The number of effective observations

The number of effective observations N_{eff} characterises how sensitive a region is to systematic sampling, which can occur due to for example a cloud field covering part of a grid cell. The lower the value for N_{eff}, the more sensitive a region is to systematic sampling. In a region where the uncertainty in NO_2 columns is dominated by noise, usually a relatively unpolluted region, the N_{eff} is high. In such a region, the draw of one retrieval can be enough to characterise the spatial ensemble on a whole. In an area with strong pollution gradients, usually a polluted area, multiple draws are needed for proper characterisation. For example, a grid cell covering a city effectively consists of two types of NO_2 observations: those over the polluted city and those over the unpolluted nearby rural area, and such areas are thus sensitive to systematic sampling.

The value for N_{eff} was estimated by Rijsdijk et al. (2025) 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 representativeness (the absolute difference between the mean of observed pixels and the true mean of the fully covered grid cell when including all pixels). The representativeness function f in Eq. (6) was then fitted to the mean of these experiments, providing the number of effective observations N_{eff} for this superobservation grid cell (Fig. B1a). For this polluted grid cell, with resolution $0.2° \times 0.2°$ and TROPOMI pixels of $5.5 \times 3.5\,\text{km}^2$, the number of effective observations N_{eff} is somewhat smaller than 13.

This method was repeated for multiple superobservation grid cells. The results showed that polluted regions ($x_{\text{o,t}} \geq 1.8\,\text{molecules cm}^{-2}$) have a lower value for N_{eff} than unpolluted regions ($x_{\text{o,t}} < 1.8\,\text{molecules cm}^{-2}$), making it more sensitive to systematic sampling (Rijsdijk et al., 2025). The ratio N/N_{eff}, where N is the number of pixels (valid or not) in a superobservation, was determined for NO_2 retrievals from TROPOMI in Rijsdijk et al. (2025) to be linearly dependent on the grid resolution. This linear relationship was determined for both polluted and unpolluted superobservations by Rijsdijk et al. (2025). For a superobservation resolution of $0.2° \times 0.2°$, this results in $N/N_{\text{eff}} = 1.376$ for unpolluted areas which are not sensitive to systematic sampling and $N/N_{\text{eff}} = 3.933$ for polluted areas which are more sensitive to systematic sampling (Table B1). These ratios are used to determine the number of effective observations N_{eff} for each grid cell in the superobservation. When the N_{eff} is equal to the number of pixels N, and thus the ratio N/N_{eff} equals 1, systematic sampling has no effect. The ratios N/N_{eff} would need to be redetermined for another substance or sensor (with a different pixel resolution) following the methods in Rijsdijk et al. (2025).

