Precursors_cci+

ESA Climate Change Initiative (CCI)

D1.1 User Requirement Document

Title	D1.1 User Requirement Document
Reference	Precursors_cci+_D1.1_URD_01_01
Issue	01
Revision	01
Status	Final
Date of issue	25/01/2023
Document type	Deliverable

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

Generated by KNMI

	FUNCTION	NAME	DATE
LEAD AUTHOR(S)	Co-Science Leader Contributing authors	Folkert Boersma I. De Smedt, N. Theys. M. George, C. Clerbaux, L. Clarisse, C. Lerot, S. Compernolle, J.-C. Lambert	25/01/2023
REVIEWED BY	Science Leader	M. Van Roozendael	25/01/2023
	ESA Project Officer	Simon Pinnock	21/12/2022
DISTRIBUTED TO	ESA Precursors_cci+ Consortium		

Table of Contents

1. Purpose and scope

1.1. Purpose

The Global Climate Observing System (GCOS) Implementation Plan provides high level user requirements specification for 4 of the ECV precursor gases (NO_2, HCHO, SO_2, CO) in this CCI project. The most recent and relevant GCOS requirements have been published in documents GCOS-244 and GCOS-245 and were presented at the 27th session of the Conference of Parties of the UNFCCC (COP-27) on 11 November 2022. Quantitative requirements for the other 2 ECV precursors in this CCI project (ammonia (NH_3) and glyoxal (CHOCHO)) included in this project might appear later in 2023 for application areas - other than climate - currently discussed in WMO's Rolling Review of Requirements. This User Requirements Document critically summarizes the most recent requirements for long-term ECV precursor gas data records identified for the GCOS, and also expressed by individual users that were interviewed, and from a preliminary assessment of peer-reviewed literature on climate modelers using satellite ECV precursor data to evaluate and improve their models. This URD contains a detailed and thorough specification of the requirements, including also specific requirements which may vary according to different applications of the ECV for climate.

1.2. Scope

The URD is meant to inform algorithm developers on what the users need from the level-2 (L2) and level-3 (L3) data products to be developed in this CCI project. Ongoing feedback from various users and communities will be considered as this project evolves. The URD is also an essential input to the product validation plan (D1.3): in addition to the classical quantitative requirements on data uncertainty and stability, the users' needs set the stage for which particular aspects of the data products need to be validated.

The URD document:

- Elaborates on the GCOS IP 2022 specification [RD-6] to provide a fully detailed specification of the user requirements for the ECV.
- Provides details of all sources of user requirements considered, e.g., surveys, workshops, URDs from other projects, and scientific papers.

1.3. Reference documents

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

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

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

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

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

[RD-6] The 2022 CGOS Implementation Plan, GCOS-244, October 2022. Available online at: https://library.wmo.int/index.php?lvl=notice_display&id=22134#.Y3JSyOy0u1R

[RD-7] The 2022 CGOS ECV Requirements, GCOS-245, October 2022. Available online at: https://library.wmo.int/index.php?lvl=notice_display&id=22135#.Y3JS9Oy0u1Q

1.4. List of acronyms

AC-SAF	Satellite Application Facility on Atmospheric Composition Monitoring
ADP	Algorithm Development Plan
AK	Averaging Kernel
AMF	Air-mass factor
ATBD	Algorithm Theoretical Basis Document
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
BIRA-IR	BIRA-IASB Infrared Observations Team
BIRA-SYN	BIRA-IASB Atmospheric Data Synergies Team
BIRA-UVVIS	BIRA-IASB UV-Visible Observations Team
BIRA-MOD	BIRA-IASB Tropospheric Chemistry Modelling Team
CAMS	Copernicus Atmospheric Monitoring Service
C3S	Copernicus Climate Change Monitoring Service
CCI	ESA Climate Change Initiative
CCI+	Climate Change Initiative Extension (CCI+), is an extension of the CCI over the period 2017-2024.
CDR	Climate Data Record
CEOS	Committee on Earth Observation Satellites
CMUG	Climate Modelling User Group
CO	Carbon monoxide
COBRA	COvariance-Based Retrieval Algorithm
CRDP	Climate Research Data Package
CRG	Climate Research Group
DLR	German Aerospace Centre
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-range Weather Forecast
ECV	Essential Climate Variable
ENVISAT	Environmental Satellite (ESA)
EO	Earth Observation
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
FRM	Fiducial Reference Measurement
GCOS	Global Climate Observation System
GOME	Global Ozone Monitoring Instrument (aboard ERS-2)
GOME-2	Global Ozone Monitoring Instrument – 2 (aboard MetOp-A, -B and -C)
HRI	Hyperspectral Range Index
IASI	Infrared Atmospheric Sounding Interferometer
IGAC	International Global Atmospheric Chemistry project
IP	Implementation Plan
KNMI	Royal Netherlands Meteorological Institute
LEO	Low Earth Orbit
LUT	Look-up table
MetOp	Meteorological Operational Platform (EUMETSAT)
MOPITT	Measurement of Pollution in the Troposphere
NASA	National Aeronautics and Space Administration

NDACC	Network for the Detection of Atmospheric Composition Change
NH_3	Ammonia
NN	Neural Network
NO_2	Nitrogen dioxide
NRT	Near-Real Time
OCRA	Optical Cloud Recognition Algorithm)
OMI	Ozone Monitoring Instrument (aboard EOS-Aura)
PCA	Principal Component Analysis
QA4ECV	Quality Assurance for Essential Climate Variables
QA4EO	Quality Assurance framework for Earth Observation
R&D	Research and Development
ROCINN	Retrieval of Cloud Information using Neural Networks
SAF	Satellite Application Facility
SCIAMACHY	Scanning Imaging Absorption spectroMeter for Atmospheric CHartography
S5P	Sentinel-5 Precursor
SoW	Statement of Work
STREAM	STRatospheric Estimation Algorithm from Mainz
SZA	Solar Zenith Angle
TEMIS	Tropospheric Emission Monitoring Internet Service
TIR	Thermal Infrared spectral range
TROPOMI	Tropospheric Monitoring Instrument (aboard Sentinel-5 Precursor)
TOA	Top-of-atmosphere
TOAR-II	Tropospheric Ozone Assessment Report Phase-II
ULB	Université Libre de Bruxelles
IUP-UB	Institute of Environmental Physics, University of Bremen
UPAS	Universal Processor for UV/Vis Atmospheric Sensors
URD	User Requirement Document
UV-Vis	Ultraviolet and visible spectral range
VZA	Viewing Zenith Angle
WP	Work Package

2. Results from user requirements analysis

2.1 Analysis of QA4ECV user requirements

Documents analysed:
QA4ECV Deliverable 1.1, Results from the QA4ECV User Requirements Survey on quality assurance in satellite data, September 2015.
QA4ECV Deliverable 1.3, Update of the User Requirements Report, February 2018.

In the EU FP7 QA4ECV project, a wide user survey was done to investigate what satellite data users need to put available atmospheric data products to good use. *By nature of the QA4ECV-project, the survey and analysis of user requirements focused on satellite data quality assurance for users, in particular the availability and usefulness of quality flags, traceability information, uncertainties, and evidence of validation.* These aspects are all relevant, but, in the context of the ESA CCI+ Precursors project, the question of why users require ECV precursor data, and for what purposes or applications, should come first.

Another lesson from the QA4ECV user requirement analysis is that *asking specifically what users are missing in existing data data-products helps to address omissions in the new satellite data products* generated towards the end of the project. The QA4ECV NO_2, HCHO, and CO data-products have been extended with more quality indicators relative to earlier versions of these retrievals and formatted in line with the data format selected for NO_2, HCHO, SO_2 and CO from the S5P mission. The ESA CCI+ Precursor project will take on board what has been achieved in QA4ECV and for S5P in terms of data content, data format, and accessibility.

The third clear-cut lesson from QA4ECV is that while the very broad user survey sent out via email reached a lot of data users of diverse backgrounds, the answers were often not specific enough. A follow-up up on the written survey with *targeted interviews by phone or in person gave better insight in the needs of users*.

The above considerations shaped the form and emphasis of the interview questions to better understand users' requirements to use long-term satellite data of ECV precursors.

2.2 Users interviewed

Table 1 summarizes all users who have shared their opinions on what they need in terms of data record, data content, availability, representativeness, flagging, uncertainties, and validation. The users have been selected for their previous work with satellite ECV precursor data, and especially the long-term aspects thereof (trend analysis, atmospheric composition reanalysis). The users received specific questionnaires by email and were then asked to complete the form in writing, after which follow-up video-calls were held with them to discuss their answers in more detail.

Table 1: Overview of users interviewed on user requirements for ECV precursor in August-September 2022

User	Trace gas(es)	Affiliation	Research Focus
Antje Inness	NO_2, CO, HCHO, SO_2	ECMWF	Data assimilation for atmospheric composition reanalysis
Kazuyuki Miyazaki	NO_2, CO, HCHO, SO_2	NASA JPL	Decadal chemical data assimilation
Christoph Riess	NO_2, SO_2	Wageningen University	Interannual variability in ship emissions
Aris Georgoulias	NO_2	Aristotle University of Thessaloniki	Long-term trend analysis
Lok Lamsal	NO_2	NASA, University of Maryland	Generation of NO_2 climate data record
Vincent Huijnen	CHOCHO, HCHO, NO_2, CO, NH_3	KNMI	Global atmospheric composition model evaluation
Daniel Jacob	CHOCHO, HCHO, NO_2, NH_3	Harvard University	Emission monitoring, chemical regime
Arjo Segers	CHOCHO, NO_2, SO_2, HCHO, CO	TNO	model validation and assimilation
Xiaomeng Jin	HCHO, NO_2	University of California, Berkeley	Trend analysis, emissions of wildfires, ozone chemical regime, public outreach
Xueying Yu	HCHO, CO	Stanford University	Inverse analyses to study atmospheric chemistry
Karn Vohra	HCHO, NH_3	University College London	Trend analysis and chemical regime for urban areas

Gonzalo Abad/Caroline Nowlan	HCHO/CHOCHO	NASA Center for Astrophysics, Harvard & Smithsonian	satellite/sensor intercomparison and validation
Simon Rosanka	CHOCHO	Institute of Energy and Climate Research IEK	Model evaluation
Can Li/Nickolay Krotkov	SO_2, NO_2	NASA	Satellite product developer, long-term trend, intercomparison
Vitali Fioletov	SO_2	Environment Canada	Long-term emissions estimation
Zhen Qu	SO_2, NO_2	University of Colorado	Long-term top-down emission estimates of pollutants
Zhenqi Luo	NH_3	Cornell University	Emissions, Modeling
Jonathan Hickman	NH_3	NASA Goddard Institute for Space Studies	NH_3 emission, seasonality and trends in Africa
Audrey Fortems/Gaëlle Dufour	NH_3	Université Paris-Est Créteil	NH_3 modeling, particulate matter
Dang Ruijin	NH_3	Harvard University	Model evaluation
Enrico Dammers	NH_3	TNO	Validation, point sources, inverse modeling
Maureen Beaudor	NH_3	LSCE-IPSL	Emissions, modeling
Benjamin Gaubert	CO	NCAR	Trend analysis, comparison with MOPITT
Christoph Keller	CO	NASA	Evaluation of CO in NASA GEOS-CF system
Maarten Krol	CO	Wageningen University	Data assimilation in TM5
Antoine Ehret	CO	LMD	Fire studies

2.3 User requirements summary for NO_2

Tropospheric NO_2 columns are widely used to monitor NO_x (NO_x = NO + NO_2) emissions (e.g. Zhang et al. (2023), Lange et al. (2022)), to analyse long-term trends in NO_2 pollution (e.g. Zara et al. (2021), Jiang et al. (2022)), to constrain decadal-scale atmospheric composition in data assimilation systems, to map chemical regimes (in combination with the HCHO product. e.g. Jin et al. (2020)), to evaluate models (e.g. Shah et al. (2020)), to help constrain surface pollution levels and surface deposition processes (e.g. Wei et al. (2022)), and for public outreach purposes. Users also use NO_2 data in combination with satellite observations of CO and SO_2 to obtain a comprehensive assessment of the (multi-species) chemical nature of air pollution (e.g. Peng et al. (2018), Miyazaki et al. (2015)).

We sent out questionnaires (see Appendix 1) to various users and obtained 5 written responses from NO_2 users as listed in Table 1. Following up on these responses, we held 5 follow-up video calls for clarifications and to elaborate more on specific topics.

Users indicated that they (want to) use level-2 (L2) and level-3 (L3) data, with most users preferring L2 data, so that they can impose their own flagging and averaging/gridding procedures. They welcome the provision of a multi-sensor data product and L3 data, especially for the purpose of trend analysis and reanalysis of atmospheric composition but demanded extension of the L3 data products beyond just gridded and averaged data. Also gridded information on flagging and sampling, measurement time, uncertainties, and averaging kernels is required in L3 data. Current data products are considered generally well accessible, although the large data volume is considered troublesome in the case of S5P.

Rec-NO2-1. Improve availability and accessibility of data

Users consider current tropospheric NO_2 data products to be findable and accessible. Some users requested that L2 data can be found via high-level generic entrance points (such as GCOS), along with tools (such as *wget* scripts) to quickly download subsets of the product.

Rec-NO2-2. Make available well-documented and complete L3 data

Although (monthly mean) L3 data is already available for tropospheric NO_2, users indicated that they require better documented (flags used, cloud screening used, native resolution) L3 data, extended with information on observation times, number of observations gridded, L3 uncertainty estimates, and preferably on a daily and monthly mean basis.

Rec-NO2-3. Improve awareness and usage of uncertainty estimates and flagging

Users agreed that current NO_2 products and their documentation contain useful information on how to use uncertainty estimates and flagging. New, less experienced users recommend providing information on the satellite data in a clear and brief manner than in the ATBD or PUM. This can be achieved with a one-pager highlighting the essentials of proper data usage: representativeness and limitations of the measurement, how to reduce (systematic) uncertainties, usage of flagging.

