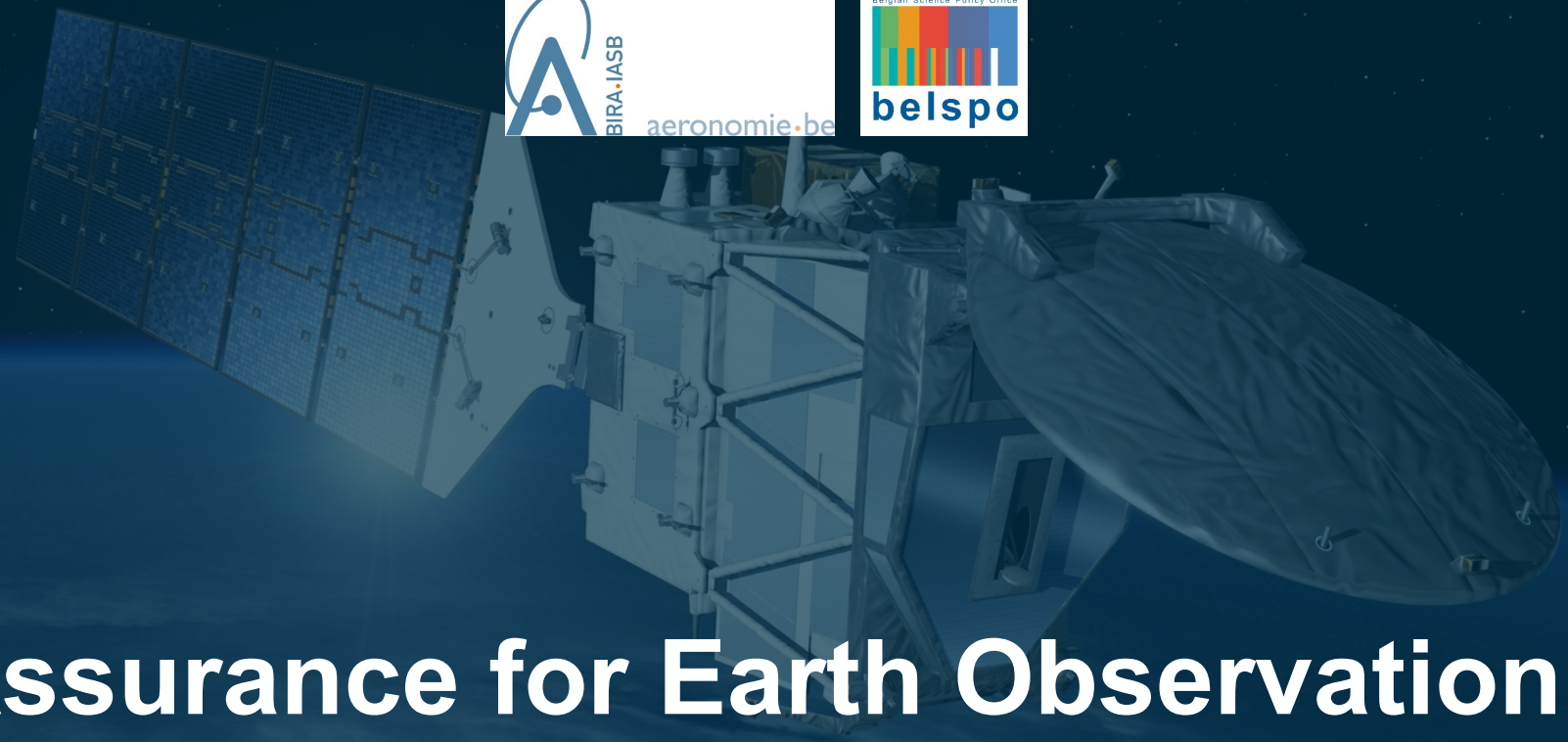


Quality Assurance for Earth Observation (QA4EO) concept and example

J.-C. Lambert, S. Compernolle, D. Hubert, A. Keppens, and T. Verhoelst

Royal Belgian Institute for Space Aeronomy (BIRA-IASB)

2nd ESA EarthCARE Validation Workshop

25-28 May 2021 (online)