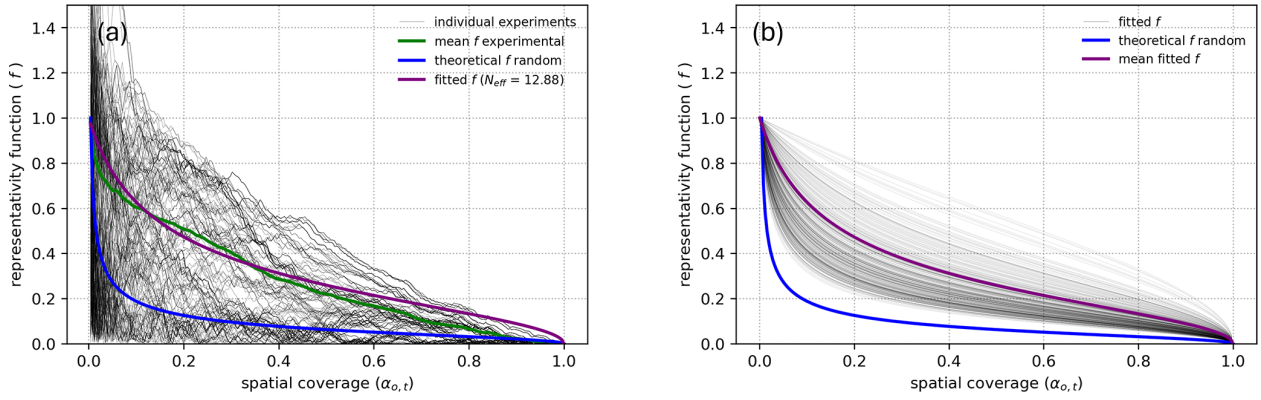

Figure B1. Results of repeatedly sampling grid cells in a systematic way to calculate the representativeness function f. **(a)** One single grid cell (polluted, $x_{\text{o,t}} \geq 1.8 \times 10^{15}$ molecules cm^{-2}) sampled 200 times. The grey lines represent 200 individual experiments sampling the same superobservation grid cell. The green line is the mean of the individual experiments, and the purple curve is the representativeness function in Eq. (6) fitted to the green line. The blue curve shows the theoretical representativeness function 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 are 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.

Table B1. The ratio N/N_{eff} for TROPOMI NO_2 in unpolluted and polluted ($x_{\mathrm{o,t}} > 1.8 \times 10^{15}$ molecules cm^{-2}) regions for different spatial resolutions of the superobservation grid. The ratio increases with superobservation area ($^{\circ 2}$).

	N/N_{eff} unpolluted areas	N/N_{eff} polluted areas
$0.2^\circ \times 0.2^\circ$	1.376	3.933
$0.5^\circ \times 0.5^\circ$	1.890	7.392
$1.0^\circ \times 1.0^\circ$	3.724	19.746
$2.0^\circ \times 2.5^\circ$	13.508	85.634

Appendix C: Contents of L3 dataset

Table C1. Overview of the variables, units, and types in the main L3 data output product file.

name/data	symbol	unit	dataset name
air mass factor	M^{trop}	1	`tropospheric_NO2_column_number_density_amf`
	M	1	`total_NO2_column_number_density_amf`
averaging kernel	$\mathbf{A}$	1	`NO2_averaging_kernel`
cloud radiance fraction	$w_{\mathrm{NO_2}}$	1	`cloud_fraction`
cloud pressure	p_{c}	hPa	`cloud_pressure`
grid cell coordinates	δ_{geo}	$^\circ$	`latitude`
	ϑ_{geo}	$^\circ$	`longitude`
grid cell corners	δ_{geo}	$^\circ$	`latitude_bounds`
	ϑ_{geo}	$^\circ$	`longitude_bounds`
land/water classification	–	1	`land_water_mask`
number of pixels used	–	1	`no_observations`
in averaging	–	1	`tropospheric_NO2_column_number_density_count`
profile layers	N_{l}	1	`layer`
quality assurance value	–	1	`qa_L3`
slant column density	$N_{\mathrm{s,NO_2}}$	molecules cm^{-2}	`NO2_slant_column_number_density`
	$N^{\mathrm{trop}}_{\mathrm{s,NO_2}}$	molecules cm^{-2}	`NO2_slant_column_number_density_troposphere`
	$\sigma_{N_{\mathrm{s,NO_2}}}$	molecules cm^{-2}	`NO2_slant_column_density_uncertainty`
surface albedo	A_{s}	1	`surface_albedo`
surface pressure	p_{s}	hPa	`surface_pressure`
time	–	days	`eff_date`
	–	1	`eff_frac_day`
	t	date	`time`
TM5 pressure level	$A^{\mathrm{TM5}}_{\mathrm{l}}$	hPa	`tm5_sigma_a`
coefficients	$B^{\mathrm{TM5}}_{\mathrm{l}}$	1	`tm5_sigma_b`
vertical column density	$N^{\mathrm{trop}}_{\mathrm{v,NO_2}}$	molecules cm^{-2}	`tropospheric_NO2_column_number_density`
	$N^{\mathrm{strat}}_{\mathrm{v,NO_2}}$	molecules cm^{-2}	`stratospheric_NO2_column_number_density`
vertical column uncertainty	$\bar{\sigma}$	molecules cm^{-2}	`tropospheric_NO2_column_number_density_temporal_std`
	$\bar{\sigma}_{\mathrm{total}}$	molecules cm^{-2}	`tropospheric_NO2_column_number_density_total_uncertainty`
	$\bar{\sigma}_{\mathrm{total,kernel}}$	molecules cm^{-2}	`tropospheric_NO2_column_number_density_total_uncertainty_kernel`