Rec-NO2-4. Provide code snippets and examples on how to apply the averaging kernels

Along with the data, users indicated they would be helped if code (e.g. in Python) would be made available that demonstrates how to apply the averaging kernel, or how to recalculate air mass factors with different a priori profiles than those used in the retrieval product. A demo of the application of such a piece of code on an exemplary NO_2 profile would help users to avoid pitfalls with units, vertical interpolation, and interpretation.

Rec-NO2-5. Cross-validate QA4ECV and AC-SAF NO_2 algorithms
Users recommended to concentrate validation efforts on simultaneous validation of the current QA4ECV and AC-SAF NO_2 data products, in order to obtain guidance on how to improve the NO_2 retrieval algorithm within this project. L1 data quality, retrieval algorithms, and use of auxiliary information should be as consistent as possible in generating a multi-sensor record.

2.4 User requirements summary for HCHO

HCHO columns are widely used to monitor and constrain non-methane volatile organic compounds (NMVOCs) emissions using tropospheric models, to analyse long-term trends in anthropogenic and biogenic NMVOC emissions (e.g. Stavrakou et al., 2018; Sun et al., 2020; Bauwens et al., 2022; Morfopoulos et al., 2022; Wang, Y. et al., 2022), to evaluate models (Cao et al., 2018; Zhao et al., 2021; Opacka et al., 2022; Wang, P. et al., 2022) and for mapping and public outreach purposes (e.g. on Twitter, Copernicus, ESA, Terrascope). Users often use HCHO in combination with satellite observations of NO2, but also CHOCHO to map chemical regimes at urban and regional scales (Jin et al.,m2020; Li et al., 2020, Jung et al., 2022) or in wildfire studies (e.g. Stavrakou et al., 2016; Alvarado et al., 2020; Theys et al., 2020).

Users indicated that they use L2 and L3 data, with the more advanced users preferring L2 data, so that they can impose their own flagging, apply vertical averaging kernels and a priori profile substitution. They welcome the provision of a multi-sensor data product, for the purpose of trend analysis, reanalysis of atmospheric composition, and past and future sensor evaluation. Accessibility to the data was found to be very good, via the QA4ECV website, the S5P platform, or by contacting the product developer. However, the large S5P data volume and associated processing costs are problematic, while no L3 product is provided for TROPOMI. Quick visualization tools such as in the QA4ECV website are also appreciated.

Users were overall satisfied with the data, but some weaknesses were identified. Many users pointed out the difficulties and challenges associated with data filtering, uncertainties and errors, interpretation of the data, especially when averaging the observations.

Rec-HCHO-1: Improve documentation on data interpretation, use of quality flags and uncertainty estimates

Users suggested a better or more direct way to understand the data. Algorithm Theoretical Baseline Documents (ATBDs) and Product User Manuals (PUMs) are very detailed and a short and easy to digest information on the data product, the uncertainties calculation and quality flagging, is desirable.

As the noise of HCHO retrievals is known to be large, clear instructions on how to average the data and how to use negative values are needed. User guidance is also needed on how to use the different reported uncertainties. Is the uncertainty decreasing when averaging and what component of bias it contains? How to derive weighted means? Are there correlations between the uncertainties of different molecules (how to calculate ratio uncertainties)?

The single *qa_value* of TROPOMI is easy to use but not detailed enough for most of the advanced users, and not traceable. Flags should allow testing different filtering for viewing geometries, surface reflectance, cloud fraction, aerosol contamination as well as sensor issues and gap filling. Harmonization of the quality flags among the sensors could be improved.

Rec-HCHO-2: Long term stability and across sensors

Users have underlined the importance of having clean data sets with artifacts filtered out (with improved quality flags), stable in time for trend analysis and well characterized biases. The consistency between different sensors could still be improved. Spatial and temporal resolution differences should be documented. Clear information on the optimal spatial and temporal resolution for each sensor is requested.

Rec-HCHO-3: Make available well-documented and complete L3 data

Users requested better documented and more complete level-3 data products. Diagnostic variables from the L2 should be propagated in the L3 as much as possible (e.g the number of observations or the weights, propagated errors, mean AKs, cloud fraction, surface properties). Daily and monthly products are requested.

Rec-HCHO-4: Continuous validation

Several validation studies have been published (Wang et al., 2019; Zhu et al., 2020; Vigouroux et al., 2020; De Smedt et al., 2015; 2021). The conclusions on precision and bias estimates should be provided with the product. Some users are interested in guidance on how to correct the satellite data with the reported linear bias. Transferring local validation results to other regions is seen as a difficulty. Validation is seen as something important and that should be repeated and improved at regular intervals.

2.5 User requirements summary for SO_2

Satellite columns for SO_2 are widely used to estimate SO_2 emissions emissions (Beirle et al., 2014; Fioletov et al., 2023 and references therein), sometimes in conjunction with NO_2 (Qu et al., 2019), investigate long-term trends in SO_2 (Krotkov et al., 2016; van der A et al., 2017), monitor volcanic activity (Carn et al., 2017; Theys et al., 2019), follow volcanic SO_2 plume dispersion (Brenot et al., 2014), studying the impact of SO_2 on air quality (Schmidt et al., 2015), and for outreach. Most of the interviewed users indicate a preference for L2 SO_2 data, as it provides them more flexibility in the data selection/flagging and averaging. L3 data can be useful for long-term trend analysis but should include averaging kernel information. Current data products are considered generally well accessible.

Rec-SO2-1. Improve documentation and use of quality flags and uncertainty estimates

Users suggested a better or more direct way to understand the data. ATBDs and PUMs are very detailed and a short and easy to digest information on the data product, the uncertainties calculation and quality flagging, is desirable.

Rec-SO2-2. Produce traceable, stable, and consistent multi-mission L2 SO_2 product

Long-term stability and consistency are the most important aspects raised by the users. Processing of data from multiple sensors using a common algorithm to form L2 data from individual sensors is highly desirable. For historical sensors like GOME, the pixel size and noise level are a concern but data are considered potentially useful for large sources.

Rec-SO2-3. Make available additional information in L2 SO_2 data sets

Users requested to output useful information such as wind data, snow/ice information in addition to the existing variables (quality flags, cloud fractions, solar zenith angles), most useful for data screening.

Rec-SO2-4. Produce merged L3 data product including uncertainty and averaging kernel information

Although level-2 data products are considered the most useful, a multi-sensor L3 SO_2 data product could be considered e.g., for trend analysis. The L3 product should provide averaging kernels, uncertainty estimates (random and systematic) and number of observations.

2.6 User requirements summary for CO

The IASI CO product is currently used widely for a large variety of applications such as mapping distributions (Georges et al., 2015), trend analysis (Buchholz et al., 2021), assimilation in models and forecasts (Klonecki et al., 2012; Inness et al., 2013), deriving emissions (Krol et al., 2013; Kovalov et al. 2014, Nechita Banda et al, 2017), and for studies related to fires (Turquety et al., 2009), transport (Sodemann et al. 2011), and pollution. Both L2 and L3 data are used. L3 is being preferred for long-term trend analysis and by new users who want a first easy access to the data. Accessibility to the data was found to be very good, via the AERIS platform or directly from the CICLAD IPSL server for those who have access. Quick looks (daily and monthly maps) on the AERIS platform are appreciated.

Users need consistent products, with observation errors (well characterized). Observations need to come with averaging kernels and a-priori profiles. Quality flags are essential. Users are satisfied with the NCDF format, along with the provided documentation. Users trust the development team for the definition of quality flags to filter the data. MOPITT users are naturally interested in IASI data (many validation papers published).

Rec-CO-1. Users need consistent and harmonized CO dataset

The CO product available on AERIS relies on the same FORLI version but on different temperature/cloud or humidity information, provided by EUMETSAT. For long term studies, users require consistent datasets. In 2022, EUMETSAT reprocessed the full IASI period with the latest version for temperature/cloud and humidity information. A new consistent IASI CO dataset exists now but is not public yet.

Rec-CO-2. Additional L3 information would be appreciated

In the future, it would be great to get additional monthly grids containing number of points, DOFS or averaging kernels.

Rec-CO-3. Suggestion: list of users for contact

It would be great to have somewhere a list of users that have created their own L3 products from L2. For example, Antoine Ehret (LMD-IPSL) generated L3 daily CO grids from L2 data. Maybe other users would be interested and could contact Antoine Ehret to get feedback.

2.7 User requirement summary of NH_3

The IASI NH_3 product is currently used widely for a large variety of applications such as mapping of point sources (Van Damme et al., 2018) , trend analysis (Van Damme et al., 2021), air quality (Lachatre et al., 2019), wildfires (Lutsch et al., 2019), quantifying emissions (Chen et al., 2021) or study of seasonality (Wang et al., 2021). An example of where IASI NH_3 satellite is analysed together with NO_2 and HCHO data in the context of fine particulate pollution can be found in Vohra et al. (2022).

Both L2 and L3 data are used, with the more advanced users preferring the L2. Accessibility to the data was found to be very good, via the AERIS platform or directly from the developing team at the ULB. Users were overall satisfied with the data, but some weaknesses were identified. Many users pointed out the difficulties and challenges associated with uncertainties and errors, data filtering and interpretation of the data. While large retrieval uncertainty is inherent to the current-state-of-the art of NH_3 satellite measurements, improvement can be achieved in four areas:

Rec-NH3-1. Better traceability/documentation for both L2 and L3

While the NH_3 retrieval algorithm and its updates are well documented in a series of paper, a public ATBD describing the algorithm in full and a PUM would be welcomed. This should include example cases of using the data with advice/recommendations on data handling/quality filtering. Traceability is very important for the L2, but even more important for the L3.

Rec-NH3-2. Availability of unfiltered data

Many expert users preferred to be able to choose which quality flags (on retrieval sensitivity/cloudiness etc..) to apply, rather than a product that comes pre-filtered (which is currently the case).

Rec-NH3-3. Additional auxiliary variables

The addition of extra auxiliary variables would be welcomed by most users, to improve the interpretation and range of applicability of the data. The following were mentioned:

- Information which describes the vertical sensitivity of the measurement, such as a total column averaging kernel
- A more detailed uncertainty budget (total + first and second most important contributor)
- Indicator of aerosol/PM presence to complement quality flagging
- Metric on potential retrieval interferences (e.g. emissivity)
- NH_3 profile assumption

Rec-NH3-4. Additional validation

Several validation studies have been published (Van Damme et al., 2015, Dammers et al., 2016; Guo et al., 2021), but not enough robust conclusions were drawn from these related to the difficulty in obtaining representative (for NH_3 spatio-temporal variability) and reliable reference data (existing reference data have their own limitations). Validation is seen as something important and that should be repeated and improved at regular intervals.

2.8 User requirements summary for CHOCHO

Glyoxal data products are currently used for a range of applications, covering mapping, inter-sensor comparison (e.g. Hoque et al., 2018; Kluge et al., 2022; Lerot et al. 2021), evaluation of models (e.g. Stavrakou et al., 2009; Cao et al., 2018; Fu et al., 2008), data assimilation, trend analyses (Vrekoussis et al., 2009), monitoring of emissions (e.g. Li et al., 2021; Stavrakou et al., 2021, 2016; Liu et al., 2019) and identification of chemical regimes (e.g. Chan Miller et al., 2017; Kaiser et al., 2015, Guo et al., 2021). The TROPOMI glyoxal product has been developed as part of the ESA GLYRETRO innovation project and is distributed via a website hosted at BIRA-IASB. Most interviewed users relied on this product to provide their feedback. However, many of their answers apply to other species as well since many applications rely on multi-species analyses, in particular glyoxal is often used in complement of the HCHO and NO2 products (e.g. Guo et al., 2021; Alvarado et al., 2020; Chan Miller et al., 2017; Zarzana et al., 2017; Chan Miller et al., 2016; Liu et al., 2012). Below are the main recommendations and requests to make the generated products of extra interest.

Rec-CHOCHO-1. Diagnostic and traceability variables

Advanced users clearly stated a preference for L2 data sets in order to be able to use their own filtering methods, to apply the averaging kernels or to co-locate appropriately the satellite measurements with their own reference data. For this, the data products should contain a comprehensive set of diagnostic and traceability variables as

- Averaging kernels and prior estimates as well as the pressure level grid,
- Uncertainty estimates, possibly with random and systematic components separated,
- Simple *qa_value* variables complemented by a series of other flags or diagnostic variables to allow the application of personalized filtering schemes, particularly related to cloudiness.

Users interested by L3 data also requested some additional diagnostic variables such as the number of observations binned or the sum of weights to be included in the products in order to facilitate further regridding treatments. Where applicable, diagnostic variables from the L2 should be propagated in the L3 (e.g., propagated errors, mean AKs)

Rec-CHOCHO-2. Clean and stable data sets

Users have underlined the importance of having clean data sets, i.e. with artifacts filtered out, stable in time for trend analysis and with biases well characterized and removed if possible in order to have good inter-sensor consistency and to avoid jumps between data sets. Distribution of applied correction factors would also be an interesting piece of information. Even if it has greatly improved over the past years, it has been raised that the inter-product consistency might be further enhanced, for example in terms of formatting and error reporting.

Rec-CHOCHO-3. Guidance on data use

A general demand was to have better guidance on how to appropriately use the available diagnostic variables for filtering purposes. Some practical help on the way to apply the averaging kernels would also be welcome. Similarly, the correct interpretation of the provided error estimates remains unclear to some users and some more information on how they propagate would be useful too (e.g. how do they evolve when several observations are combined? Are they reduced or do they remain at the same level?) This additional guidance can be done via documentation (e.g. PUM, readme files), even if most users recognize it is already very informative. Another suggestion was to provide code snippets, which would illustrate some simple data uses.

Rec-CHOCHO-4. Traceability and documentation

Users judge the documentation to be an important aspect allowing them to evaluate the product performance (e.g. validation results). Overall, they find the current documentation of good quality. Some of them privilege technical reports; others search for information in peer-review literature, while another part of users prefers informal contacts (email exchanges, presentations). It is therefore recommended to carry on with this multi-channel communication.