Quality Assurance for Earth Observation (QA4EO) concept and example

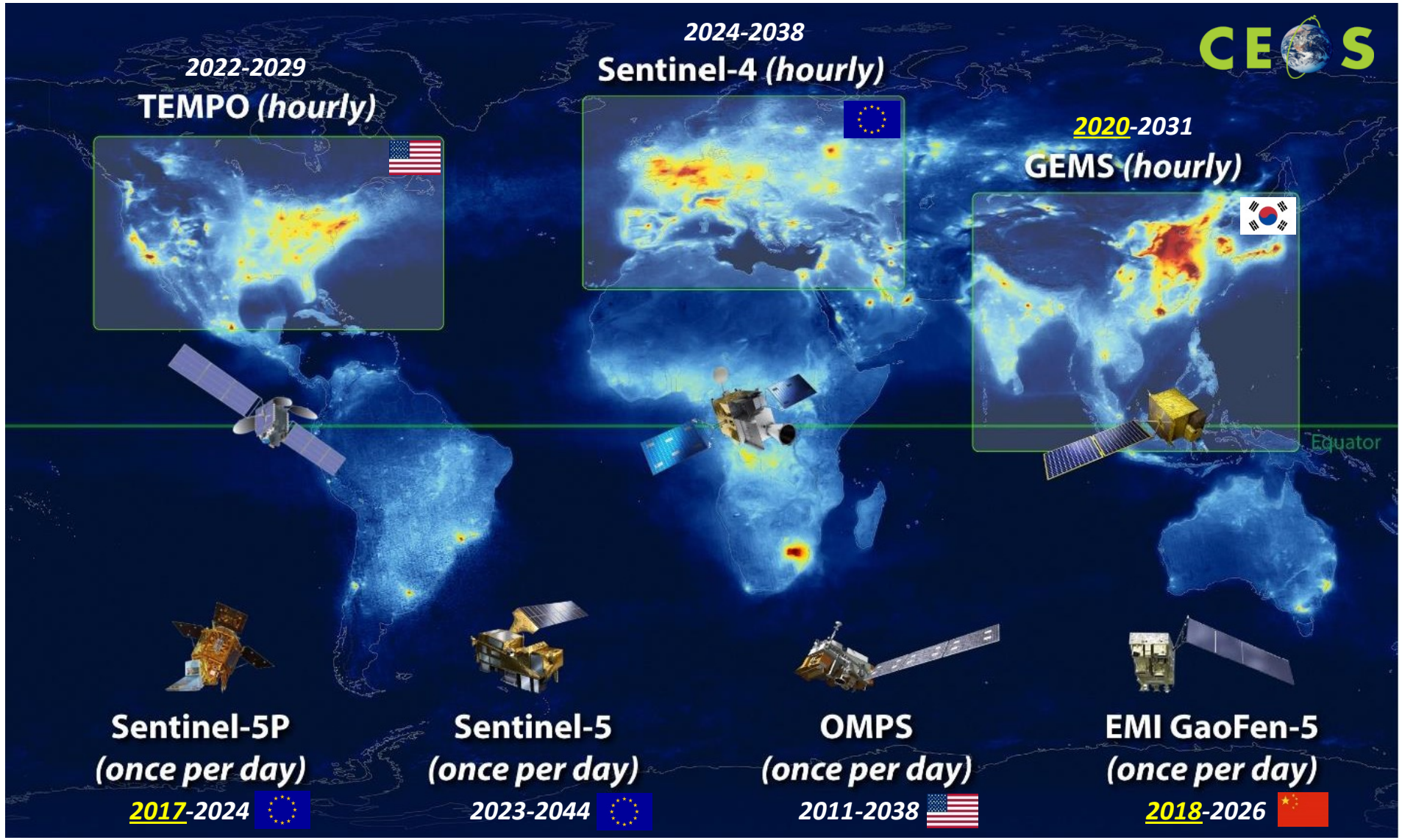
1. QA4EO framework

2. QA4EO implementation methods and tools

3. Example 1: QA4ECV QA system applied to multi-mission NO_2 Climate Data Record

4. Example 2: Validation of Sentinel-5p NO_2 intermediate retrieval parameters

LEO+GEO Satellite Constellation for Air Quality



QA4EO

A QUALITY ASSURANCE FRAMEWORK FOR EARTH OBSERVATION

Quality Indicator

traceability

users

evidence

QA4EO Principle

Data and derived products shall have associated with them a fully traceable indicator of their quality

Quality Indicator

A Quality Indicator (QI) shall provide sufficient information to allow all users to readily evaluate the "fitness for purpose" of the data or derived product

Traceability

A QI shall be based on a documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards

fitness-for-purpose

documentation

reference standards

BIRA•IASB

QA4EO

A QUALITY ASSURANCE FRAMEWORK FOR EARTH OBSERVATION

Quality Indicator

traceability

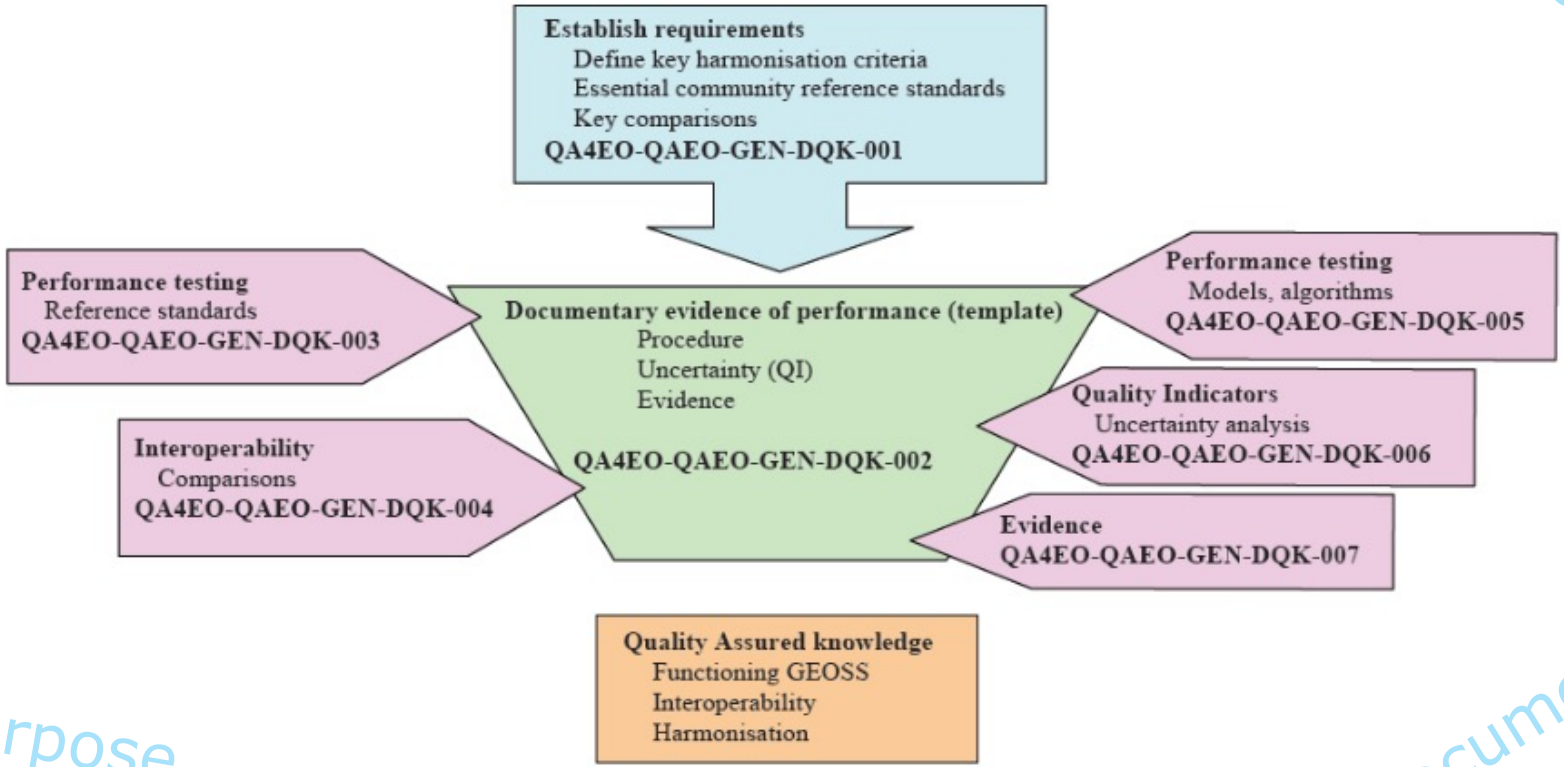
users

evidence

fitness-for-purpose

reference standards

documentation



CEOS WGISS Data Management & Stewardship Maturity Matrix

Usability criteria: encoding, documentation, traceability, validation, data uncertainty, QA/QC