Author contributions. IG developed the methodology, created software, and carried out the analysis and writing of the initial draft. KFB co-developed the methodology and contributed to the conceptualisation and writing of the manuscript. IA and HE contributed to the conceptualisation and analysis of results. PR developed the methodology and software for the spatial averaging procedure. TV, SC, GP, JCL, and MVR were involved in the validation of the dataset. All authors contributed to reviewing and editing of the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This work was carried out as part of the ESA CCI+ Precursors for Aerosols and Ozone project, funding the contributions from Isolde Glissenaar, Isidora Anglou, Klaas Folkert Boersma, Henk Eskes, Tijl Verhoelst, Steven Compernolle, Gaia Pinardi, Jean-Christopher Lambert, and Michel Van Roozendael.

Sentinel-5 Precursor is a European Space Agency (ESA) mission on behalf of the European Commission (EC). The TROPOMI payload is a joint development by ESA and the Netherlands Space Office (NSO). The Sentinel-5 Precursor ground segment development has been funded by the ESA and with national contributions from the Netherlands, Germany, and Belgium. This work contains Sentinel-5P TROPOMI data, processed in the operational framework or locally at KNMI.

Financial support. This research has been supported by the European Space Agency (contract no. 4000138243).

Review statement. This paper was edited by Iolanda Ialongo and reviewed by two anonymous referees.

References

Bauwens, M., Compernolle, S., Stavrakou, T., Müller, J.-F., Van Gent, J., Eskes, H., Levelt, P. F., Van Der A, R., Veefkind, J., Vlietinck, J., Yu, H., and Zehner, C.: Impact of coronavirus outbreak on NO_2 pollution assessed using TROPOMI and OMI observations, Geophys. Res. Lett., 47, e2020GL087978, https://doi.org/10.1029/2020GL087978, 2020.

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., Eskes, H., and Brinksma, E.: Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res., 109, D04311, https://doi.org/10.1029/2003JD003962, 2004.

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, https://doi.org/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. C.: Improving algorithms and uncertainty estimates for satellite NO_2 retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project, Atmos. Meas. Tech., 11, 6651–6678, https://doi.org/10.5194/amt-11-6651-2018, 2018.

Bognar, K., Zhao, X., Strong, K., Boone, C., Bourassa, A., Degenstein, D., Drummond, J., Duff, A., Goutail, F., Griffin, D., Jeffery, P., Lutsch, E., Manney, G., McElroy, C., McLinden, C., Millán, L., Pazmino, A., Sioris, C., Walker, K., and Zou, J.: Updated validation of ACE and OSIRIS ozone and NO_2 measurements in the Arctic using ground-based instruments at Eureka, Canada, J. Quant. Spectrosc. Ra., 238, 106571, https://doi.org/10.1016/j.jqsrt.2019.07.014, 2019.

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.

Chan, K. L., Valks, P., Heue, K.-P., Lutz, R., Hedelt, P., Loyola, D., Pinardi, G., Van Roozendael, M., Hendrick, F., Wagner, T., Kumar, V., Bais, A., Piters, A., Irie, H., Takashima, H., Kanaya, Y., Choi, Y., Park, K., Chong, J., Cede, A., Frieß, U., Richter, A., Ma, J., Benavent, N., Holla, R., Postylyakov, O., Rivera Cárdenas, C., and Wenig, M.: Global Ozone Monitoring Experiment-2 (GOME-2) daily and monthly level-3 products of atmospheric trace gas columns, Earth Syst. Sci. Data, 15, 1831–1870, https://doi.org/10.5194/essd-15-1831-2023, 2023.