Rec-CHOCHO-5. Improvement of data distribution

The research CHOCHO product is distributed via the GLYRETRO website. Experience has shown that it is relatively slow and currently doesn't allow automatization. Downloading massive amounts of TROPOMI level-2 data turns out to be cumbersome. Having one single and efficient platform to distribute all products would be beneficial. Options to easily download some data subset (subregions or limited period) would be of help.

2.9 Analysis of user requirements for different applications

Many users expressed similar and overlapping requirements for ECV precursor data. In Table 2 we zoom in on requirements that differ most for different applications. One striking difference between users who assimilate data into models or monitor emissions is that these require level-2 data, whereas users analyzing long-term trends require level-3 data, but much more advanced and complete data than what is currently provided.

Table 2: Differences in requirements for different user applications

User application	Users	Outstanding requirements
Decadal-scale data assimilation	Inness, Miyazaki, Qu	L2 data required with AK and uncertainties; L3 could be considered depending on how AK is provided
Long-term trend analysis	Georgoulias, Lamsal, Li,Gaubert	L3 data preferred; Intra-sensor corrections required for resolution, sampling and error differences; L3 data need extension with uncertainties, pixels used, time indicator
Emission monitoring	Riess, Fioletov	L2 data required with AK and uncertainties; more info on representativeness satellite observation (snapshot, vertical sensitivity) and additional data as wind, snow/ice

2.10 Analysis of climate modelers' requirements in the literature

Although we approached a variety of relevant users of ECV Precursor data, not many of those are from the climate user community. In general, climate modelers are not the most well-known users within the algorithm development teams. Still, users interested in long-term trend analysis and data assimilation have similar user requirements as climate modelers. To verify this, we reached out to one climate modeler at KNMI (dr. Twan van Noije) and asked him about his knowledge of studies that used ECV Precursor data sets in the climate modelling community. Dr. van Noije provided us with a list of 5 recent research papers that test, improve and evaluate their climate models with satellite ECV Precursor data. In Table 3 we analyze those studies specifically focusing on the question of what climate modelers require from the ECV Precursor data products for their meaningful use.

In general, climate modellers are aware of the need to evaluate their models by using averaging kernels, and by ensuring that the model and satellite measurements are co-sampled in (overpass) time and space. In their evaluation, they also require guidance and data on time-averaged (level-3) ECV Precursor observational uncertainty. Archibald et al. (2020) noted that the absence of information on tropopause height in MOPITT CO satellite data limited their analysis to some extent, however tropopause layer height are routinely provided in (level-2) ECV Precursor products such as from QA4ECV. Kluge et al. (2022) and Klonecki et al. (2012) applied additional flagging procedures in using the glyoxal (2022) and CO (2012) products, which were relatively new at the time of study. Especially for CO, a lot of experience with proper data usage (flagging) has been gained since, and two-way exchange between users and data providers has led to good consensus on what flagging and QA values to include in the data products.

Table 3: Overview of climate model papers using ECV Precursor satellite data

Study	Climate model	Purpose	ECV Precursor	Requirements expressed	Key result
Pozzer et al. (2022)	EMAC	Improve description of VOC oxidation for gas/aerosol	MOPITT CO	Monthly mean gridded L3 data Usage of AKs	Model likely overestimates fire emissions
Archibald et al. (2020)	UKESM1	Evaluation of all-atmosphere chemistry scheme	MOPITT CO OMI NO_2	L2 data Usage of AKs	Too little CO and too much NO_2 point at missing VOC chemistry
Kluge et al. (2022)	EMAC	Evaluate VOC emissions and oxidation scheme	S5P glyoxal	L2 data (own flagging, filtering) Usage of AKs	Model simulates too little glyoxal, indicating shortcomings in VOC chemistry

Klonecki et al. (2012)	LMDz-INCA	Improve CO concentrations	IASI CO	L2 data (own flagging, gridding) Usage of AKs	Improvement of model CO in the free troposphere, especially where CO emissions are too low (Asia)
Michou et al. (2011)	CNRM-CCM	Stratospheric chemistry evaluation	TOMS O_3 MIPAS NO_2 MLS CO	Monthly mean profile climatologies (L3 data)	Good simulation of stratospheric profiles

2.11 High-level Summary of User Requirements

Analyzing the overall response from all interviewed users of ECV precursor satellite data as well as a selection of peer-reviewed papers on climate model evaluation with ECV precursors, we draw the following conclusions:

- Most users prefer L2 data as this allows them to apply own flagging, gridding, averaging and uncertainty procedures
- Users investigating long-term trends and re-analysis of atmospheric composition indicated that L3 data would be useful to them
- Climate modelers use both L2 and L3 data to evaluate their decadal model runs

The main and most frequent requirements brought forward by *L2 users* are listed below (ranked by the number of times it was brought up):

(1) users request stable and consistent multi-sensor data records (common algorithm/inputs, bias estimates),
(2) they request simple (one-pager) guidance on proper data usage and interpretation.
(3) Users require demo code and examples on using averaging kernels, uncertainties, and *qa_values*. The newer ECV products (NH_3, CO, CHOCHO) are requested to include *qa_value*, averaging kernels, and uncertainty estimates, which are (partly) missing from existing datasets.
(4) The large data volume of S5P data is a general concern, whereas the data content and format of the S5P data products is the recommended standard for this ESA CCI ECV Precursor project.

L3 users expressed the need that

(1) L3 products be extended with time stamps, sampling and flagging details, co-gridded time-averaged uncertainties, and averaging kernels, which are currently lacking from L3 data files.

(2) They also request that multi-sensor L3 products are accompanied with correction factors to account for differences in spatial resolution, sampling, overpass time, and intra-sensor biases.

Other, more generic user requirements are the recommendation that satellite data records become findable via a central point of access and are accompanied with custom download tools (even though accessibility and findability via www.temis.nl, www.qa4ecv.eu, and https://iasi-aeris-data.fr is being appreciated). Users value the multiple communication channels available for the various ECV Precursor data products, such as technical documentation, scientific papers, as well as personal communication with the algorithm team. Last but not least, users request continuous validation and some recommend quantitative conclusions from the validation to accompany the ECV Precursor data products, for purposes of bias correction of the data, noting that they are aware of the potential pitfalls of such bias 'earmarks'.

3. GCOS requirements

3.1. GCOS Implementation Plan 2022: Status

Building on the 2021 GCOS Status Report that reviewed the state of climate observations and identified gaps and issues, GCOS finalized the new version 2022 of the GCOS Implementation Plan [RD-6], aiming to guide the development and improvement of the global climate observing system. A new update of the GCOS-IP is produced every 5-6 years. The 2022 version includes revised observing system requirements of all the Essential Climate Variables (ECVs), including ECV precursors, now with more details than available for past versions. The final version of the GCOS Implementation Plan 2022 [RD-6] was published in October 2022 on the GCOS website. Following this review all the comments have been examined by the GCOS expert panels and writing team before approval of the final version by the GCOS Steering Committee for submission to the UNFCCC, WMO and IOC. GCOS-IP 2022 was presented to the UNFCCC at COP 27 in Sharm el-Sheikh, Egypt, on 11 November 2022, and can now be considered final.

3.2. Existing requirements for NO_2, HCHO, SO_2, and CO

Here we elaborate on the GCOS specifications for a specification of the main user requirements. As the intended climate data records from this project will be static data sets (i.e., not operationally updated), we do not discuss the GCOS requirement on timeliness. The two levels in need of assessment are 'goal' and 'threshold', which span a range from the minimum requirement to be met to ensure that data are useful ('threshold'), and the 'ideal' requirement ('goal') above which no further improvements are necessary. Data products from the current and upcoming sensors are approaching 'goal', whereas the 'threshold' level is in most cases met by data products from historical sensors. We therefore consider a detailed evaluation of all data products against the GCOS level 'breakthrough', an intermediate between the 'goal' and 'threshold' requirements less useful (or even confusing) and refrain from including that evaluation here.

We note that most atmospheric composition users interviewed pay little attention to the GCOS requirements, as they are seldomly aware of them. There is thus only a weak link between how users perceive the usefulness of satellite data and the requirements stipulated by GCOS. Nevertheless, we reiterate the current requirements here, and evaluate the state-of-science retrieval products for NO_2, HCHO, SO_2 and CO against the current requirements in section 3.4.

3.2.1 GCOS specification for NO_2

Table 4 below lists the ECV requirements for tropospheric NO_2 columns as documented in the 2022 GCOS ECVs Requirements [RD-7] (Section 3.3.6).

Table 4: ECV requirements according [RD-7] (GCOS-245) for NO_2 Tropospheric Column. 'Threshold' requirement is the minimum to be met so that data are useful, and 'Goal' is an ideal requirement above which further improvements are not necessary and which would result in a significant improvement for satellite data applications, including climate monitoring.

NO_2 tropospheric column	**Goal**	**Threshold**
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	Once per 30 days
Measurement uncertainty (2-sigma)	max(20%, 1E+15 molec. cm^{-2})	max(100%, 5E+15)
Stability	max(4%, 1E+15 molec. cm^{-2}/decade)	max(20%, 1E+15 molec. cm^{-2}/decade)

3.2.2 GCOS specification for HCHO

Table 5: ECV requirements according [RD-7] (GCOS-245, section 3.3.3) for HCHO Tropospheric Column. 'Threshold' requirement is the minimum to be met so that data are useful, and 'Goal' is an ideal requirement above which further improvements are not necessary and which would result in a significant improvement for satellite data applications, including climate monitoring.

HCHO tropospheric column	**Goal**	**Threshold**
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	Once per 30 days
Measurement uncertainty (2-sigma)	max(20%, 8E+15 molec. cm^{-2})	max(100%, 40E+15 molec. cm-2)
Stability	max(4%, 8E+15 molec. cm^{-2}/decade)	max(20%, 8E+15 molec. cm^{-2}/decade)

3.2.3 GCOS specification for SO_2

Table 6: ECV requirements according [RD-7] (GCOS-245, section 3.3.4) for SO_2 Tropospheric Column. 'Threshold' requirement is the minimum to be met so that data are useful, and 'Goal' is an ideal requirement above which further improvements are not necessary and which would result in a significant improvement for satellite data applications, including climate monitoring.

SO2 tropospheric column	Goal	Threshold
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	30 days
Measurement uncertainty (2-sigma)	max(30%, 6E+15 molec. cm^{-2})	max(100%, 20E+15 molec. cm-2)
Stability	max(6%, 1.2E+15 molec. cm^{-2}/decade)	max(20%, 4E+15 molec. cm^{-2}/decade)

3.2.4 GCOS specification for CO

Table 7: ECV requirements according [RD-7] (GCOS-245, Section 3.3.1) for CO Tropospheric Column. 'Threshold' requirement is the minimum to be met so that data are useful, and 'Goal' is an ideal requirement above which further improvements are not necessary and which would result in a significant improvement for satellite data applications, including climate monitoring.

CO tropospheric column	Goal	Threshold
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	30 days
Measurement uncertainty (2-sigma)	1 ppb	10 ppb
Stability	< 1 ppb/decade	< 2 ppb/ decade

3.3. Design of requirements for NH_3 and CHOCHO

Quantitative requirements for other ECVs in this CCI project (NH_3 and glyoxal) not considered in the GCOS Implementation Plan 2022 [RD-6] nor in [RD-7] might appear later in 2023 for application areas -other than climate- currently discussed in WMO's Rolling Review of Requirements. Based on expert knowledge from algorithm developers and applications of the current state-of-science NH_3 and glyoxal satellite data, 'GCOS-equivalent' requirements are being proposed by the consortium in Tables 8 and 9 below.

3.3.1. Consortium specification for NH_3

Table 8 contains a proposal for 'GCOS-equivalent specifications' for NH_3 requirements, based on user consultation and expert judgment.

Table 8: Proposal for NH_3 requirements.

NH_3 tropospheric column	**Goal**	**Threshold**
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	Monthly
Measurement uncertainty (2-sigma)	max(50% , 2.5e+15 molec. cm^{-2})	max(100%, 1e+16 molec. cm^{-2})
Stability	max(2%, 1e+15 molec. cm^{-2} /decade)	max(10%, 2e+15 molec. cm^{-2} /decade)

3.3.2. Consortium specification for CHOCHO

In [RD-6] nor [RD-7], no requirement was defined for glyoxal. Glyoxal shares many applications with HCHO and similar numbers can be given for the horizontal and temporal resolutions. Values are proposed herebased on user feedback and data provider experience.

Table 9: Proposal for glyoxal requirements.

CHOCHO tropospheric column	**Goal**	**Threshold**
Horizontal resolution	10 km	100 km
Vertical resolution	Column	Column
Temporal resolution	Hourly	Once per 30 days
Measurement uncertainty (2-sigma)	max(20%, 4E+14 molec. cm^{-2})	max(100%, 8E+14 molec. cm^{-2})
Stability	max(10%, 1E+14 molec. cm^{-2}/decade)	max(25%, 2E+14 molec. cm^{-2}/decade)

3.4 Critical Assessment of GCOS requirements

3.4.1 Evaluation of GCOS requirements for NO_2

Here we compare the performance of two state-of-science NO_2 retrieval algorithms (KNMI/QA4ECV and KNMI TROPOMI) against the required performance specified by [RD-6] and [RD-7] listed in Table 3. For that purpose, we analysed scientific papers from the peer-reviewed literature published in the last 5 years. All NO_2 products fulfil the vertical resolution requirement. The frequency and resolution requirements are not met by any of the sensors capable of measuring NO_2 over the last 27 years, because of limitations in instrumental design, although TROPOMI nadir pixels (3.5km x 5.5km) do fulfil the resolution requirement set as 'goal'. Should the Nitrosat mission be selected by ESA, then the 'goal' requirement for horizontal resolution can even be considered as unambitious, since Nitrosat is intended to measure tropospheric NO_2 columns at the sub-kilometre scale.