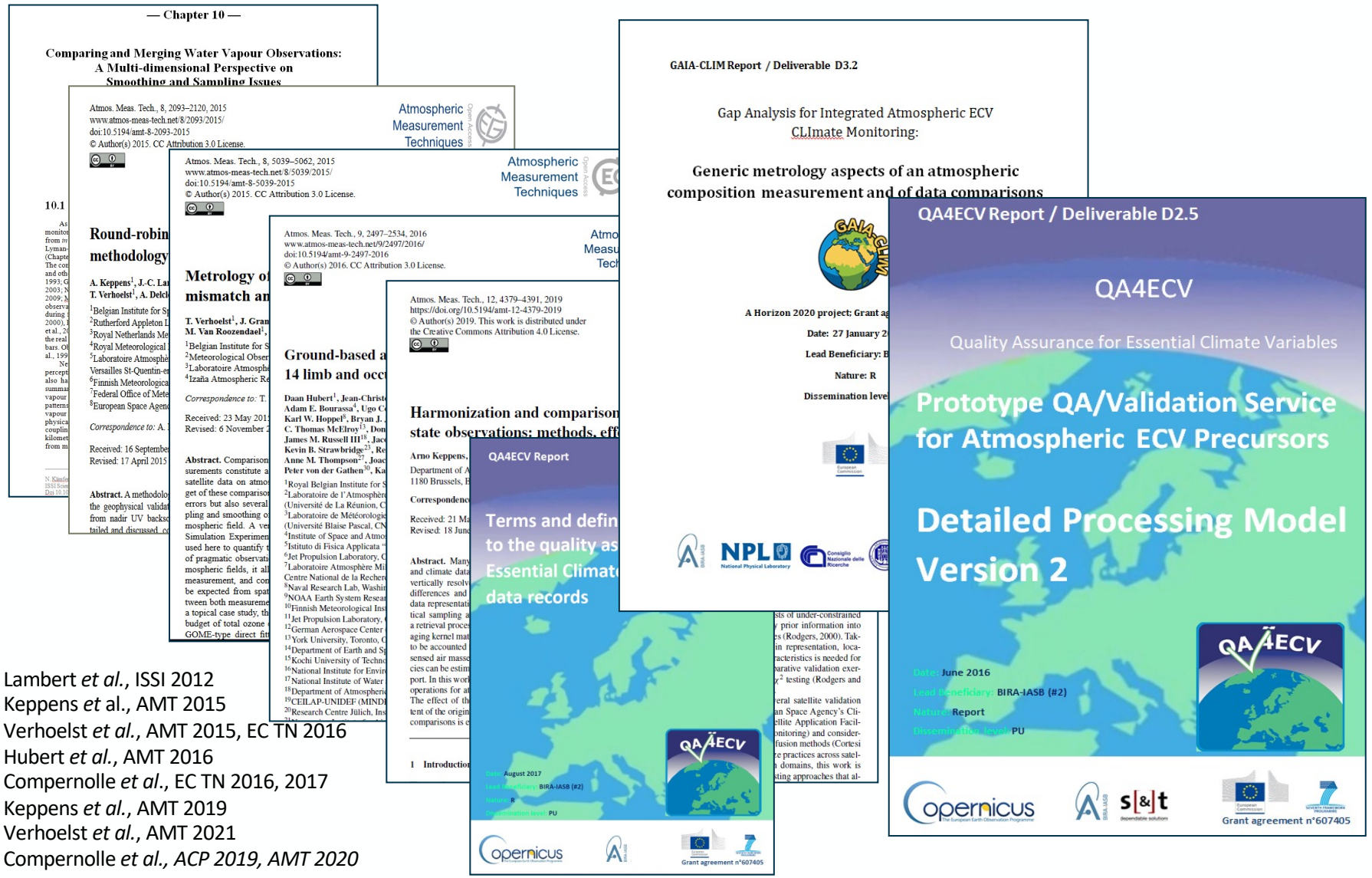
not managed

fully managed

	DISCOVERABILITY	ACCESSIBILITY	USABILITY						PRESERVATION		CURATION	
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceability	MMP6 Data Validation	MMP7 Data Uncertainty	MMP8 Data Quality Control	MMP9 Data Preservation	MMP10 Data Verification	MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable Identifier
Level-0 Not Managed	1) No catalogue available 2) No advertising available	Data and metadata are not accessible online	1) Data Not Structured 2) Non-standard or proprietary data format, or, poorly-documented standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	1) Reference Data Representativeness - No validation 2) Reference Data Quality - No validation 3) Validation Method - No validation 4) Validation Results - No validation	1) Uncertainty Method: Uncertainty characterisation not performed, or method not documented. 2) Uncertainty Sources: Uncertainty characterisation not performed, or sources analysed not documented. 3) Uncertainty Values: No uncertainty information provided.	1) No control and monitoring check 2) No quality indicator in metadata 3) No procedures documentation	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check	1) No reprocessing activities planned 2) Pre-flight calibration & characterisation not documented or information not available. 3) Post-launch calibration & characterisation not documented or not available. 4) Processing: Additional processing steps not documented.	No persistent and resolvable identifiers available
Level-1 Partially Managed	1) Advertising available 2) Catalogue search available at product level	Basic online services available for data and metadata access	1) Basic schema for automated data use 2) Data in documented standard file format. Non-standard naming conventions used.	1) Already existent mission documentation available and preserved for the long term 2) No link between mission documentation and data records	Product information available (not online)	1) Reference Data Representativeness: measurements assessed to be mostly representative of the satellite measurements 2) Reference Data Quality: single uncertainty for the entire dataset. 3) Validation Method: simple uncertainty estimated 4) Validation Results: Validation results show good agreement between satellite and reference measurements within uncertainties in most cases.	1) Uncertainty Method: Limited use of GUM approach, and/or, an expanded comparison to measurements by other sensors. 2) Uncertainty Sources: Most important sources of uncertainty included. 3) Uncertainty Values: Single uncertainty value provided for subsets of data	1) Basic data quality control and monitoring check 2) Minimal set of quality control procedures documented and available	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	1) Minor updates and bugs corrections of data records implemented 2) Data Records repackaging and/or reformatting 3) Pre-flight calibration & characterisation misses some important aspects 4) Post-launch calibration & characterisation misses some important aspects of instrument behaviour and/or is not entirely of a level of quality to be judged fit for purpose. 5) Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.	1) Persistent identifier assignment only for particular Data Records Collections 2) Basic landing pages management
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product metadata oriented towards an international standard 3) Data Collection and Associated Information searchable. 4) International standard for Collection metadata	1) Simple Access Architecture through metadata 2) Data access system oriented towards an international standard	1) Use of non-proprietary international standards encodings for syntactic interoperability. 2) Periodically repackaging/ reformatting of archived data. 3) Data in well-documented standard file format, community naming convention standards.	1) Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	Dataset tested for presence of correct provenance metadata. Well described product information available online	1) Reference Data Representativeness: measurements assessed to be well representative of the satellite measurements 2) Reference Data Quality: full uncertainty information 3) Validation Methods assess satellite measurements 4) Validation Results show excellent agreement between satellite and reference measurements, within uncertainties.	1) Uncertainty Method: GUM approach to estimate measurement uncertainty with full breakdown of components and separated as Type A or B classification. 2) Uncertainty Sources: All important sources of uncertainty included. 3) Uncertainty Values: Total uncertainty per pixel is provided, with basic breakdown of key components no error-covariance.	1) Quality indicator post-processing available 2) Quality control procedures documented and available online	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Reprocessing for calibration and/or algorithm improvement 2) Pre-flight calibration & characterisation covers all reasonable aspects 3) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance and uses appropriate community infrastructure/methods (CEOS/FRMs). 4) Additional processing steps documented.	1) Persistent identifier assignment to all disseminated Data Records Collections and metadata 2) Automatic landing page generation and extensive management of landing pages
Level-3 Fully Managed	1) International standard for Product metadata 2) International standard for Collection metadata 3) Catalogue accessible via international or community agreed standards protocol 4) Data policy available in metadata 5) Periodic updates of metadata in the catalogue 6) Quality indicator metadata available and discoverable 7) Search results relevancy. 8) Seamless transition from discovery to access	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 4) Reporting system 5) Hosted processing 6) Quick adoption to new technologies and standards evolution 7) Data and metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadata uses FAIR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	1) Automatic metadata generation for provenance documentation 2) Complete and updated data provenance available online	1) Reference Data Representativeness: Reference measurements independently assessed to be fully representative of the satellite measurements, covering the satellite's full range of measurements and with full assessment of uncertainties and carried out on a regular basis determined by product performance. 2) Reference Data Quality: full uncertainty and error-correlation information, assessed following the GUM and traceable to SI 3) Validation Methods assess satellite measurements and reference data w.r.t. their error- covariance and validates those uncertainties. 4) Validation Results show excellent agreement between satellite and reference measurements, within uncertainties.	1) Uncertainty Method: GUM approach to estimate measurement uncertainty, including a treatment of error-covariance. 2) Uncertainty Sources: All reasonable sources of uncertainty included. 3) Uncertainty Values: Uncertainties per pixel provided with error-covariance information for all appropriate components.	1) Data quality control fully compliant with an international standard 2) Quality indicator pre and post processing available in the metadata 3) Quality metadata assessed	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting	1) Reprocessing for time-series creation 2) Roadmap for technology evolution 3) Plurality of accurate and relevant attributes are provided to allow reuse 4) Metadata includes information about the licence 5) Pre-Flight: As Level-2, additionally calibration and characterisation includes the measurements needed to assess uncertainties at component level and their impact on the final product. 6) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance. 7) All additional processing steps fully documented and state-of-the-art.	1) Persistent identifier created for all accessible data records and metadata 2) Metadata includes the identifier for the data 3) Metadata is offered in such a way that it can be harvested and indexed