Cifuentes, F., Eskes, H., Piters, A., Gomez, J., Douros, J., Pinardi, G., Dammers, E., Cede, A., and Boersma, K. F.: Uncertainty assessment of TROPOMI NO_2 over Europe using ground-based remote sensing observations, in preparation, 2025.

Compernolle, S., Verhoelst, T., Pinardi, G., Granville, J., Hubert, D., Keppens, A., Niemeijer, S., Rino, B., Bais, A., Beirle, S., Boersma, F., Burrows, J. P., De Smedt, I., Eskes, H., Goutail, F., Hendrick, F., Lorente, A., Pazmino, A., Piters, A., Peters, E., Pommereau, J.-P., Remmers, J., Richter, A., van Geffen, J., Van Roozendael, M., Wagner, T., and Lambert, J.-C.: Validation of Aura-OMI QA4ECV NO_2 climate data records with ground-based DOAS networks: the role of measurement and comparison uncertainties, Atmos. Chem. Phys., 20, 8017–8045, https://doi.org/10.5194/acp-20-8017-2020, 2020.

Copernicus Sentinel-5P: TROPOMI Level 2 Nitrogen Dioxide total column products, version 02, European Space Agency [data], https://doi.org/10.5270/S5P-9bnp8q8, 2021a.

Copernicus Sentinel-5P: TROPOMI Level 2 Nitrogen Dioxide total column products, Version 02, European Space Agency [data], https://doi.org/10.5270/S5P-9bnp8q8, 2021b.

Dirksen, R., Boersma, K., Eskes, H., Ionov, D., Bucsela, E., Levelt, P., and Kelder, H.: Evaluation of stratospheric NO_2 retrieved from the Ozone Monitoring Instrument: Intercomparison, diurnal cycle, and trending, J. Geophys. Res.-Atmos., 116, D08305, https://doi.org/10.1029/2010JD014943, 2011.

ESA Validation Data Center (EVDC): https://evdc.esa.int/, last access: 30 August 2025.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, 2003.

Eskes, H., Velthoven, P., Valks, P., and Kelder, H.: Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model, Q. J. Roy. Meteorol. Soc., 129, 1663–1681, https://doi.org/10.1256/qj.02.14, 2003.

Eskes, H., Tsikerdekis, A., Ades, M., Alexe, M., Benedictow, A. C., Bennouna, Y., Blake, L., Bouarar, I., Chabrillat, S., Engelen, R., Errera, Q., Flemming, J., Garrigues, S., Griesfeller, J., Huijnen, V., Ilić, L., Inness, A., Kapsomenakis, J., Kipling, Z., Langerock, B., Mortier, A., Parrington, M., Pison, I., Pitkänen, M., Remy, S., Richter, A., Schoenhardt, A., Schulz, M., Thouret, V., Warneke, T., Zerefos, C., and Peuch, V.-H.: Technical note: Evaluation of the Copernicus Atmosphere Monitoring Service Cy48R1 upgrade of June 2023, Atmos. Chem. Phys., 24, 9475–9514, https://doi.org/10.5194/acp-24-9475-2024, 2024.

Faustini, A., Rapp, R., and Forastiere, F.: Nitrogen dioxide and mortality: review and meta-analysis of long-term studies, European Respirator Journal, 44, 744–753, https://doi.org/10.1183/09031936.00114713, 2014.

Fischer, P. H., Marra, M., Ameling, C. B., Hoek, G., Beelen, R., de Hoogh, K., Breugelmans, O., Kruize, H., Janssen, N. A., and Houthuijs, D.: Air pollution and mortality in seven million adults: the Dutch Environmental Longitudinal Study (DUELS), Environ. Health Perspect., 123, 697–704, https://doi.org/10.1289/ehp.1408254, 2015.