Box 1: addressing 2-sigma uncertainties in [RD-6] and [RD-7] vs. 1-sigma uncertainties provided in the literature.

Various papers have discussed uncertainty requirements for the OMI QA4ECV NO_2 product. Those papers reported *single-pixel* 1-sigma uncertainties, i.e. the range of dispersion wherein 68% of the measured values would fall. However, [RD-6] and [RD-7] define the GCOS requirements as 2-sigma uncertainties, i.e. the range wherein 95% of the measured values would fall. Assuming that the tropospheric NO_2 data follows a normal distribution, we translate the reported *single-pixel* 1-sigma uncertainties into 2-sigma uncertainties by simply multiplying them with a factor 2. For the systematic component of the uncertainty however (the uncertainty remaining when averaging *multiple pixels* over space and time), we do not apply such a translation, as this uncertainty component represents the general, average deviation from the 'true' value.

Boersma et al. (2018) discussed that although the then applicable GCOS requirements (from 2015) on (1-sigma) uncertainty were not within reach at the time, they were an inspiration to reduce retrieval uncertainties, and to better estimate the systematic and random components of the retrieval uncertainty. QA4ECV NO_2 measurement uncertainties at pixel level are 70-90% (2-sigma), exceeding GCOS' goal, but they are generally well within the GCOS threshold requirement.

In Boersma et al. (2018) and in other papers (Miyazaki et al., 2015; Boersma et al., 2016) it is discussed that a substantial fraction of the uncertainty can be reduced by averaging individual pixels over space and time, as the errors are only partially correlated. To account for persistent errors in spatial or temporal means arising from error correlation (for example because the error in the a priori NO_2 profile on day 1 is correlated with that on day 2 because they are simulated with the same model), it is recommended to apply the principle of super-observations. For super-observations, the overall (reduced) measurement uncertainty is estimated as $\sigma_o = \sigma\sqrt{(1-c)/n + c}$ with c = 0.15 to account for correlation between (air mass factor) errors in surface reflectivity, clouds, a priori NO_2 profile, and aerosols at the spatiotemporal scales of models. After averaging, an unknown systematic uncertainty on the order of 20% remains. This kind of bias is related to air mass factor calculations and their related ancillary databases. It has led the product developers to state that for OMI QA4ECV

NO_2 “the typical measurement uncertainty of averaged columns is at or just above the GCOS requirement of 20%”.

In a follow-up validation study, Compernolle et al. (2020) found that the QA4ECV OMI NO_2 product is generally biased low relative to independent (MAX-DOAS) measurements, partly due to horizontal smoothing errors and differences in vertical sensitivity between the satellite and MAX-DOAS measurements. In a general sense, linking validation results such as presented in Compernolle et al. (2020) is difficult for NO_2 because of the strong spatial heterogeneity and variability. Nevertheless, the findings by Compernolle et al. (2020) indicate that the systematic uncertainty is indeed on the order or above the GCOS uncertainty requirement of 20%. *Further improvements in surface reflectivity, cloud retrievals, and (the resolution of) a priori NO_2 profiles are needed to reduce the systematic (air mass factor) uncertainty of (QA4ECV) NO_2 retrievals.*

Zara et al. (2018) discussed the stability of the QA4ECV NO_2 (slant) columns and found that these have a better than 0.2E+15 molec.cm^{-2}/decade stability. However, stability in tropospheric NO_2 columns is also influenced by instability in air mass factors, and a full stability analysis, including a long-term validation exercise, is much needed.

KNMI S5P NO_2 product

Since the beginning of nominal operation in April 2018, compliance of S5P NO_2 has been evaluated against ground-based validation measurements and satellite data from OMI and GOME-2. An extensive validation study by Verhoelst et al. (2021) points at a negative bias in the TROPOMI NO_2 tropospheric columns (version 1.4) of -20% to -50%, which is above the 20% ‘goal’ (target) uncertainty set by GCOS [RD-7]. More recent studies indicate that improvements in TROPOMI’s cloud pressure retrieval (Riess et al., 2022) have reduced the negative bias considerably, such that the systematic component of the uncertainty is estimated to now be close (but still slightly above) to 20% for version 2 of the TROPOMI NO_2 product (van Geffen et al. (2022); Riess et al. (2022)).

A recent study by Douros et al. (2022) demonstrated that replacing the global 1° resolution a priori information by the regional 0.1° resolution profiles of CAMS in the S5P NO_2 retrieval leads to increases of up to 30% in pollution hotspots. With these high-resolution profiles, they generated a new S5P NO_2 level-2 data product for Europe. This updated S5P product compares favorably to ground-based Pandora and MAX-DOAS instruments compared to the standard S5P NO_2 product (v1.2-1.4) with a 5-18% smaller (negative) bias in the tropospheric NO_2 columns. The data record of S5P is too short for a proper stability analysis.

Table 10: Evaluation of state-of-science tropospheric NO_2 product characteristics against GCOS requirements for NO_2. The achieved performance is indicated in red if the threshold level is not met, orange if the performance in in between the goal and threshold, and green if the goal is met.

	Horizontal resolution	**Temporal resolution**	**Measurement uncertainty (2-sigma)**	**Stability**
OMI QA4ECV NO_2 *Compernolle et al. (2020); Lorente et al. (2017); Boersma et al. (2018); Zara et al. (2018)*	13-60 km	Once per 2-3 days	**Random (2-sigma)** 70-90% / 1E15 - 10E15 molec.cm^{-2} **Random** Pixel: 1E+15 100km: <2E+14 **Systematic*** 20-30%	0.2+E15 molec.cm^{-2} /decade
TROPOMI NO_2 (version 1.4)	3.5-10 km	Once per 1-2 days	**Random (2-sigma)** 50%-100% **Random** Pixel: 1E+15 10km: 4E+14 100km: <2E+14 **Systematic*** +-20-30%	Not quantified
TROPOMI NO2 (version 2) *Riess et al. (2022), van Geffen et al. (2022), Douros et al., 2022*	3.5-10 km	Once per 1-2 days	**Random (2-sigma)** 15%-50% **Systematic*** +-15-25%	Not quantified
Required (G/T)	<10 / < 100	Hourly / once per 30 days	max(20%, 1E+15 molec. cm^{-2}) / max(100%, 5E+15)	max(4%/20%, 1E+15 molec. cm^{-2}/ decade)

***: Column-dependent bias**

3.4.2 Evaluation of GCOS requirements for HCHO

Since nominal operations started in April 2018, compliance of S5P HCHO has been evaluated against ground-based validation measurements and satellite data from OMI. Vigouroux et al. (2020) validated the operational TROPOMI HCHO product using a global network of Fourier transform infrared (FTIR) instruments. The study concluded that the TROPOMI HCHO columns present a negative bias over high concentrations sites (−31 % for HCHO columns larger than 8E+15 molec/cm²) and a positive bias for clean sites (+26 % for HCHO columns lower than 2.5E+15 molec/cm²). Based on clean sites, an upper limit of 1.3E+15 molec/cm² (1-sigma) was estimated for the deviation of daily observations at a spatial resolution of 20 km.

De Smedt et al. (2021) compared the OMI and TROPOMI HCHO observations using a global network of MAX-DOAS instruments to validate both satellite sensors for a large range of HCHO columns. Consistent with the FTIR validation study, it is reported that for elevated HCHO columns, TROPOMI data are systematically low (−25 % for HCHO columns larger than 8E+15 molec/cm²), while no significant bias is found for medium-range column values. OMI and TROPOMI data present equivalent biases for large HCHO levels. This kind of low-bias is related to AMF calculations and their related ancillary databases.

TROPOMI significantly improves the precision of the HCHO observations at short temporal scales and for low HCHO columns. Compared to OMI, the precision of the TROPOMI HCHO columns is improved by 25 % for individual pixels and by up to a factor of 3 when considering daily averages in 20 km radius circles. The validation precision obtained with daily TROPOMI observations is comparable to the one obtained with monthly OMI observations.

Table 11: Evaluation of state-of-science tropospheric HCHO products against GCOS requirements for HCHO. The achieved performance is indicated in red if the threshold level is not met, orange if the performance is in between the goal and threshold, and green if the goal is met

	Horizontal resolution	**Temporal resolution**	**Measurement uncertainty (2-sigma)**	**Stability**
OMI QA4ECV HCHO *De Smedt et al. (2021); Lorente et al. (2017); Zara et al. (2018)*	13-60 km	Once per 2-3 days	**Random** Pixel: 15-22E+15 10km: 10-20E+15 100km: 4-12E+15 **Systematic*** ±40%	Not quantified
TROPOMI HCHO *Vigouroux et al. (2020); De Smedt et al. (2021);*	3.5-10 km	Once per 1-2 days	**Random** Pixel: 10-15E+15 10km: 5-10E+15 100km: 2-6E+15 **Systematic*** ±40%	Not quantified
Required (G/T)	<10 / < 100	Hourly / once per 30 days	max(20%, 8E+15molec. cm-2) / max(100%, 40E+15)	max(4%, 1.6E+15 molec. cm-2/decade) / max(20%, 8E+15 molec. cm-2/decade)

***: Column-dependent bias**

3.4.3 Evaluation of GCOS requirements for SO_2

Because the envisaged algorithm for SO_2 retrievals (COBRA) is relatively new and currently only applied to TROPOMI, we compare here the performance of the two SO_2 retrieval algorithms of TROPOMI (operational DOAS-based and COBRA) against the required performance specified by GCOS listed in Table 12. For the other sensors (GOME, SCIAMACHY, OMI), an analysis of the space and time resolution requirement is possible, but we cannot really conclude about measurement uncertainty and stability requirements. All SO_2 products are column products and therefore fulfil the vertical resolution requirement. The threshold frequency requirement of 30 days is met by all sensors but none of the sensors are meeting the goal requirement of one hour. However, geostationary instruments like GEMS, TEMPO and Sentinel-4 UVN measure with an hourly sampling, but only over dedicated regions (East Asia, North America, and Europe). The horizontal resolution of the 100 km threshold is met by all sensors except GOME. The horizontal resolution requirements set as 'goal' are generally not fulfilled, except for TROPOMI nadir pixels (3.5km x 5.5km).

Validation of satellite SO_2 products is challenging especially for polluted scenes. Traditional ground-based validation only works at a very limited number of locations. Since the beginning of nominal operation of TROPOMI, the S5P SO_2 product has been evaluated against ground-based validation measurements and satellite data. The general conclusion is that large-scale positive biases are present in the operational product with VCD errors of up to 5E+15 molec/cm² (Fioletov et al., 2020), and are due to limitations from the spectral fitting step. This motivated the development of an alternative retrieval scheme (COBRA; Theys et al., 2021) that enabled the reduction of systematic offsets in the VCD data to less than 1E+15 molec/cm². Besides systematic errors from the spectral fitting part, both DOAS and COBRA algorithms are affected by errors on the air mass factors, which translate to VCD systematic errors of about 30-50%. For individual pixels, the random error sources dominate the total error budget, and account for 5-10E+16 molec/cm² and 2.5-5E+16 molec/cm² (at 2-sigma level), for DOAS and COBRA respectively. However, the random errors can be reduced (in principle) by data averaging by a factor $1/\sqrt{N}$, where N is the number of pixels considered.

Table 12 summarizes the TROPOMI uncertainty estimates and compares them to the required measurement uncertainty. Generally, the TROPOMI SO_2 products fulfil the threshold requirement, except for the random error for the operational data. As for the stability requirement, this has not been quantified partly because the TROPOMI data record is too short. It should be noted that for a species like SO_2 characterized by a small atmospheric background, the question of long-term stability is difficult to address. Over source regions, SO_2 emissions can change rapidly with time, making it difficult to assess the stability of a data record.

Table 12: Evaluation of state-of-science SO_2 products against GCOS requirements for SO_2. The achieved performance is indicated in red if the threshold level is not met, orange if the performance is in between the goal and threshold, and green if the goal is met.

	Horizontal resolution	**Temporal resolution**	**Measurement uncertainty (2-sigma)**	**Stability**
TROPOMI SO_2 DOAS (operational)	3.5-10 km	Once per 1-2 days	**Random*** (2-sigma): 5-10E+16 10 km: 2.5-5E+16 100 km: <6E+15 **Systematic:** 5e15 30%-50%	Not quantified
TROPOMI SO_2 COBRA *Theys et al., 2020*	3.5-10 km	Once per 1-2 days	**Random*** (2-sigma): 2.5-5E+16 10 km: 1.2-2.5E5+15 100 km: <6E+15 **Systematic:** 1E+15 30%-50%	Not quantified
Required (G/T)	<10 / < 100	Hourly / once per 30 days	max(30%, 6E+15 molec. cm^{-2}) / max(100%, 2E+16)	max(6%, 1.2E+15 molec. cm^{-2}/decade) (G), max(20%, 4E+15 molec. cm^{-2}/decade)

*single-pixel random uncertainty.

3.4.4 Evaluation of GCOS requirements for CO

Here we compare the performance of the IASI FORLI-CO product against the required performance specified by GCOS listed in Table 13. For that purpose, we analyzed "Validation Reports" generated in the framework of the AC SAF project. IASI CO products are validated with NDACC CO total column products, as well as with MOPITT satellite CO.

The AC SAF IASI CO validation report of 2021 (https://acsaf.org/docs/vr/Validation_Report_IASI-C_CO_May_2021.pdf) compares IASI CO L2 products against ground based FTIR measurement data available from 22 stations from NDACC (Network for the Detection of Atmospheric Composition Change) during the 4-year period 2017 – 2020. With an average of the relative differences of 3.7% (Metop-B) and 2.7% (Metop-C) as mentioned in Table 2.2 of the document, total column accuracy is similar or better than the uncertainty goal (12%) of CO GCOS requirements. Regarding the stability of the IASI CO product, Figure 3.12 of the report shows the temporal evolution of the IASI-A product since 2008 with the addition of the IASI-B in March 2013 and IASI-C in October 2019. IASI-C ensures the continuity of the mission as it is stable, and agrees very well globally, and on different latitude bands with the other two instruments. The stability of the IASI mission is also clearly shown with no apparent drift to any of the three instruments.