https://ceos.org

Generic validation chain, state-of-the-art co-locators, harmonized terminology, common metrics…



Lambert *et al.*, ISSI 2012
Keppens *et* al., AMT 2015
Verhoelst *et al.*, AMT 2015, EC TN 2016
Hubert *et al.*, AMT 2016
Compernolle *et al.*, EC TN 2016, 2017
Keppens *et al.*, AMT 2019
Verhoelst *et al.*, AMT 2021
Compernolle *et al., ACP 2019, AMT 2020*

GEO CEOS WGCV-28/…/31

Cross GEOSS EO Cal/Val harmonization initiatives

esa cci ozone

Prototyping in Multi-TASTE and ESA CCI_Ozone

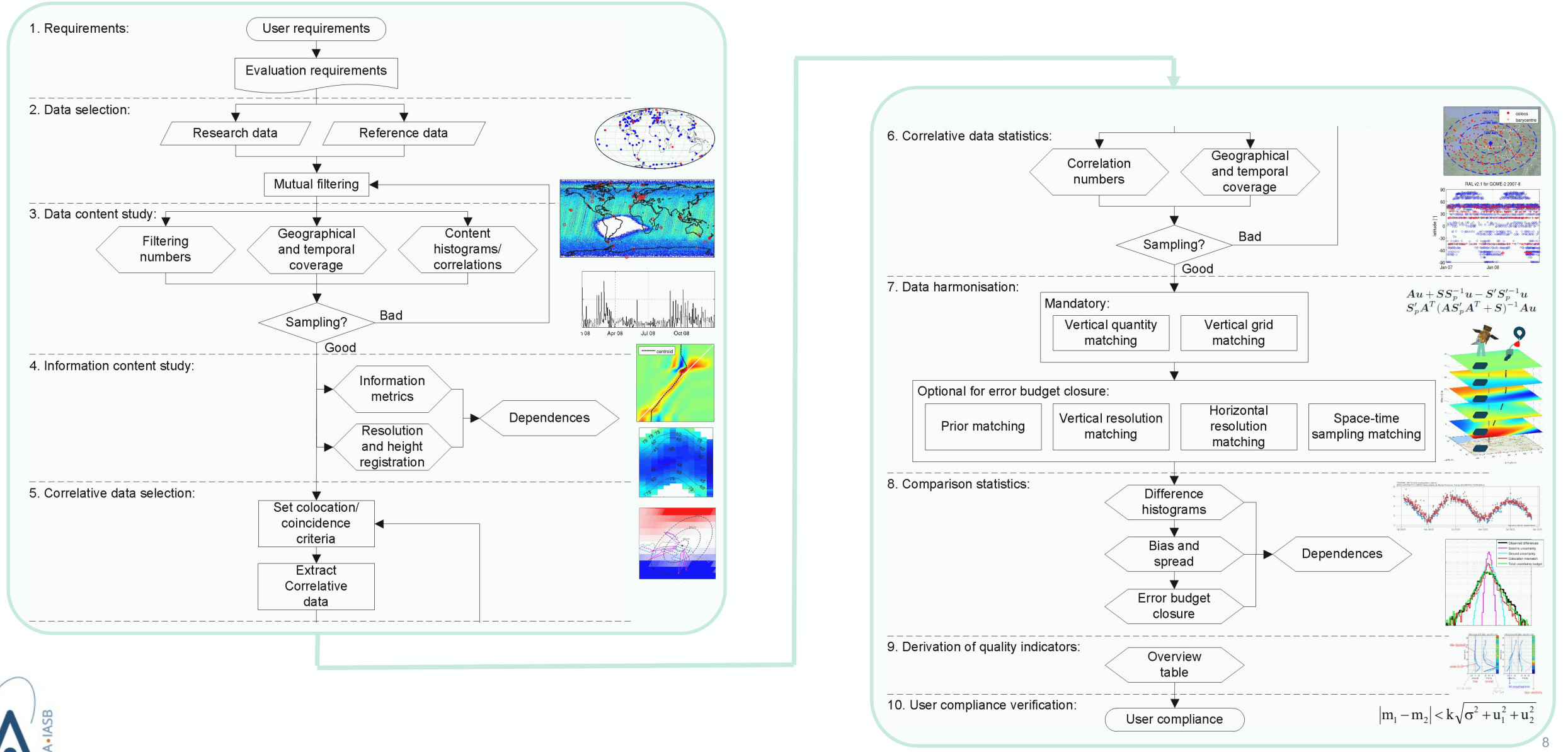
CEOS ACC-8/…/11

Community feedback & endorsement

TROPOMI
Climate Change Service

Implementation in FP7 QA4ECV, H2020 GAIA-CLIM, C3S Ozone, HARP, Sentinel-5p MPC, CAMS-NDACC, CCI+, CCVS…