Geddes, J. A., Murphy, J. G., O'Brien, J. M., and Celarier, E. A.: Biases in long-term NO_2 averages inferred from satellite observations due to cloud selection criteria, Remote Sen. Environ., 124, 210–216, https://doi.org/10.1016/j.rse.2012.05.008, 2012.

Glissenaar, I.: L3_NO2: NO_2 temporal mean L3 (v1.2), Zenodo [code], https://doi.org/10.5281/zenodo.14505525, 2024.

Hendrick, F., Barret, B., Van Roozendael, M., Boesch, H., Butz, A., De Mazière, M., Goutail, F., Hermans, C., Lambert, J.-C., Pfeilsticker, K., and Pommereau, J.-P.: Retrieval of nitrogen dioxide stratospheric profiles from ground-based zenith-sky UV-visible observations: validation of the technique through correlative comparisons, Atmos. Chem. Phys., 4, 2091–2106, https://doi.org/10.5194/acp-4-2091-2004, 2004.

Hendrick, F., Müller, J.-F., Clémer, K., Wang, P., De Mazière, M., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G., Stavrakou, T., Vlemmix, T., and Van Roozendael, M.: Four years of ground-based MAX-DOAS observations of HONO and NO_2 in the Beijing area, Atmos. Chem. Phys., 14, 765–781, https://doi.org/10.5194/acp-14-765-2014, 2014.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., and Abuhassan, N.: NO_2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, J. Geophys. Res.-Atmos., 114, D13307, https://doi.org/10.1029/2009JD011848, 2009.

Hönninger, G. and Platt, U.: Observations of BrO and its vertical distribution during surface ozone depletion at Alert, Atmos. Environ., 36, 2481–2489, https://doi.org/10.1016/S1352-2310(02)00104-8, 2002.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Isserlis, L.: On the Value of a Mean as Calculated from a Sample, J. Roy. Stat. Soc., 81, 75–81, http://www.jstor.org/stable/2340569 (last access: 30 August 2025), 1918.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res.-Atmos., 113, D18308, https://doi.org/10.1029/2008JD010290, 2008.

KNMI: ESA CCI+ TROPOMI L3 monthly mean NO_2 columns, Royal Netherlands Meteorological Institute (KNMI) [data set], https://doi.org/10.21944/CCI-NO2-TROPOMI-L3, 2025.

Labzovskii, L. D., van Geffen, J., Liu, M., van der A, R., de Laat, J., Leune, B., Eskes, H., Lin, X., Ding, J., and Richter, A.: NO_2 satellite retrievals biased by absorption in water, Nat. Geosci., 17, 972–975, https://doi.org/10.1038/s41561-024-01545-8, 2024.

Lambert, J.-C., Van Roozendael, M., Granville, J., Gerard, P., Peeters, P., Simon, P., Claude, H., and Stahelin, J.: Comparison of the GOME ozone and NO_2 total amounts at mid-latitude with ground-based zenith-sky measurements, edited by: Bojkov, R. and Visconti, G., Atmospheric Ozone – 18th Quad. Ozone Symp., L'Aquila, Italy, 1996, 1, 301–304, 1997.

Lambert, J.-C., Keppens, A., Compernolle, S., Eichmann, K.-U., de Graaf, M., Hubert, D., Langerock, B., Ludewig, A., Sha, M., Verhoelst, T., Wagner, T., Ahn, C., Argyrouli, A., Balis, D., Chan, K., Coldewey-Egbers, M., Smedt, I. D., Eskes, H., Fjæraa, A., Garane, K., Gleason, J., Goutail, F., Granville, J., Hedelt, P., Ahn, C., Heue, K.-P., Jaross, G., Kleipool, Q., Koukouli, M., Lutz, R., Velarte, M. M., Michailidis, K., Nanda, S., Niemeijer, S., Pazmiño, A., Pinardi, G., Richter, A., Rozemeijer, N., Sneep, M., Zweers, D. S., Theys, N., Tilstra, G., Torres, O., Valks, P., van Geffen, J., Vigouroux, C., Wang, P., and Weber, M.: Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #21: April 2018 – November 2023, Tech. rep., S5P MPC Routine Operations Consolidated Validation Report series, 2023.