In the AC SAF IASI L3 CO validation report (under review), we compared IASI-A, B and IASI-C L3 products with MOPITT CO L3 version 8T products. An excellent agreement is found between the two products. Mean biases for IASI-A/MOPITT comparison range between -6.9% with a standard deviation of approximately 11% for thirteen 5°x5° regions distributed over the globe and between -7 and +5% for five cities (2°x2° boxes). For IASI-B/MOPITT, mean biases range between -7% and 13% for the 5°x5° regions and between -6% and +6% for the cities. As discussed in George et al. (2015), large differences are found during the winter months in the Northern Hemisphere (Canada, Eastern-USA, Europe, Siberia) and are partly explained by the different *a priori's* used in the different retrieval algorithms.

Table 13: Evaluation of state-of-science IASI CO products against GCOS requirements for CO. The achieved performance is indicated in red if the threshold level is not met, orange if the performance is in between the goal and threshold, and green if the goal is met.

	Horizontal resolution	Temporal resolution	Vertical resolution	Measurement uncertainty (2-sigma)	Stability
IASI CO	10 km	0.5 day	0-10/10-25 km	<12%	1% / decade
Required (G/T)	10 km / 200 km	Hourly / 30 days	column	1 ppb / 10 ppb	< 1 ppb/decade / < 3 ppb / decade

3.4.5 Evaluation of preliminary GCOS requirements for NH_3

In this section, we assess the proposed NH_3 GCOS requirements and evaluate how the current IASI product (ULB/Latmos) compares.

As far as the spatial requirement is concerned, the proposed goal and threshold values have been set in line with SO_2 and NO_2, the other main inorganic contributors to particulate matter. These were likely, at least in part, motivated by the satellite landscape at the time the requirements were first set out. NH_3 has a short atmospheric lifetime and is therefore both temporally as spatially highly variable. To constrain local variations and the multitude of small emission (point) sources the goal of 10 km is inadequate. A similar argument holds for NO_2. This motivated the Earth Explorer mission Nitrosat (currently still in competitive phase A) that would measure both species in sub-kilometre spatial resolution. As it stands, both IASI and the CrIS sounder almost meet the horizontal resolution requirement. The sounder IRS onboard MTG, to be launched in 2023, will already be capable of measuring at a resolution of 4 km (but geographically limited mainly to Europe/Africa).

The temporal resolution requirements, ranging from hourly to monthly are appropriate, as NH_3 both has large diurnal and seasonal variations. Depending on the specific application, even the threshold requirement can be considered enough (e.g. constraining the seasonal variations of agriculture), while for other applications (PM predictions), diurnal profiles are desirable. The sun-synchronous polar orbiting IASI and CrIS instruments both have bidaily global coverage, even though the evening/night-time overpass is characterized, on average, by less sensitivity. They thus fall in between the threshold and goal requirement. With 11 overpasses per day the Chinese GIIRS instrument onboard the geostationary FYII-A platform currently comes closest to meeting the goal temporal resolution requirement (Clarisse et al., 2021). IRS/MTG will offer overpasses each 30 min over its geographical target area, easily meeting the goal temporal resolution requirement.

Assessing the measurement uncertainty of NH_3 in a single number is unrealistic, as the measurement conditions (thermal contrast and vertical profile NH_3) are highly variable. In the current product, random uncertainty varies from as low as 30% to above 200%. As a general statement, the proposed measurements uncertainty requirements are reasonable, and achievable when measurement conditions are favourable. Validation, which could be used to assess systematic uncertainty is equally challenging. A recent, but limited validation dataset showed that the current IASI product might have systematic uncertainties of the order of 20% with a bias of the order of 2.10^{15} molec.cm^{-2} (Guo et al., 2021).

While the stability of the current IASI product has not been quantified, current instabilities are largely driven by instabilities in the input parameters. Currently, an IASI product is available that uses ERA5 data as far as vertical temperature profiles and humidity profiles are concerned, however there still exists inconsistencies related to the use of inhomogeneous proxies on cloud and surface temperature (Van Damme et al., 2017). These types of temporal inconsistencies are going to be addressed within the framework of the CCI+ project, notably with the inclusion of a new cloud product (Whitburn et al., 2022). As for L1 (radiance spectra), the IASI instrument itself is extremely stable, and the official release of the reprocessed L1C (http://doi.org/10.15770/EUM_SEC_CLM_0014) also guarantees that any changes in the processing from L0 to L1C data have been homogenized.

Table 14: Evaluation of state-of-science IASI NH3 products against GCOS requirements for NH_3. The achieved performance is indicated in red if the threshold level is not met, orange if the performance is in between the goal and threshold, and green if the goal is met.

	Horizontal resolution	**Temporal resolution**	**Measurement uncertainty (2-sigma)**	**Stability**
IASI NH_3	12-20 km	Bidaily	**Random*** (2-sigma): Highly variable **Systematic:** (Limited assessment) 22% + 5E+15 molec. cm^{-2}	Not quantified
Required (G/T)	<10 / < 100	Hourly / Monthly	max(50%, 5E+15 molec. cm^{-2}) / max(100%, 1E+16)	max(2%, 1E+15 molec. cm^{-2}/decade) (G), max(10%, 2E+15 molec. cm^{-2}/decade)

3.4.6 Evaluation of preliminary GCOS requirements for glyoxal

As they are relatively preliminary, it is of interest to compare the GCOS requirements with mission requirements defined for the future Sentinel-4 and 5 missions as well as for TEMPO.

No requirement on horizontal resolution and revisit time has been defined specifically for glyoxal but we can use those defined for the formaldehyde columns in the context of the S4/5 missions as those two species are useful for similar applications. The spatial requirement for HCHO has been set to 5/20 km (goal/threshold) for air quality applications and relaxed to 10/50 km for climate applications. It is therefore more stringent than the GCOS spatial requirement. The revisit time requirement is 0.5/2 hours for air quality applications in the context of S4/S5 and 2 measurements per day for the TEMPO mission. They can obviously not be met for space instruments boarded on LEO platforms such as TROPOMI but should be reachable in future with Sentinel-4 and TEMPO aboard geostationary platforms.

Regarding the requirements on uncertainties, they are defined differently for each mission and sometimes applicable only in defined validity domains (see Table 15). Like for the GCOS requirement, one single total uncertainty requirement is defined for Sentinel-4. On the contrary, two separate values are defined for the random and systematic components of the uncertainty in Sentinel-5. As for TEMPO, a requirement is defined only for the random component of the uncertainty.

In the following, those different requirements, including the GCOS values, are discussed regarding the current performance of the TROPOMI glyoxal product developed as part of the ESA innovation programme and of which the retrieval algorithm will serve as basis for further development within this project. Owing to the faint glyoxal signal, its associated (1-sigma) random uncertainty is large on individual measurements, in the range 6-8E+14 molec.cm^{-2} for TROPOMI (Lerot et al., 2021). Although generally reaching the S5 random uncertainty requirement, spatio-temporal averaging of several observations is needed to reduce this error component and meet the other defined values. For example, if one wants to meet a spatial resolution requirement of 20 km, we can afford to average 20-25 individual TROPOMI observations, which corresponds to a reduction of a random Gaussian error by a factor 4-5, making the defined requirement more easily reachable. For older sensors with coarser spatial resolution, the defined requirements can only be met by reducing the random uncertainty component also at the cost of the temporal resolution. Based on the proposed GCOS spatial and time threshold requirements (100 km, 30 days), the random error component can be reduced to small values even for heritage missions.

On the other hand, systematic errors cannot be reduced and are estimated by Lerot et al. (2021) to be in the range 1-3E+14 molec.cm^{-2} for clear sky pixels, corresponding to about 30-60% for emission regimes (columns larger than 2E+14 molec.cm^{-2}). This implies a compliance with the different requirements (including the proposed GCOS threshold values) for most cases.

Validation of satellite glyoxal column measurements is difficult and challenging because of scarcity of reference data and because retrievals from the ground suffer from the same limitations as from space. In addition, the retrieval strategy across the ground-based network is far from being homogenized, leading to inter-stations discrepancies. The TROPOMI glyoxal product has been validated in Lerot et al. (2021) with 8 MAX-DOAS stations. In general, the variabilities observed from space and from the ground agree well and overall absolute biases are reported to be less than 1E+14 molec.cm^{-2} for stations with moderate columns (within requirements). However, the bias exceeds the requirements at a few stations, which deserves further investigation.

Table 15: Evaluation of state-of-science tropospheric CHOCHO products against the proposed GCOS requirements for CHOCHO (Table 9). The achieved performance is indicated in red if the threshold level is not met, orange if the performance is in between the goal and threshold, and green if the goal is met

	Horizontal resolution	**Temporal resolution**	**Measurement uncertainty (2-sigma)**	**Stability**
TROPOMI CHOCHO	3.5x5.5 km²	Once a day	**Random** Pixel: 16E+14 10km: 8E+14 100km/Month: <1E+13	Not quantified
OMI CHOCHO	13-60x24 km²	Once per 2-3 days	Pixel: 20E+14 10km: 20E+14 100km/Month: 2E+14	
GOME2	80x40 km²	Once per 2 days	Pixel: 16E+14 molec.cm^{-2} 10km: 16E+14 100km/Month: 2.4E+14	
Lerot et al. (2021)			**Systematic (all sensors)** 2-3E+14 (40-60%)	
Required (G/T)	<10 / < 100	Hourly / once per 30 days	max(20%, 4E+14molec. cm-2) / max(100%, 8E+14)	max(10%, 1E+14 molec.cm^{-2}/deca) / max(25%, 2E+14 molec. cm^{-2}/deca)

3.5 Conclusions on GCOS requirements and their feasibility for the 6 ECV Precursors

The GCOS requirements on horizontal, vertical, and temporal resolution for the 6 ECV Precursors appear to be based on historical and existing sensor capabilities for 'threshold' requirements. The more ambitious 'goal' requirements on resolution would come within reach with the successful launch of future geostationary sensors (TEMPO, Sentinel-4), at least over the Northern Hemisphere continental regions. Some 'goal' GCOS requirements (on spatial resolution) are even rather unambitious if the Nitrosat mission is selected by ESA to measure NO_2 and NH_3 at the sub-kilometre scale.

Official GCOS uncertainty requirements have been established for 4 ECV Precursors (NO_2, HCHO, SO_2, and CO) [RD-7]. Error propagation studies, validation efforts, and intensive use for applications in the air quality, climate modelling, and data assimilation fields suggest that the ‘threshold’ uncertainty requirements are met for all 4 ECV precursors, and that all 4 ECV Precursors are approaching ‘goal’ uncertainty requirements.

An essential step in interpreting the uncertainty requirements, which is recommended to be included in future updates of the GCOS requirements, is to make the distinction between the random and systematic components of the reported uncertainties. Uncertainty estimates at individual pixel level contain contributions from random and systematic error contributions and seldomly meet ‘goal’ requirements. Via spatio-temporal averaging however the random components tend to cancel, so that users in practice need to be concerned mainly with systematic uncertainties which are currently at or exceeding the GCOS ‘goal’ requirements for NO_2, HCHO, SO_2, and CO. This consideration implies that requirements for climate applications (with relatively low spatio-temporal resolution of 25-50 km) are met more easily than requirements for air quality applications (high spatio-temporal resolution of 5-10 km).

There is a hierarchy in GCOS requirements for the DOAS-type retrievals of NO_2, HCHO, SO_2, and CO, with the more ambitious requirements set for NO_2 and the less ambitious for SO_2 and CO. This hierarchy merely reflects differences in vertical sensitivity to lower atmospheric pollutants. Owing to their absorption lines occurring towards the UV-part (~340 nm) of the satellite reflectance spectrum, satellite retrieval of SO_2 and HCHO contends with weaker vertical sensitivity than retrieval of NO_2, whose key absorption signatures are concentrated in the VIS (~440 nm) part of the spectrum, where photons propagate deeper into the lowest layers of the atmosphere (see e.g. Lorente et al. (2017)). For CO, the vertical information requirement (column only) is too loose for modelers who wish to use the CO data for inverse modelling of emission sources.

To be able to meet the ‘goal’ GCOS requirements, further improvements in satellite retrieval are needed. These mostly concern the better description of ancillary data and full physics for the purpose of better air mass factor calculations (for NO_2, HCHO, and SO_2). Scientific efforts to achieve the GCOS ‘goal’ requirements should focus on improving retrieval procedures in terms of the spatial resolution and physical realism of surface reflectivity, cloud parameters, a priori profile shapes, etc. Another widely felt gap is the lack of long-term, decadal-scale validation efforts for ECV precursors. Without continuous validation it is difficult to establish whether ECV Precursors fulfil the GCOS stability requirements.

3.6 Recommendations for future usage of GCOS requirements

The below list provides a summary of key remarks made during the discussion of the GCOS requirements at the CCI+ ECV Precursors 1st Progress Meeting on 4 October 2022.

	Issue	Recommendation
1.	GCOS requirements are currently generic, and not specific to identified user applications	Differentiate GCOS requirements specific to user application areas (e.g., global climate, linkage between climate and air quality, air quality)
2.	It is not clear where the quantitative GCOS requirements originate from (for each ECV: technical feasibility, scientific papers on uncertainty)	Visibility and and traceability of how GCOS specific requirements have been derived for each ECV should improve
3.	The definition for 'Goal' as an 'ideal requirement above which further improvements are not necessary' is not a good definition	Redefine 'Goal' requirements as "the median requirement of the top 20% users"
4.	Not clear if various types of requirements (on resolution, timeliness, uncertainty, and stability) are linked	GCOS IP should make clear if the types of requirements within a classification are linked or can be regarded as independent
5.	GCOS 'Goal' requirements on spatial resolution are not ambitious enough	The GCOS IP team should be aware that there are very good reasons why future sounders should target better spatial resolution than now considered as 'ideal'
6.	It is not sufficiently clear what GCOS measurement uncertainty refers to. Individual pixel uncertainties contain random and systematic components, but for many applications, the random contribution is averaged out and only the systematic uncertainty remains. This is not sufficiently recognized in GCOS vocabulary.	GCOS measurement uncertainty requirements should be defined for individual pixels (random + systematic), as well as for multiple, averaged pixels (systematic only)

4. References

Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Kalisz Hedegaard, A. B., Schneising, O., and Burrows, J. P.: Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus Sentinel-5 Precursor satellite during the 2018 Canadian wildfires, Atmos Chem Phys, 20, 2057–2072, https://doi.org/10.5194/acp-20-2057-2020, 2020.