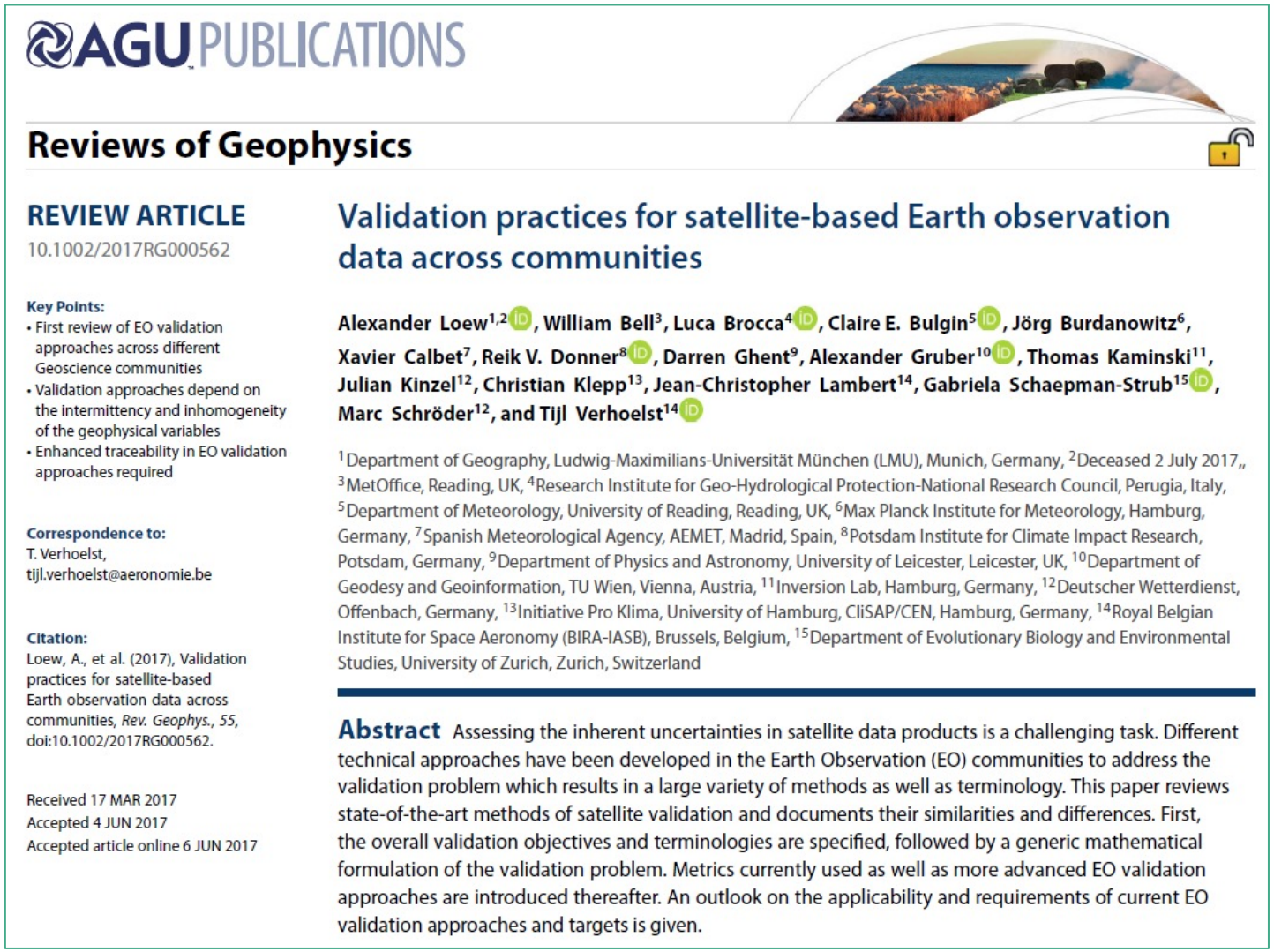
Generic round-robin validation protocol → set of Quality Indicators



Keppens *et al.*, AMT 2015

Validation practices across EO domains

- Objectives of Cal/Val
- Terminology
- Mathematical formulation
- Validation metrics
- Advanced methods and strategies



AGU PUBLICATIONS

Reviews of Geophysics

REVIEW ARTICLE
10.1002/2017RG000562

Key Points:
- First review of EO validation approaches across different Geoscience communities
- Validation approaches depend on the intermittency and inhomogeneity of the geophysical variables
- Enhanced traceability in EO validation approaches required

Correspondence to:
T. Verhoelst,
tijl.verhoelst@aeronomie.be

Citation:
Loew, A., et al. (2017), Validation practices for satellite-based Earth observation data across communities, *Rev. Geophys.*, *55*, doi:10.1002/2017RG000562.

Received 17 MAR 2017
Accepted 4 JUN 2017
Accepted article online 6 JUN 2017

Validation practices for satellite-based Earth observation data across communities

Alexander Loew[1,2], William Bell[3], Luca Brocca[4], Claire E. Bulgin[5], Jörg Burdanowitz[6], Xavier Calbet[7], Reik V. Donner[8], Darren Ghent[9], Alexander Gruber[10], Thomas Kaminski[11], Julian Kinzel[12], Christian Klepp[13], Jean-Christopher Lambert[14], Gabriela Schaepman-Strub[15], Marc Schröder[12], and Tijl Verhoelst[14]

[1]Department of Geography, Ludwig-Maximilians-Universität München (LMU), Munich, Germany, [2]Deceased 2 July 2017,, [3]MetOffice, Reading, UK, [4]Research Institute for Geo-Hydrological Protection-National Research Council, Perugia, Italy, [5]Department of Meteorology, University of Reading, Reading, UK, [6]Max Planck Institute for Meteorology, Hamburg, Germany, [7]Spanish Meteorological Agency, AEMET, Madrid, Spain, [8]Potsdam Institute for Climate Impact Research, Potsdam, Germany, [9]Department of Physics and Astronomy, University of Leicester, Leicester, UK, [10]Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria, [11]Inversion Lab, Hamburg, Germany, [12]Deutscher Wetterdienst, Offenbach, Germany, [13]Initiative Pro Klima, University of Hamburg, CliSAP/CEN, Hamburg, Germany, [14]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, [15]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

Abstract Assessing the inherent uncertainties in satellite data products is a challenging task. Different technical approaches have been developed in the Earth Observation (EO) communities to address the validation problem which results in a large variety of methods as well as terminology. This paper reviews state-of-the-art methods of satellite validation and documents their similarities and differences. First, the overall validation objectives and terminologies are specified, followed by a generic mathematical formulation of the validation problem. Metrics currently used as well as more advanced EO validation approaches are introduced thereafter. An outlook on the applicability and requirements of current EO validation approaches and targets is given.

Loew *et al., Rev. Geophys.* 2017

Quality Assurance system for EO ECVs

FP7-SPACE-2013-1
Project No 607405

- FP7 project 2014-2018
- QA4EO implementation for ECVs
- Applied to 6 pilot ECVs:

  Land
  - surface spectral albedo
  - FAPAR
  - LAI

  Atmosphere
  - NO_2
  - HCHO
  - CO

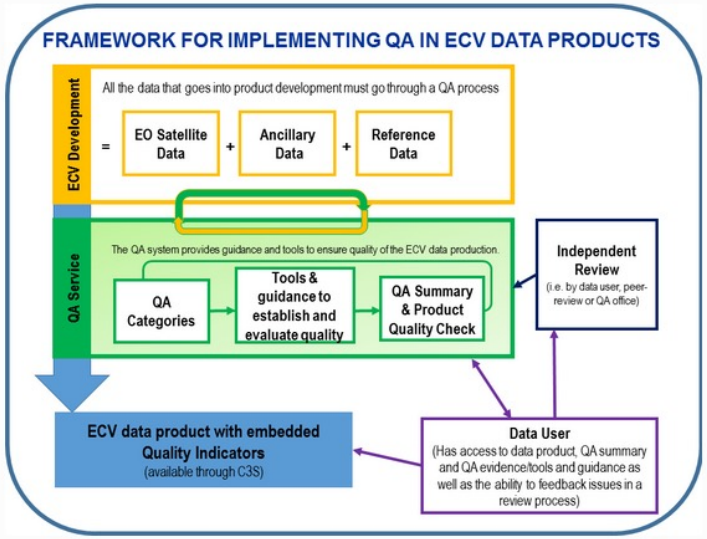
QA4ECV QA System

http://www.qa4ecv.eu

The QA4ECV Quality Assurance (QA) System is now available for use by data providers. The function of the QA4ECV QA service is to:

- Provide ECV producers with resources to generate ECV products with embedded QA information
- Provide data users with QA information to assess the fitness-for-purpose of data for their applications

Quality System | Training



Design of the QA System

The QA system has been developed through consideration of current practices. An overview of the system and support studies are available:

- Overview of the QA system
- Review of Other Projects
- Review of Quality Flags in Products
- Review of Validation in Products

Guidance for Users

Guidance for various aspects are available:

- Algorithm Theoretical Basis: Guidance and Template
- Product User Manual: Guidance and Template
- Traceability Chain: Guidance and Tool Manual
- Validation Report: Guidance and Template

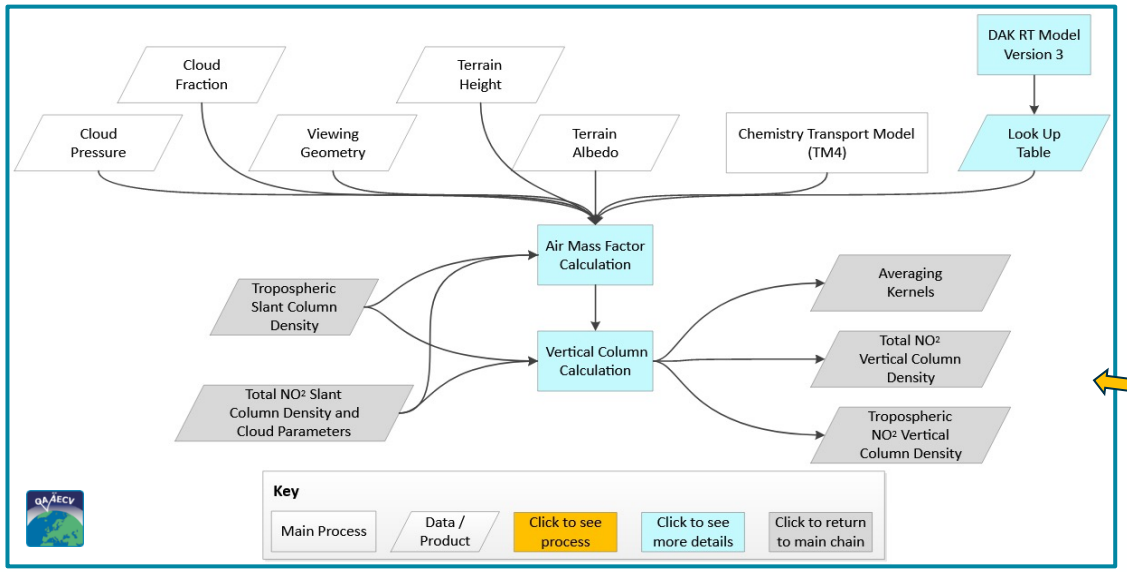
Checking Process

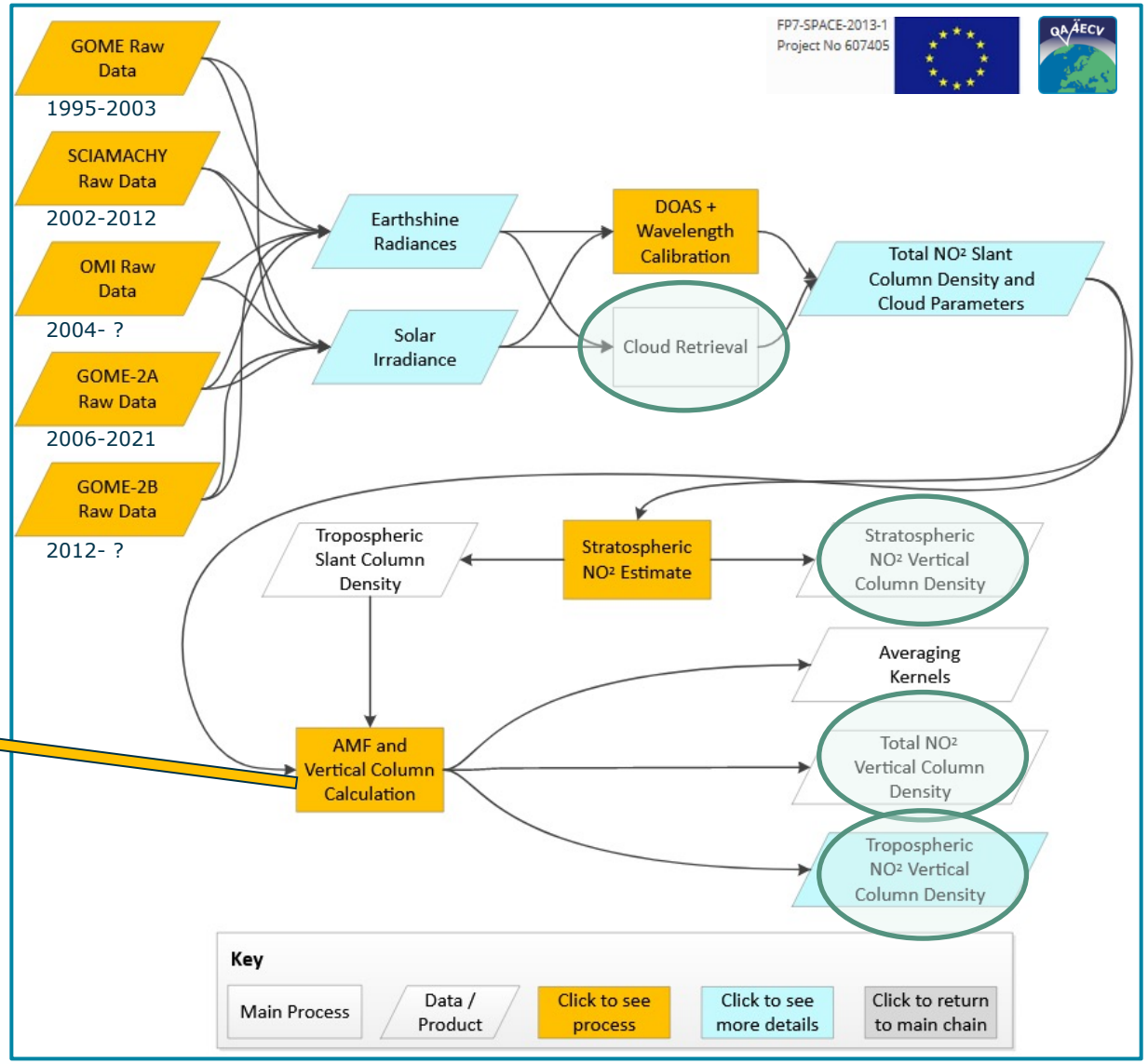
Once information is submitted to the QA system, it will be subject to checking to ensure consistency of the information provided with other products. The following describe this process:

- Quality Checklist
- Quality Checking Procedure

QA4ECV QA system for EO ECVs: atmospheric NO_2

First cross-calibrated, multi-sensor NO_2 Climate Data Record, spanning 20+ years, with very detailed quality information embedded.