Liu, M., Shang, F., Lu, X., Huang, X., Song, Y., Liu, B., Zhang, Q., Liu, X., Cao, J., Xu, T., Wang, T., Xu, Z., Xu, W., Liao, W., Kang, L., Cai, X., Zhang, H., Dai, Y., and Zhu, T.: Unexpected response of nitrogen deposition to nitrogen oxide controls and implications for land carbon sink, Nat. Commun., 13, 3126, https://doi.org/10.1038/s41467-022-30854-y, 2022.

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., Sofieva, 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.

Miller, A., Brown, C., and Franson, K. D.: Photochemical reactions of nitrogen dioxide and hydrocarbons in air, Science, 123, 934, https://doi.org/10.1126/science.123.3204.934.a, 1956.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO_2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012.

NDACC: Measurement Stations, https://ndacc.larc.nasa.gov/, last access: 30 August 2025.

Nightingale, J., Boersma, K. F., Muller, J.-P., Compernolle, S., Lambert, J.-C., Blessing, S., Giering, R., Gobron, N., De Smedt, I., Coheur, P., George, M., Schulz, J., and Wood, A.: Quality Assurance framework development based on six new ECV data products to enhance user confidence for climate applications, Remote Sens., 10, 1254, https://doi.org/10.3390/rs10081254, 2018.

Pandonia Global Network (PGN): https://www.pandonia-global-network.org/, last access: 30 August 2025.

Pinardi, G., Van Roozendael, M., Hendrick, F., Theys, N., Abuhassan, N., Bais, A., Boersma, F., Cede, A., Chong, J., Donner, S., Drosoglou, T., Dzhola, A., Eskes, H., Frieß, U., Granville, J., Herman, J. R., Holla, R., Hovila, J., Irie, H., Kanaya, Y., Karagkiozidis, D., Kouremeti, N., Lambert, J.-C., Ma, J., Peters, E., Piters, A., Postylyakov, O., Richter, A., Remmers, J., Takashima, H., Tiefengraber, M., Valks, P., Vlemmix, T., Wagner, T., and Wittrock, F.: Validation of tropospheric NO_2 column measurements of GOME-2A and OMI using MAX-DOAS and direct sun network observations, Atmos. Meas. Tech., 13, 6141–6174, https://doi.org/10.5194/amt-13-6141-2020, 2020.

Pommereau, J. and Goutail, F.: O_3 and NO_2 ground-based measurements by visible spectrometry during Arctic winter and spring 1988, Geophys. Res. Lett., 15, 891–894, https://doi.org/10.1029/GL015i008p00891, 1988.

Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P.: An improved NO_2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech., 4, 1147–1159, https://doi.org/10.5194/amt-4-1147-2011, 2011.

Richter, A., Boersma, F., De Smedt, I., Theys, N., Pinardi, G., Heue, K.-P., Seo, S., Argyrouli, A., Verhoelst, T., and Compernolle, S.: Product Validation and Algorithm Selection Report (PVASR), Tech. rep., University of Bremen, https://climate.esa.int/en/projects/precursors-for-aerosols-and-ozone/ (last access: 30 August 2025), 2024.

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. and Eskes, H.: Superobservation software (1.0), Zenodo [code], https://doi.org/10.5281/zenodo.10726645, 2024.