Archibald, A. T., O'Connor, F. M., Abraham, N. L., Archer-Nicholls, S., Chipperfield, M. P., Dalvi, M., Folberth, G. A., Dennison, F., Dhomse, S. S., Griffiths, P. T., Hardacre, C., Hewitt, A. J., Hill, R. S., Johnson, C. E., Keeble, J., Köhler, M. O., Morgenstern, O., Mulcahy, J. P., Ordóñez, C., Pope, R. J., Rumbold, S. T., Russo, M. R., Savage, N. H., Sellar, A., Stringer, M., Turnock, S. T., Wild, O., and Zeng, G.: Description and evaluation of the UKCA stratosphere–troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1, Geosci. Model Dev., 13, 1223–1266, https://doi.org/10.5194/gmd-13-1223-2020, 2020.

Beirle, S., Hörmann, C., Penning de Vries, M., Dörner, S., Kern, C., and Wagner, T.: Estimating the volcanic emission rate and atmospheric lifetime of SO_2 from space: a case study for Kīlauea volcano, Hawaii, Atmos. Chem. Phys., 14, 8309–8322, https://doi.org/10.5194/acp-14-8309-2014, 2014.

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

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., et al.: 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.

Brenot, H., Theys, N., Clarisse, L., van Geffen, J., van Gent, J., Van Roozendael, M., van der A, R., Hurtmans, D., Coheur, P.-F., Clerbaux, C., Valks, P., Hedelt, P., Prata, F., Rasson, O., Sievers, K., and Zehner, C.: Support to Aviation Control Service (SACS): an online service for near real-time satellite monitoring of volcanic plumes, Nat. Hazards Earth Syst. Sci., 14, 1099-1123, doi:10.5194/nhess-14-1099-2014, 2014.

Buchholz R., Worden H., Park M., Francis G., Deeter M., Edwards D., Emmons L., Gaubert B., Gille J., Martínez-Alonso S., Tang W., Kumar R., Drummond J. R., Clerbaux C., George M., Coheur P.-F., Hurtmans D., Bowman K. W., Luo M., Payne V. H., Worden J. R., Chin M., Levy R. C., Warner J., Wei Z. and Kulawik S. S.: Air pollution trends measured from Terra: CO and AOD over industrial, fire-prone, and background regions, Remote Sen. Environ., Elsevier, 256 (April), doi: 10.1016/j.rse.2020.112275, 2021.

Cao, H., Fu, T.-M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., de Smedt, I., Zhang, Q., van Roozendael, M., Hendrick, F., Chance, K., Li, J., Zheng, J., and Zhao, Y.: Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, Atmos Chem Phys, 18, 15017–15046, https://doi.org/10.5194/acp-18-15017-2018, 2018.

Carn, S. A., Fioletov, V. E., McLinden, C. A., Li, C., and Krotkov, N. A.: A decade of global volcanic SO2 emissions measured from space, Sci. Rep.-UK, 7, 44095, https://doi.org/10.1038/srep44095, 2017.

Chan Miller, C., Jacob, D. J., González Abad, G., and Chance, K.: Hotspot of glyoxal over the Pearl River delta seen from the OMI satellite instrument: implications for emissions of aromatic hydrocarbons, Atmos Chem Phys, 16, 4631–4639, https://doi.org/10.5194/acp-16-4631-2016, 2016.

Chan Miller, C., Jacob, D. J., Marais, E. A., Yu, K., Travis, K. R., Kim, P. S., Fisher, J. A., Zhu, L., Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Kaiser, J., Min, K. E., Brown, S. S., Washenfelder, R. A., González Abad, G., and Chance, K.: Glyoxal yield from isoprene oxidation and relation to formaldehyde: Chemical mechanism, constraints from

SENEX aircraft observations, and interpretation of OMI satellite data, Atmos Chem Phys, 17, 8725–8738, https://doi.org/10.5194/acp-17-8725-2017, 2017.

Chen, Y., Shen, H., Kaiser, J., Hu, Y., Capps, S. L., Zhao, S., ... & Russell, A. G. High-resolution hybrid inversion of IASI ammonia columns to constrain US ammonia emissions using the CMAQ adjoint model. *Atmospheric Chemistry and Physics*, 21(3), 2067-2082, 2021.

Clarisse, L., Van Damme, M., Hurtmans, D., Franco, B., Clerbaux, C., & Coheur, P. F.: The Diel Cycle of NH_3 Observed From the FY-4A Geostationary Interferometric Infrared Sounder (GIIRS). *Geophysical Research Letters*, *48*(14), e2021GL093010, 2021.

Compernolle, S., Verhoelst, T., Pinardi, G., Granville, J., Hubert, D., et al.: Validation of Aura-OMI QA4ECV NO2 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.

Dammers, E., Palm, M., Van Damme, M., Vigouroux, C., Smale, D., Conway, S., ... & Erisman, J. W. An evaluation of IASI-NH3 with ground-based Fourier transform infrared spectroscopy measurements. *Atmospheric Chemistry and Physics*, 16(16), 10351-10368, 2016.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendael, M.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmos. Chem. Phys., 15, 12519–12545, https://doi.org/10.5194/acp-15-12519-2015, 2015.

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

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A.-M., Peuch, V.-H., Colette, A., and Veefkind, P.: Comparing Sentinel-5P TROPOMI NO2 column observations with the CAMS-regional air quality ensemble, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2022-365, 2022.

Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., ... & Ziemba, L. D. (2014). Satellite data of atmospheric pollution for US air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmospheric Environment*, *94*, 647-662.

Eskes, H., van Geffen, J., Boersma, K.F., Eichmann, K.-U., Apituley, A., Pedergnana, M., Sneelp, M., Veefkind, J. P., and Loyola, D., Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Nitrogen dioxide, S5P-KNMI-L2-0021-MA, CI-7570-PUM, issue 4.1.0, 2022.

Fioletov, V., McLinden, C. A., Griffin, D., Theys, N., Loyola, D. G., Hedelt, P., Krotkov, N. A., and Li, C.: Anthropogenic and volcanic point source SO2 emissions derived from TROPOMI on board Sentinel-5 Precursor: first results, Atmos. Chem. Phys., 20, 5591–5607, https://doi.org/10.5194/acp-20-5591-2020, 2020.

Fioletov, Vitali, Chris A. McLinden, Debora Griffin, Ihab Abboud, Nickolay Krotkov, Peter J. T. Leonard, Can Li, Joanna Joiner, Nicolas Theys, and Simon Carn, Version 2 of the global catalogue of large anthropogenic and volcanic SO2 sources and emissions derived from satellite measurements, Earth Syst. Sci. Data, 15, 75–93, https://doi.org/10.5194/essd-15-75-2023, 2023.

Fu, T.-M. T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M., and Henze, D. K.: Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, J Geophys Res, 113, D15303, https://doi.org/10.1029/2007JD009505, 2008.

George, M., Clerbaux, C., Bouarar, I., Coheur, P.-F., Deeter, M. N., Edwards, D. P., Francis, G., Gille, J. C., Hadji-Lazaro, J., Hurtmans, D., Inness, A., Mao, D. and Worden, H. M.: An examination of the long-term CO records from MOPITT and IASI: comparison of retrieval methodology, Atmospheric Measurement Techniques, 8, 4313-4328, doi: 10.5194/amt-8-4313-2015, 2015.

Guo, X., Wang, R., Pan, D., Zondlo, M. A., Clarisse, L., Van Damme, M., et al.: Validation of IASI satellite ammonia observations at the pixel scale using in situ vertical profiles, Journal of Geophysical Research: Atmospheres, 126, e2020JD033475. https://doi.org/10.1029/2020JD033475, 2021.

Guo, Y., Wang, S., Zhu, J., Zhang, R., Gao, S., Saiz-Lopez, A., and Zhou, B.: Atmospheric formaldehyde, glyoxal and their relations to ozone pollution under low- and high-NOx regimes in summertime Shanghai, China, Atmos Res, 105635, https://doi.org/10.1016/j.atmosres.2021.105635, 2021.

Hoque, H. M. S., Irie, H., and Damiani, A.: First MAX-DOAS Observations of Formaldehyde and Glyoxal in Phimai, Thailand, Journal of Geophysical Research: Atmospheres, 123, 9957–9975, https://doi.org/10.1029/2018JD028480, 2018.

Jiang, Z., Zhu, R., Miyazaki, K., McDonald, B. C., Klimont, Z., Zheng, B., Boersma, K. F., Zhang, Q., Worden, H., Worden, J. R., Henze, D. K., Jones, D. B. A., Denier van der Gon, H. A. C., Eskes, H.: Decadal variabilities in tropospheric nitrogen oxides over United States, Europe, and China. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035872. https://doi.org/10.1029/2021JD035872, 2022.

Jin, X., Fiore, A., Boersma, K. F., Smedt, I. D., & Valin, L. (2020). Inferring changes in summertime surface Ozone–NO x–VOC chemistry over US urban areas from two decades of satellite and ground-based observations. *Environmental science & technology*, *54*(11), 6518-6529.

Jung, J., Choi, Y., Mousavinezhad, S., Kang, D., Park, J., Pouyaei, A., Ghahremanloo, M., Momeni, M., & Kim, H.: Changes in the ozone chemical regime over the contiguous United States inferred by the inversion of Nox and VOC emissions using satellite observation. *Atmospheric Research*, *270*, 106076. https://doi.org/10.1016/J.ATMOSRES.2022.106076, 2022.

Kaiser, J., Wolfe, G. M., Min, K. E., Brown, S. S., Miller, C. C., Jacob, D. J., deGouw, J. A., Graus, M., Hanisco, T. F., Holloway, J., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Washenfelder, R. A., and Keutsch, F. N.: Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation, Atmos Chem Phys, 15, 7571–7583, https://doi.org/10.5194/acp-15-7571-2015, 2015.

Klonecki, A., Pommier, M., Clerbaux, C., Ancellet, G., Cammas, J.-P., Coheur, P.-F., Cozic, A., Diskin, G. S., Hadji-Lazaro, J., Hauglustaine, D. A., Hurtmans, D., Khattatov, B., Lamarque, J.-F., Law, K. S., Nedelec, P., Paris, J.-D., Podolske, J. R., Prunet, P., Schlager, H., Szopa, S., and Turquety, S.: Assimilation of IASI satellite CO fields into a global chemistry transport model for validation against aircraft measurements, Atmos. Chem. Phys., 12, 4493–4512, https://doi.org/10.5194/acp-12-4493-2012, 2012.

Kluge, F., Hüneke, T., Lerot, C., Rosanka, S., Rotermund, M. K., Taraborrelli, D., Weyland, B., and Pfeilsticker, K.: Airborne glyoxal measurements in the marine and continental atmosphere: Comparison with TROPOMI observations and EMAC simulations, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2022-416, in review, 2022.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J.

W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C. and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmospheric Chemistry and Physics, 13, 4073-4109, doi: 10.5194/acp-13-4073-2013, 2013.

Konovalov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Beekmann, M., Hadji-Lazaro, J., Clerbaux, C., Andreae, M. O., Kaiser, J. W. and Schulze, E.-D.: Constraining CO2 emissions from open biomass burning by satellite observations of co-emitted species: a method and its application to wildfires in Siberia, Atmospheric Chemistry and Physics, 14, 10383-10410, doi: 10.5194/acp-14-10383-2014, 2014.

Krol, M., Peters, W., Hooghiemstra, P., George, M., Clerbaux, C., Hurtmans, D., McInerney, D., Sedano, F., Bergamaschi, P., El Hajj, M., Kaiser, J. W., Fisher, D., Yershov, V. and Muller, J.-P.: How much CO was emitted by the 2010 fires around Moscow?, Atmospheric Chemistry and Physics, 13, 4737-4747, doi: 10.5194/acp-13-4737-2013, 2013.

Krotkov, N. A., McLinden, C. A., Li, C., Lamsal, L. N., Celarier, E. A., Marchenko, S. V., Swartz, W. H., Bucsela, E. J., Joiner, J., Duncan, B. N., Boersma, K. F., Veefkind, J. P., Levelt, P. F., Fioletov, V. E., Dickerson, R. R., He, H., Lu, Z., and Streets, D. G.: Aura OMI observations of regional SO2 and NO2 pollu- tion changes from 2005 to 2015, Atmos. Chem. Phys., 16, 4605– 4629, https://doi.org/10.5194/acp-16-4605-2016, 2016.

Lachatre, M., Fortems-Cheiney, A., Foret, G., Siour, G., Dufour, G., Clarisse, L., ... & Beekmann, M. The unintended consequence of SO_2 and NO_2 regulations over China: increase of ammonia levels and impact on PM 2.5 concentrations. *Atmospheric Chemistry and Physics*, 19(10), 6701-6716, 2019.

Lange, K., Richter, A., and Burrows, J. P.: Variability of nitrogen oxide emission fluxes and lifetimes estimated from Sentinel-5P TROPOMI observations, Atmos. Chem. Phys., 22, 2745–2767, https://doi.org/10.5194/acp-22-2745-2022, 2022.