Boersma *et al.*, AMT 2018

QA4ECV QA system for EO ECVs: atmospheric NO_2

NO_2

The QA4ECV NO_2 ECV precursor product has achieved advanced status for the information provided for product details, traceability, quality flags, and assessment against standards; intermediate status for uncertainty assessment; and basic status for validation.

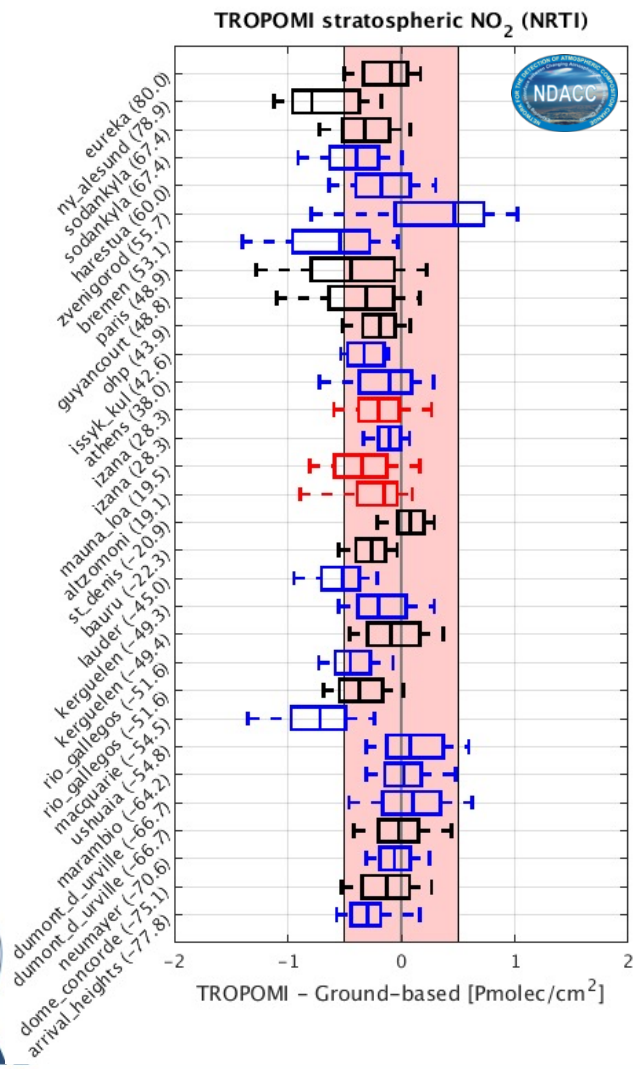
Further work (inspired among others by outcome of QA4ECV audits):

- Improved retrieval method
- L1 calibration and L2 retrieval
- Tailoring of ground-based validation measurements (FRM)
- Validation of NO_2 intermediate and influence quantities
- Uncertainty assessment, representativeness

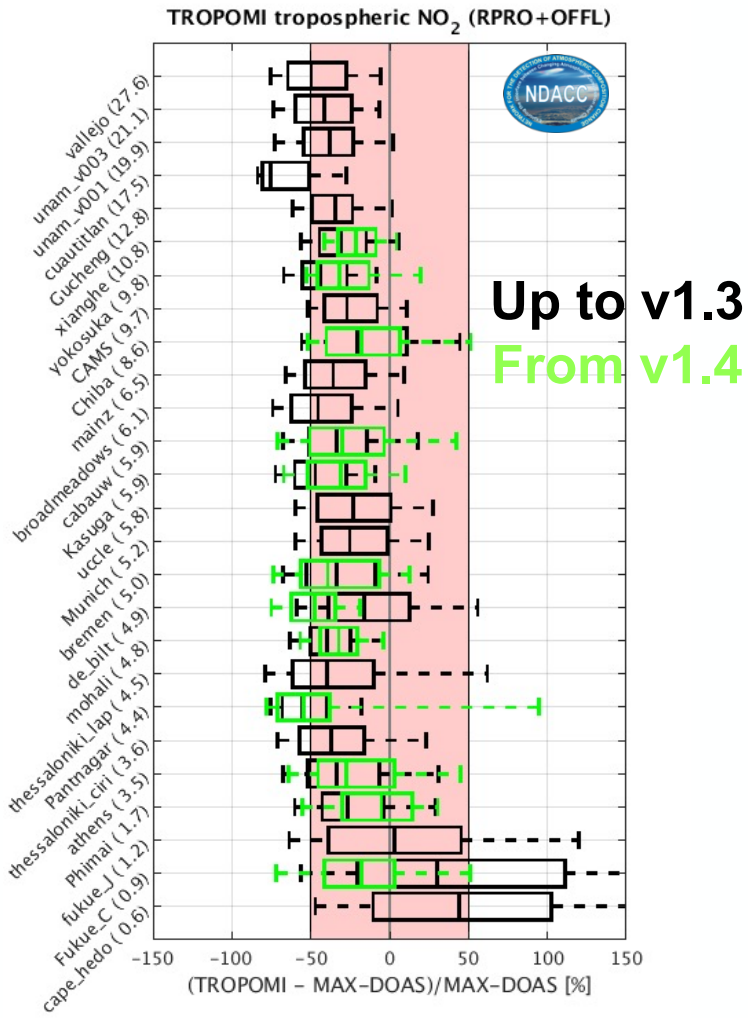
Advances in NO_2 validation

Multi-network validation of Sentinel-5p TROPOMI NO_2 intermediate quantities

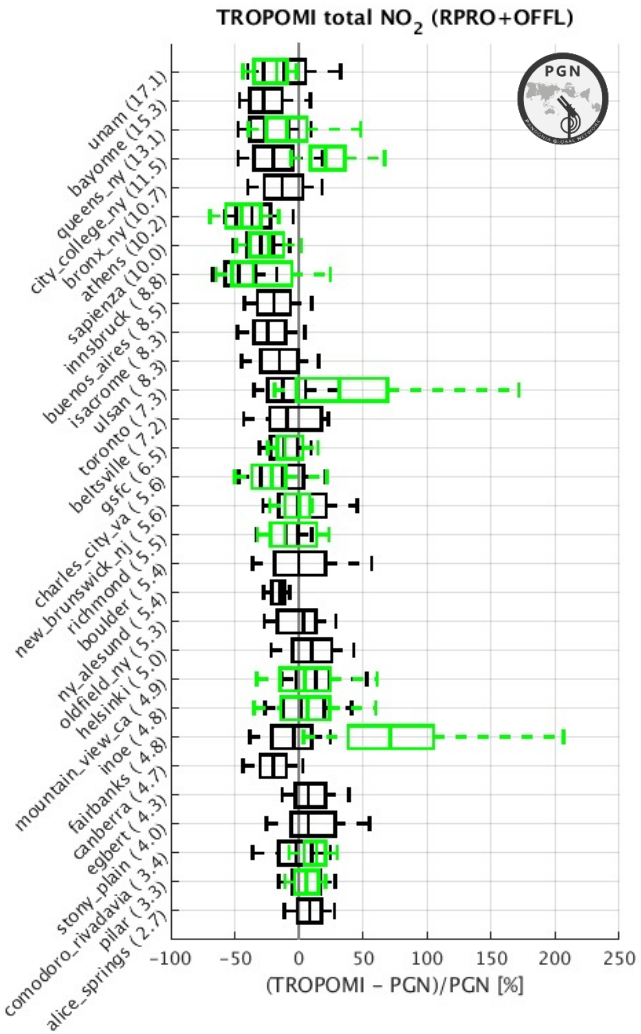
Stratospheric: vs. zenith-sky DOAS



Tropospheric: vs. multi-axis DOAS



Total: vs. direct sun DOAS



Verhoelst *et al.*, AMT 2021

Advances in NO_2 validation

Validation of Sentinel-5p TROPOMI cloud properties as influence quantities for trace gas retrievals