Rijsdijk, P., Eskes, H., Dingemans, A., Boersma, K. F., Sekiya, T., Miyazaki, K., and Houweling, S.: Quantifying uncertainties in satellite NO_2 superobservations for data assimilation and model evaluation, Geosci. Model Dev., 18, 483–509, https://doi.org/10.5194/gmd-18-483-2025, 2025.

Sekiya, T., Miyazaki, K., Eskes, H., Sudo, K., Takigawa, M., and Kanaya, Y.: A comparison of the impact of TROPOMI and OMI tropospheric NO_2 on global chemical data assimilation, Atmos. Meas. Tech., 15, 1703–1728, https://doi.org/10.5194/amt-15-1703-2022, 2022.

Shah, V., Jacob, D. J., Li, K., Silvern, R. F., Zhai, S., Liu, M., Lin, J., and Zhang, Q.: Effect of changing NO_x lifetime on the seasonality and long-term trends of satellite-observed tropospheric NO_2 columns over China, Atmos. Chem. Phys., 20, 1483–1495, https://doi.org/10.5194/acp-20-1483-2020, 2020.

Solomon, S., Schmeltekopf, A. L., and Sanders, R. W.: On the interpretation of zenith sky absorption measurements, J. Geophys. Res., 92, 8311–8319, https://doi.org/10.1029/JD092iD07p08311, 1987.

Tan, J., Fu, J. S., and Seinfeld, J. H.: Ammonia emission abatement does not fully control reduced forms of nitrogen deposition, P. Natl. Acad. Sci., 117, 9771–9775, https://doi.org/10.1073/pnas.1920068117, 2020.

Tilstra, L. G., de Graaf, M., Trees, V. J. H., Litvinov, P., Dubovik, O., and Stammes, P.: A directional surface reflectance climatology determined from TROPOMI observations, Atmos. Meas. Tech., 17, 2235–2256, https://doi.org/10.5194/amt-17-2235-2024, 2024.

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, 2022a.

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO_2 data products, Tech. Rep. S5P-KNMI-L2-0005-RP, version 2.4.0, Royal Netherlands Meteorological Institute, https://sentiwiki.copernicus.eu/web/s5p-products (last access: 30 August 2025), 2022b.

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO_2 data products, Tech. Rep. S5P-KNMI-L2-0005-RP, version 2.8.0, Royal Netherlands Meteorological Institute, https://sentiwiki.copernicus.eu/_attachments/1673595/S5P-KNMI-L2-0005-RP - Sentinel-5P TROPOMI ATBD NO2 data products 2024 - 2.8.0.pdf?inst-v=aae16534-0abf-4c5e-b66d-6774e8ba38d4 (last access: 30 August 2025), 2024.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H., De Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J.-P., Redondas, A., Van Roozendael, M., and Lambert, J.-C.: Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons, At-

mos. Meas. Tech., 8, 5039–5062, https://doi.org/10.5194/amt-8-5039-2015, 2015.

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J.-C., Eskes, H. J., Eichmann, K.-U., Fjæraa, A. M., Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A., Boersma, K. F., Bognar, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter de la Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagkiozidis, D., Kivi, R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., Piters, A. J. M., Pommereau, J.-P., Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Rivera Cárdenas, C., Saavedra de Miguel, L., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendael, M., Veefkind, J. P., Wagner, T., Wittrock, F., Yela González, M., and Zehner, C.: Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks, Atmos. Meas. Tech., 14, 481–510, https://doi.org/10.5194/amt-14-481-2021, 2021.

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.

Wei, J., Liu, S., Li, Z., Liu, C., Qin, K., Liu, X., Pinker, R. T., Dickerson, R. R., Lin, J., Boersma, K., Sun, L., Li, R., Xue, W., Cui, Y., Zhang, C., and Wang, J.: Ground-level NO_2 surveillance from space across China for high resolution using interpretable spatiotemporally weighted artificial intelligence, Environ. Sci. Technol., 56, 9988–9998, 2022.