Lerot, C., Hendrick, F., Van Roozendael, M., Alvarado, L. M. A., Richter, A., De Smedt, I., Theys, N., Vlietinck, J., Yu, H., Van Gent, J., Stavrakou, T., Müller, J.-F., Valks, P., Loyola, D., Irie, H., Kumar, V., Wagner, T., Schreier, S. F., Sinha, V., Wang, T., Wang, P., and Retscher, C.: Glyoxal tropospheric column retrievals from TROPOMI – multi-satellite intercomparison and ground-based validation, Atmos. Meas. Tech., 14, 7775–7807, https://doi.org/10.5194/amt-14-7775-2021, 2021.

Li, B., Ho, S. S. H., Li, X., Guo, L., Chen, A., Hu, L., Yang, Y., Chen, D., Lin, A., and Fang, X.: A comprehensive review on anthropogenic volatile organic compounds (VOCs) emission estimates in China: Comparison and outlook, Environ Int, 156, 106710, https://doi.org/10.1016/j.envint.2021.106710, 2021.

Li, K., Jacob, D. J., Shen, L., Lu, X., de Smedt, I., & Liao, H.: 2013—2019 increases of surface ozone pollution in China: anthropogenic and meteorological influences. Atmospheric Chemistry and Physics Discussions, 2020, 1–18, 2020.

Liu, T., Marlier, M. E., Karambelas, A., Jain, M., Singh, S., Singh, M. K., Gautam, R., and DeFries, R. S.: Missing emissions from post-monsoon agricultural fires in northwestern India: regional limitations of MODIS burned area and active fire products, Environ Res Commun, 1, 011007, https://doi.org/10.1088/2515-7620/AB056C, 2019.

Liu, Z., Wang, Y., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., Shao, M., Chang, C.-C., Liu, S.-C., Wang, H., and Chen, C.: Exploring the missing source of glyoxal (CHOCHO) over China, Geophys. Res. Lett., 39, L10812, https://doi.org/10.1029/2012GL051645, 2012.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J.-T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J., and Krol, M.: Structural uncertainty in air mass factor calculation for NO_2 and HCHO satellite retrievals, Atmos. Meas. Tech., 10, 759-782, doi:10.5194/amt-10-759-2017, 2017.

Lutsch, E., Strong, K., Jones, D. B., Ortega, I., Hannigan, J. W., Dammers, E., ... & Fisher, J. A.. Unprecedented atmospheric ammonia concentrations detected in the high Arctic from the 2017 Canadian wildfires. *Journal of Geophysical Research: Atmospheres*, 124(14), 8178-8202, 2019.

Michou, M., Saint-Martin, D., Teyssèdre, H., Alias, A., Karcher, F., Olivié, D., Voldoire, A., Josse, B., Peuch, V.-H., Clark, H., Lee, J. N., and Chéroux, F.: A new version of the CNRM Chemistry-Climate Model, CNRM-CCM: description and improvements from the CCMVal-2 simulations, Geosci. Model Dev., 4, 873–900, https://doi.org/10.5194/gmd-4-873-2011, 2011.

Miyazaki, K., Eskes, H. J., and Sudo, K.: A tropospheric chemistry reanalysis for the years 2005–2012 based on an assimilation of OMI, MLS, TES, and MOPITT satellite data, Atmos. Chem. Phys., 15, 8315–8348, https://doi.org/10.5194/acp-15-8315-2015, 2015.

Morfopoulos, C., Müller, J., Stavrakou, T., Bauwens, M., de Smedt, I., Friedlingstein, P., Prentice, I. C., & Regnier, P.: Vegetation responses to climate extremes recorded by remotely sensed atmospheric formaldehyde. *Global Change Biology*, *28*(5), 1809–1822. https://doi.org/10.1111/gcb.15880, 2022.

Nechita-Banda N., M. Krol, G. R. van der Werf, J. W. Kaiser, S. Pandey, V. Huijnen, C. Clerbaux, P. Coheur, M. N. Deeter, T. Röckmann: Monitoring emissions from the 2015 Indonesian fires using CO satellite data, Philosophical Transactions B, DOI: 10.1098/rstb.2017.0307, 2018.

Opacka, B., Müller, J.-F., Stavrakou, T., Miralles, D. G., Koppa, A., Pagán, B. R., Potosnak, M. J., Seco, R., de Smedt, I., & Guenther, A. B.: Impact of Drought on Isoprene Fluxes Assessed Using Field Data, Satellite-Based GLEAM Soil Moisture and HCHO Observations from OMI. *Remote Sensing*, *14*(9), 2021. https://doi.org/10.3390/rs14092021, 2022.

Peng, Z., Lei, L., Liu, Z., Sun, J., Ding, A., Ban, J., Chen, D., Kou, X., and Chu, K.: The impact of multi-species surface chemical observation assimilation on air quality forecasts in China, Atmos. Chem. Phys., 18, 17387–17404, https://doi.org/10.5194/acp-18-17387-2018, 2018.

Pozzer, A., Reifenberg, S. F., Kumar, V., Franco, B., Kohl, M., Taraborrelli, D., Gromov, S., Ehrhart, S., Jöckel, P., Sander, R., Fall, V., Rosanka, S., Karydis, V., Akritidis, D., Emmerichs, T., Crippa, M., Guizzardi, D., Kaiser, J. W., Clarisse, L., Kiendler-Scharr, A., Tost, H., and Tsimpidi, A.: Simulation of organics in the atmosphere: evaluation of EMACv2.54 with the Mainz Organic Mechanism (MOM) coupled to the ORACLE (v1.0) submodel, Geosci. Model Dev., 15, 2673–2710, https://doi.org/10.5194/gmd-15-2673-2022, 2022.

Qu, Z., Daven K. Henze, Nicolas Theys, Jun Wang, Wei Wang : Hybrid mass balance/4D-Var joint inversion of NO_x and SO_2 emissions in East Asia, J. Geophys. Res.-Atmos., https://doi.org/10.1029/2018JD030240, 2019.

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 NO2 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.

Schmidt, A., S. Leadbetter, N. Theys, E. Carboni, C. S. Witham, J. A. Stevenson, C. E. Birch, T. Thordarson, S. Turnock, S. Barsotti, *et al.* (2015), Satellite detection, long-range transport and air quality impacts of volcanic sulfur dioxide from the 2014–15 flood lava eruption at Bárðarbunga (Iceland), J. Geophys. Res. Atmos., 120, *doi:10.1002/2015JD023638.*

Shah, V., Jacob, D. J., Dang, R., Lamsal, L. N., Strode, S. A., Steenrod, S. D., Boersma, K. F., Eastham, S. D., Fritz, T. M., Thompson, C., Peischl, J., Bourgeois, I., Pollack, I. B., Nault, B. A., Cohen, R. C., Campuzano-Jost, P., Jimenez,

J. L., Andersen, S. T., Carpenter, L. J., Sherwen, T., and Evans, M. J.: Nitrogen oxides in the free troposphere: Implications for tropospheric oxidants and the interpretation of satellite NO_2 measurements, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2022-656, 2022.

Sodemann, H., Pommier, M., Arnold, S. R., Monks, S. A., Stebel, K., Burkhart, J. F., Hair, J. W., Diskin, G. S., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Schlager, H., Blechschmidt, A.-M., Kristjánsson, J. E., and Stohl, A.: Episodes of cross-polar transport in the Arctic troposphere during July 2008 as seen from models, satellite, and aircraft observations, Atmospheric Chemistry and Physics, 11, 3631-3651, doi: 10.5194/acp-11-3631-2011, 2011.

Stavrakou, T., Müller, J.-F., de Smedt, I., van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, Atmos Chem Phys, 9, 8431–8446, 2009.

Stavrakou, T., Müller, J.-F., Bauwens, M., de Smedt, I., Lerot, C., van Roozendael, M., Coheur, P.-F., Clerbaux, C., Boersma, K. F., van der A, R., and Song, Y.: Substantial Underestimation of Post-Harvest Burning Emissions in the North China Plain Revealed by Multi-Species Space Observations, Nature Publishing Group, https://doi.org/10.1038/srep32307, 2016.

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., and Guenther, A.: Impact of Short-Term Climate Variability on Volatile Organic Compounds Emissions Assessed Using OMI Satellite Formaldehyde Observations, Geophys. Res. Lett., 45, 8681–8689, 2018.

Stavrakou, T., Müller, J.-F., Bauwens, M., Doumbia, T., Elguindi, N., Darras, S., Granier, C., Smedt, I. de, Lerot, C., Roozendael, M. van, Franco, B., Clarisse, L., Clerbaux, C., Coheur, P.-F., Liu, Y., Wang, T., Shi, X., Gaubert, B., Tilmes, S., and Brasseur, G.: Atmospheric Impacts of COVID-19 on NOx and VOC Levels over China Based on TROPOMI and IASI Satellite Data and Modeling, Atmosphere 2021, Vol. 12, Page 946, 12, 946, https://doi.org/10.3390/ATMOS12080946, 2021.

Sun, W., Zhu, L., Pu, D., Chen, Y., Fu, T.-M., Wang, X., Yang, X., & de Smedt, I.: Global Changes in Formaldehyde (HCHO) Columns Observed from Space during the Early Stage of the COVID-19 Pandemic. *AGU Fall Meeting Abstracts*, *2020*, A005—0005, 2020.

Theys, N., P. Hedelt, I. De Smedt, C. Lerot, H. Yu, J. Vlietinck, M. Pedergnana, S. Arellano, B. Galle, D. Fernandez, C.J.M. Carlito, C. Barrington, B. Taisne, H. Delgado-Granados, D. Loyola, M. Van Roozendael: Global monitoring of volcanic SO2 degassing with unprecedented resolution from TROPOMI onboard Sentinel-5 Precursor, Nature Scientific Reports, volume 9, Article number: 2643, https://doi.org/10.1038/s41598-019-39279-y, 2019.

Theys, N., Volkamer, R., Müller, J. F., Zarzana, K. J., Kille, N., Clarisse, L., De Smedt, I., Lerot, C., Finkenzeller, H., Hendrick, F., Koenig, T. K., Lee, C. F., Knote, C., Yu, H., and Van Roozendael, M.: Global nitrous acid emissions and levels of regional oxidants enhanced by wildfires, Nat. Geosci., 13, 681–686, https://doi.org/10.1038/s41561-020-0637-7, 2020.

Turquety, S., D. Hurtmans, J. Hadji-Lazaro, P.-F. Coheur, C. Clerbaux, D. Josset, and C. Tsamalis: Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: analysis of the summer 2007 Greek fires, Atmospheric Chemistry and Physics, 9, 4897-4913, doi: doi.org/10.5194/acp-9-4897-2009, 2009.

Van Damme, M., Clarisse, L., Dammers, E., Liu, X., Nowak, J. B., Clerbaux, C., ... & Coheur, P. F. Towards validation of ammonia (NH_3) measurements from the IASI satellite. *Atmospheric Measurement Techniques*, 8(3), 1575-1591, 2015.

Van Damme, M., Clarisse, L., Whitburn, S., Hadji-Lazaro, J., Hurtmans, D., Clerbaux, C., & Coheur, P. F.. Industrial and agricultural ammonia point sources exposed. *Nature*, 564(7734), 99-103, 2018.

Van Damme, M., Clarisse, L., Franco, B., Sutton, M. A., Erisman, J. W., Kruit, R. W., ... & Coheur, P. F. : Global, regional and national trends of atmospheric ammonia derived from a decadal (2008–2018) satellite record. *Environmental Research Letters*, *16*(5), 055017, 2021.

van der A, R., B. Mijling, J. Ding, M. Koukouli, F. Liu, Q. Li, H. Mao, and N. Theys: Cleaning up the air: Effectiveness of air quality policy for SO_2 and NO_x emissions in China, Atmos. Chem. Phys., 17, 1775-1789, doi:10.5194/acp-17-1775-2017, 2017.

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 NO2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data, Atmos. Meas. Tech., 15, 2037–2060, https://doi.org/10.5194/amt-15-2037-2022, 2022.

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J.-C., Eskes, H. J., et al.: Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO2 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.

Whitburn, S., Clarisse, L., Crapeau, M., August, T., Hultberg, T., Coheur, P. F., and Clerbaux, C.: A CO_2-independent cloud mask from Infrared Atmospheric Sounding Interferometer (IASI) radiances for climate applications, Atmos. Meas. Tech., 15, 6653–6668, https://doi.org/10.5194/amt-15-6653-2022, 2022.

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

Vohra, K., Marais, E. A., Bloss, W. J., Schwartz, J., Mickley, L. J., Van Damme, M., ... & Coheur, P. F. Rapid rise in premature mortality due to anthropogenic air pollution in fast-growing tropical cities from 2005 to 2018. *Science Advances*, 8(14), eabm4435, 2022.

Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: Temporal and spatial variability of glyoxal as observed from space, Atmos Chem Phys, 9, 4485–4504, https://doi.org/10.5194/acp-9-4485-2009, 2009.

Wang, P., Holloway, T., Bindl, M., Harkey, M., & de Smedt, I.: Ambient Formaldehyde over the United States from Ground-Based (AQS) and Satellite (OMI) Observations. *Remote Sensing*, *14*(9), 2191. https://doi.org/10.3390/rs14092191, 2022.

Wang, R., Guo, X., Pan, D., Kelly, J. T., Bash, J. O., Sun, K., ... & Zondlo, M. A. Monthly Patterns of Ammonia Over the Contiguous United States at 2-km Resolution. Geophysical research letters, 48(5), e2020GL090579, 2021.

Wang, Y., Wang, Z., Yu, C., Zhu, S., Cheng, L., Zhang, Y., and Chen, L.: Validation of OMI HCHO Products Using MAX-DOAS observations from 2010 to 2016 in Xianghe, Beijing: Investigation of the Effects of Aerosols on Satellite Products, Remote Sens., 11, 203, https://doi.org/10.3390/rs11020203, 2019.

Wei, J., Liu, S., Li, Z., Liu, C., Qin, K., Liu, X., Pinker, R., Dickerson, 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, Environmental Science & Technology, 56(14), https://doi.org/10.1021/acs.est.2c03834, 9988–9998, 2022.