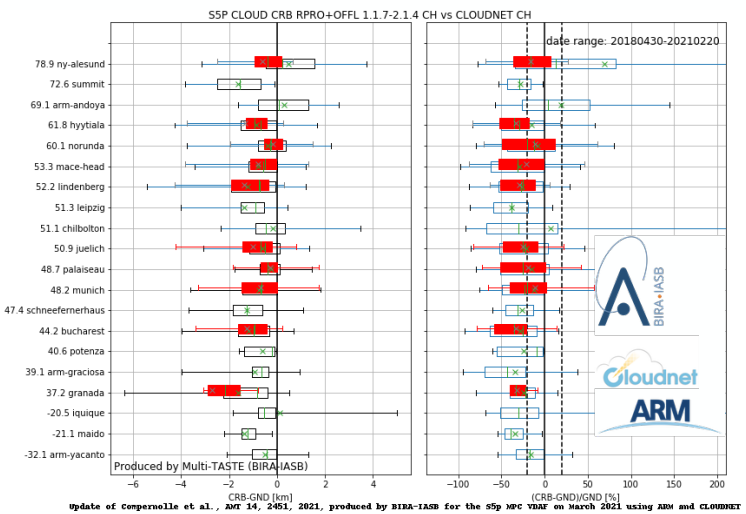
https://doi.org/10.5194/amt-2020-122
Preprint. Discussion started: 29 June 2020

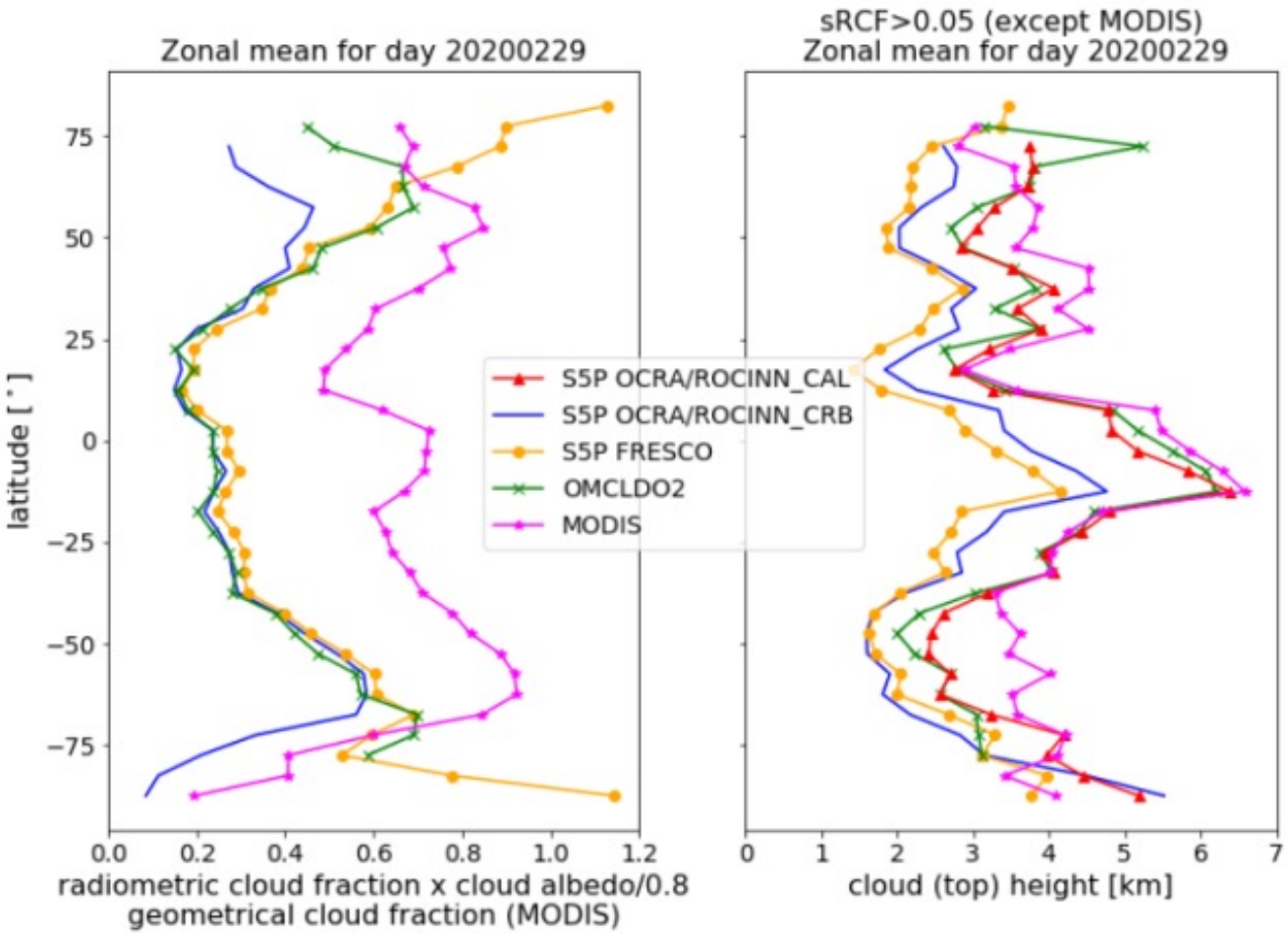
Atmospheric Measurement Techniques Discussions

Validation of the Sentinel-5 Precursor TROPOMI cloud data with Cloudnet, Aura OMI O_2-O_2, MODIS and Suomi-NPP VIIRS

Steven Compernolle[1], Athina Argyrouli[2,3], Ronny Lutz[3], Maarten Sneep[4], Jean-Christopher Lambert[1], Ann Mari Fjæraa[5], Daan Hubert[1], Arno Keppens[1], Diego Loyola[3], Ewan O'Connor[6,7], Fabian Romahn[3], Piet Stammes[4], Tijl Verhoelst[1], and Ping Wang[4]

[1]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Ringlaan 3, 1180 Uccle (Brussels), Belgium
[2]Technical University of Munich, TUM Department of Civil, Geo and Environmental Engineering, Chair of Remote Sensing Technology, Munich, Germany
[3]German Aerospace Center (DLR), Münchener Straße 20, 82234 Weßling, Germany





Compernolle *et al.*, AMT 2020

Conclusion

- QA4EO provides a set of guidelines facilitating the data quality assurance of EO missions, enhancing their value, completeness, interoperability, and heritage.
- Application and value of QA4EO demonstrated for several atmospheric EO data products and several cross-domain EO based services
- Maturity of QA4EO implementation depends on geophysical variable, data retrieval scheme, community, user needs…
- Highlights further research needs: expression and calculation of data uncertainties, representativeness issues, interoperability requirements, new validation approaches…