Wenig, M. O., Cede, A., Bucsela, E., Celarier, E., Boersma, K., Veefkind, J., Brinksma, E., Gleason, J., and Herman, J.: Validation of OMI tropospheric NO_2 column densities using direct-Sun mode Brewer measurements at NASA Goddard Space Flight Center, J. Geophys. Res.-Atmos., 113, D16S45, https://doi.org/10.1029/2007JD008988, 2008.

Williams, J. E., Boersma, K. F., Le Sager, P., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation, Geosci. Model Dev., 10, 721–750, https://doi.org/10.5194/gmd-10-721-2017, 2017.

World Meteorological Organization (WMO), United Nations Environment Programme (UNEP), International Science Council (ISC), and Intergovernmental Oceanographic Commision of the United Nations Educational, Scientific and Cultural Organization (IOC-UNESCO): The 2022 GCOS Implementation Plan, Tech. Rep. GCOS 244, World Meteorological Organization, https://library.wmo.int/idurl/4/58104 (last access: 30 August 2025), 2022a.

World Meteorological Organization (WMO), United Nations Environment Programme (UNEP), International Science Council (ISC), and Intergovernmental Oceanographic Commision of the United Nations Educational, Scientific and Cultural Organization (IOC-UNESCO): The 2022 GCOS ECVs Requirements, Tech. Rep. GCOS 245, World Meteorological Organization, https://library.wmo.int/idurl/4/58111 (last access: 30 August 2025), 2022b.

Yela, M., Gil-Ojeda, M., Navarro-Comas, M., Gonzalez-Bartolomé, D., Puentedura, O., Funke, B., Iglesias, J., Rodríguez, S., García, O., Ochoa, H., and Deferrari, G.: Hemispheric asymmetry in stratospheric NO_2 trends, Atmos. Chem. Phys., 17, 13373–13389, https://doi.org/10.5194/acp-17-13373-2017, 2017.

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 characterisation, and trends, Atmos. Meas. Tech., 11, 4033–4058, https://doi.org/10.5194/amt-11-4033-2018, 2018.

Zara, M., Boersma, K. F., Eskes, H., van der Gon, H. D., de Arellano, J. V.-G., Krol, M., van der Swaluw, E., Schuch, W., and Velders, G. J.: Reductions in nitrogen oxides over the Netherlands between 2005 and 2018 observed from space and on the ground: Decreasing emissions and increasing O_3 indicate changing NO_x chemistry, Atmos. Environ., 9, 100104, https://doi.org/10.1016/j.aeaoa.2021.100104, 2021.

Zemp, M., Chao, Q., Han Dolman, A. J., Herold, M., Krug, T., Speich, S., Suda, K., Thorne, P., and Yu, W.: GCOS 2022 implementation plan, Tech. rep., World Meteorological Organization, https://doi.org/10.5167/uzh-224271, 2022.

Zhang, Q., Pan, Y., He, Y., Walters, W. W., Ni, Q., Liu, X., Xu, G., Shao, J., and Jiang, C.: Substantial nitrogen oxides emission reduction from China due to COVID-19 and its impact on surface ozone and aerosol pollution, Sci. Total Environ., 753, 142238, https://doi.org/10.1016/j.scitotenv.2020.142238, 2021.

Zhang, Q., Boersma, K. F., Zhao, B., Eskes, H., Chen, C., Zheng, H., and Zhang, X.: Quantifying daily NO_x and CO_2 emissions from Wuhan using satellite observations from TROPOMI and OCO-2, Atmos. Chem. Phys., 23, 551–563, https://doi.org/10.5194/acp-23-551-2023, 2023.

Zhang, R., Wang, G., Guo, S., Zamora, M. L., Ying, Q., Lin, Y., Wang, W., Hu, M., and Wang, Y.: Formation of urban fine particulate matter, Chem. Rev., 115, 3803–3855, https://doi.org/10.1021/acs.chemrev.5b00067, 2015.