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

Zara, M., Boersma, K. F., Eskes, H., Denier van der Gon, H., Vilà-Guerau de Arellano, J., Krol, M., van der Swaluw, E., Schuch, W., and Velders, G. J. M.: 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, Atm. Environm. X, 9, 100104, https://doi.org/10.1016/j.aeaoa.2021.100104, 2021.

Zarzana, K. J., Min, K.-E., Washenfelder, R. A., Kaiser, J., Krawiec-Thayer, M., Peischl, J., Neuman, J. A., Nowak, J. B., Wagner, N. L., Dubè, W. P., Clair, J. M. st., Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Ryerson, T. B., and Brown, S. S.: Emissions of Glyoxal and Other Carbonyl Compounds from Agricultural Biomass Burning Plumes Sampled by Aircraft, Environ Sci Technol, 51, 11761–11770, https://doi.org/10.1021/ACS.EST.7B03517, 2017.

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.

Zhao, T., Mao, J., Simpson, W. R., de Smedt, I., Zhu, L., Hanisco, T. F., Wolfe, G. M., St Clair, J. M., González Abad, G., Nowlan, C. R., & others.: Source and variability of formaldehyde (HCHO) at northern high latitude: an integrated satellite, ground/aircraft, and model study. Atmospheric Chemistry and Physics Discussions, 1–55, 2021.

Zhu, L., González Abad, G., Nowlan, C. R., Chan Miller, C., Chance, K., Apel, E. C., DiGangi, J. P., Fried, A., Hanisco, T. F., Hornbrook, R. S., Hu, L., Kaiser, J., Keutsch, F. N., Permar, W., St. Clair, J. M., and Wolfe, G. M.: Validation of satellite formaldehyde (HCHO) retrievals using observations from 12 aircraft campaigns, Atmos. Chem. Phys., 20, 12329–12345, https://doi.org/10.5194/acp-20-12329-2020, 2020.

Annex 1: Example of a completed questionnaire

Version 3, 25 August 2022

Folkert Boersma (KNMI)
with contributions from Isabelle De Smedt, Nicolas Theys, Christophe Lerot (BIRA), Lieven Clarisse (ULB), and Maya George (LATMOS)

ESA started the Precursors for Aerosols and Ozone CCI project in July 2022. This project is developing long-term climate data records of the GCOS Precursors for Aerosol and Ozone Essential Climate Variable, including the short-lived atmospheric trace gases: nitrogen dioxide (NO_2), sulphur dioxide (SO_2), carbon monoxide (CO), formaldehyde (HCHO), and ammonia (NH_3).

The project focuses on building consistent/harmonised long-term multi-mission climate data records from satellite instruments including GOME, SCIAMACHY, GOME-2, OMI, TROPOMI, IASI, and MOPITT.

Website: https://climate.esa.int/en/projects/precursors-for-aerosols-and-ozone/

Proposed interview strategy:
(a) send the questions to users by mail,
(b) ask users to complete the answer box in writing, and
(c) make appointment with users to discuss the answers with them afterwards

Interview questions answered by dr. Antje Inness (ECMWF)

1. **Focus on what users need**

(a) *What do you use satellite data of NO_2, HCHO, SO_2, CO, or NH_3 for (mapping, trend analysis, emission monitoring, deposition studies, chemical regime, public outreach, surface estimates)?*

Assimilation in the CAMS global system, in case of the CCI data for use in reanalysis. However, for the current CAMS reanalysis we only used O3 (SCIAMACHY, MIPAS, GOME-2AB), AOD (AATSR), CO2 & CH4 (SCIAMACHY, TANSO) from CCI. We have not used any CCI NO_2, HCHO, SO_2, CO, or NH_3 data

(b) *What are your a priori requirements for using this product? In other words, do you have your own requirements on products, and if yes what are they (particularly when no GCOS requirement is defined (e.g. CHOCHO, NH3))?*

- Good quality
- Good long-term stability.
- Good consistency between data from different sensors
- Continuation of dataset into the future
- Availability close to NRT for reprocessed datasets. This is particularly important once our reanalysis production has caught up with real time and is running close to NRT. Data that are available with a month or less delay will be more useful to us than data that are available with a delay of several months. If data are not available timely enough we might be forced to change to NRT data in the reanalysis production which will affect the quality and consistency of our renalaysis
- Good documentation and validation

(c) *Are you considering the GCOS requirements for your ECV precursor in your decision to use this product? Do you think that the GCOS 'goal' requirements are achievable within the next 5-10 years (see table below, from 2022 GCOS implementation plan)?*

Sometimes. If only one dataset is available for a species there is not much choice. It is important that the datasets come with good validation reports and user guides and that bad data are flagged with clear instructions on how to make use of the data

Table 1. ECV requirements according to the 2022 update to the GCOS requirements (GCOS-200). T = 'threshold' requirement, the minimum to be met so that data are useful, and G = 'goal', an ideal requirement above which further improvements are not necessary. which would result in a significant improvement for satellite data applications, including climate monitoring. From Annex A1 to the 2022 GCOS Implementation Plan

	CO Tropospheric Column	**HCHO Tropospheric Column**	**SO_2 Tropospheric Column**	**NO_2 Tropospheric Column**

Horizontal resolution	10 km (G), 100 km (T)	10 km (G), 100 km (T)	10 km (G), 100 km (T)	10 km (G), 100 km (T)
Vertical resolution	Column	Column	Column	Column
Temporal resolution	hourly (G), 1x per 30 days (T)	hourly (G), 1x per 30 days (T)	hourly (G), 1x per 30 days (T)	hourly (G), 1x per 30 days (T)
Measurement uncertainty (2-sigma)	1 ppb (G), 10 ppb (T)	max(20%, 8E+15 molec. cm-2) (G), max(100%, 40E+15)	max(30%, 6E+15 molec. cm-2) (G), max(100%, 20E+15) (T)	max(20%, 1E+15 molec. cm-2) (G), max(100%, 5E+15) (T)
Stability	<1 ppb/decade (G), 3 ppb/decade (T)	max(4%, 8E+15 molec. cm-2/decade) (G), max(20%, 8E+15 molec. cm-2/decade) (T)	max(6%, 1.2E+15 molec. cm-2/decade) (G), max(20%, 4E+15 molec. cm-2/decade) (T)	max(4%, 1E+15 molec. cm-2/decade) (G), max(20%, 1E+15 molec. cm-2/decade) (T)

(d) Which product do you use, which instrument/period/data product? Why?
For CAMS reanalysis the best overview of what we used until 2016 is Table 2 in Inness et al. (2019) https://acp.copernicus.org/articles/19/3515/2019/
From about 2017 onwards we changed to NRT data for most species because we were running close to NRT.

(e) Why do you think this data product is fit-for-purpose?
Some of the early NRT data we had to use because no other data were available (e.g. SCIA NO2, MIPAS O3) in 2003 and 2004 were not fit for purpose.

(f) Do you know how to obtain and access the satellite data? How do you do that, and do you experience any difficulties with that?
Yes, we know how to obtain and access the data through S5p Exp Hub. However, the interface with the S5p ExpHub is a bit cumbersome to use, in particular, when you want to retrieve data from a particular collection. As now, we have the access from ODA1 at DLR and Eumetcast
for the operational data we are "safe". The main problem is when reprocessing old/test data. Sometimes the test data is limited and we cannot test all the species.

Sometimes it can be difficult to find out where to download datasets from. I think even on the CCI or QA4ECV website this wasn't always clear. A very easy to use entry point website that clearly displays where to get the data. Ideally data access via ftp.

A general overview website (not just CCI, but perhaps as part of GCOS?) with links to all available atmospheric composition satellite datasets would be great. Like a one-stop shop for AC satellite retrievals where the users could browse what is available and get working links to the website of the data producers to download the data from there. Validation and intercomparison of different products should also be available there. Now we have to spend quite some time hunting for datasets when preparing a new reanalysis. It is not always easy to find what is out there and often links are broken. I can see on the GCOS website that some information about satellite data (but if you e.g. go to Networks, Satellite and then go on to Nasa, MLS the link doesn't work. The Eumetsat link also doesn't work.

(g) Is it sufficiently clear to you what the satellite data represent in physical terms (column, vertical sensitivity, spatial resolution, snapshot in time)? Do you consider the representativity of the satellite data in your interpretation?

Yes, we apply the averaging kernels of the data in our observation operator and calculate the model equivalent of the observations at observation location and time.

If there is any additional information needed, e.g. you should average the data (including negative values) to reduce uncertainties and get meaningful results this should be clearly stated in the documentation.

(h) What is your biggest challenge in working with the satellite data (volume, processing costs, interpretation, uncertainties, flagging, meaning, …)?

Sometimes the data volumes are challenging, in particular for reprocessed data (CO) the files tend to be quite large, and we need to adapt sapp (our ECMWF data acquisition and processing system) to process these files.

From the assimilation point of view information about uncertainties and data quality are important. Good flagging of any bad data is crucial.

2. Specific questions to ESA CCI+

(a) Do you prefer to use level-2 or level-3 (daily/weekly/monthly mean gridded) data? Why?

L2 for use in data assimilation

(b) Would you consider using a harmonized, multi-sensor retrieval of ECV precursors? Why (not)?

We usually assimilate the individual sensors, but we want to look into the use of multi-sensor retrievals in the future. Good documentation of such products, including evaluation against independent observations and documentation of the improved quality/benefit compared to individual retrievals will be important. For use in NRT there is always the danger that if one sensor fails the multi-sensor retrieval will not be available anymore and by using individual retrievals we are better protected against the loss of a sensor. This is less of an issue for use in reanalysis

(c) *What is the most important aspect for us to take into account in developing a merged, multi-sensor retrieval of ECV precursors? (retrieval aspects, regridding, accounting for temporal or resolution differences, traceability, uncertainties, reporting on sensor-to-sensor discrepancy, flagging, validation, preparing for future sensors)*

From the user perspective this is validation (including showing the advantage over single sensor), flagging and clear documentation.

3. Specific questions with respect to (S5P and QA4ECV) data format/traceability/QA

Below you see an overview of the L2 data product content currently in place for QA4ECV and S5P NO2.

(a) Are you familiar with the S5P and QA4ECV data format?

I think we didn't download data from qa4ecv. Not sure though if we downloaded some data indirectly from https://d1qb6yzwaaq4he.cloudfront.net/qa4ecv/scia/v1.1/2012/04/scia_no2_qa4ecv_20120407.tar for EAC5. As we haven't processed them yet we cannot say anything about problems with the data or comment on the QA4ECV data format. Anything in HDF, netcdf is usually fine.

(b) Do you assess how the data product is being produced (traceability)?

Yes. Usually by looking at validation reports, ATBD, PUM

(c) Is the information on *__traceability__* *(ATBD, papers, web portal) enabling you to do your own assessment of the fitness-for-purpose of the data product?*

Yes

(d) Does the product have sufficient information on measurement conditions in the form of *__quality flags__* *(on cloud screening, aerosol contamination, sensor issues, algorithm problems, gap-filling, sun-glint, sea ice, etc.)?*

S5P quality flags are good. We have not used any QA4ECV data yet

(e) Does your data product documentation provide you with useful information on how to apply the data, its flags or other *__quality indicators__* *for your applications? As an example, see Table 6 below from the S5P L2 NO_2 Product User Manual, which guides users on how to use the data for different purposes.*

S5P user guides and readme files are good and give the relevant information
Clear instructions on how to use some of the information given in the data would be good.

E.g. when we wrote our initial software to process all the S5P we struggled with the averaging kernels for some data. Unit conversion, scaling factors, and at some point even if we were reading profiles in the wrong way round. A clear formula in the documentation on how to apply e.g. averaging kernels as well as an example plot that shows you a typical shape and magnitude of the values would have helped. Also, see my comment above about clear instructions if data should to be averaged before use.

(f) Which flags are most important to you, and how do they help your application?
S5P qa_flag, a good summary flag telling us whether to assimilate an observation or not
Additional information can also be used if it is known when we process the data. At ECMWF we convert the data from their original format into BUFR which we then use in the IFS. In the BUFR files usually fewer parameters are included than in the original files (e.g. for NO2 we have field of view, solar elevation, surface type, cloud cover, cloud top height). Any data issues that depend on those parameters we can handle in our data assimilation system (e.g. don't use data at low solar elevation). If there are data issues depending on other parameters that we don't have in the bufr files we can only flag those data at the time of processing by setting a quality flag that is in the BUFR files to bad or suspect.

(g) Based on the available papers, PSD, PUM, etc. do you understand how to evaluate the uncertainty of the data product for your application? Why (not)? What is missing? How do you best use uncertainty information?
We usually use the error values provided in the data.

(h) What do you think is the best way for the ESA CCI+ products to be developed to deal with aspects of data format, traceability, QA and uncertainties? What would help you most?
Good documentation and validation. It would also be helpful to have cross-validation with other data products (e.g. retrievals from same instrument from different producers). E.g. the ACSAF has similar data products to CCI, but often based on different retrievals and as a user we would like to know which one is best for our application and which one we should e.g. use in the next CAMS reanalysis without having to test both datasets ourselves. This should include information about biases, long term stability, consistency with other sensors, continuation of the dataset into the future, timeliness of dataset.
(i) What information do you need in the L3 product (spatial and temporal resolution, number of observations, uncertainty, others,…) ?
Not needed

4. Specific question on validation
(a) Has your favorite product been validated? Is it clear to you who did that?
Not always clear.

(b) How do you interpret the result of the validation efforts? Do they impact the way you use the satellite data?

Yes, for example we might not use data at high latitudes if they are shown to have problems there, or over icy surface,…

(c) Do you consider the reported satellite product uncertainties when evaluating the validation results? How?

This should be part of the validation report, e.g. error bars.

(d) Do you link the validation results to the GCOS requirements?

No

(e) What is your general opinion on the validity of the data product?

S5P validation good