Harmonization of uncertainty within QA4ECV

and

Feedback on Fiduceo

Steven Compernolle[1], Jean-Christopher Lambert[1],
Sander Niemeijer[2], Tijl Verhoelst[1]

1 BIRA-IASB aeronomie.be

2 s[&]t

FP7/H2020 context

Production

Metrological traceability and interoperability of data

Propagating uncertainty signal to end product

Mismatch uncertainty satellite – ground linking satellite data to SI

Quality-assurance of climate data, including **metrological** traceability

Users

Climate data dissemination to users, including quality metrics (e.g., **uncertainty**)

Applications

Attribution **risk** extreme weather events to climate change

Impact 2C warming, integrating **uncertainties**

Impact&risks >2C warming

Semi-operational **impact prediction** systems

Reanalysis and climate prediction

Reanalysis global climate data **+uncertainty**

Reanalysis regional climate data **+uncertainty**

Climate **forecast**

Interoperability: tools development

Satellite data
GOME, GOME-2, OMI, TROPOMI,…

Ground-based data
UV-Vis, FTIR, lidar, balloon…

Model & reanalysis data
CAMS, ERA-Interim,…

O_3, CH_4, CO, NO_2, temperature, aerosol,…
Many file formats

harpconvert
harpfilter

harpconvert
harpfilter

harpconvert
harpfilter

Format harmonization
Content harmonization
Variable selection
Filtering…

harpcollocate
harpfilter

Co-locating data pairs
Regridding & smoothing

HARP

In development

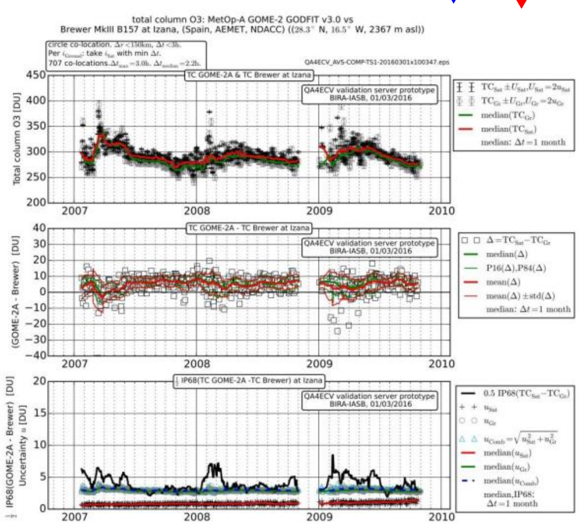


Figure 20. Plots for the comparison of GOME-2A (GODFIT v3.0) total column O3 and ground-based

Comparison
Plotting
Reporting

Metrological traceability:
Uncertainties are transformed along with the main data

Interoperability of data: tools development

BIRA-IASB

S. Compernolle
J. Granville
A. Keppens
J.-C. Lambert
B. Langerock
T. Verhoelst,…

O3M SAF
AURORA
CAMS
Copernicus – Europe's eyes on Earth
QA4ECV
GAIA-CLIM

User feedback
Sharing expertise

Including about metrology/uncertainty

S. Niemeijer
B. Rino
…

HARP tools

s[&]t dependable solutions

Open source
Dissemination

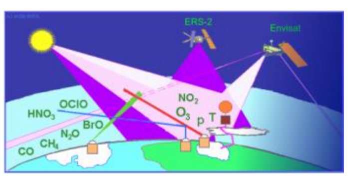
Public

GECA * 2007-2011

macc 2009-2015

* Atmosphere Monitoring Service 2015-…

GOMEval, CINAMON, TASTE, Multi-TASTE…



Since 1995…

NORS * Network of Remote Sensing Ground-Based Observations for the GMES Atmospheric Service
2011-2014

QA4ECV * 2014-2017

GAIA-CLIM 2015-2018

TROPOMI * MPC-VDAF
2016-…

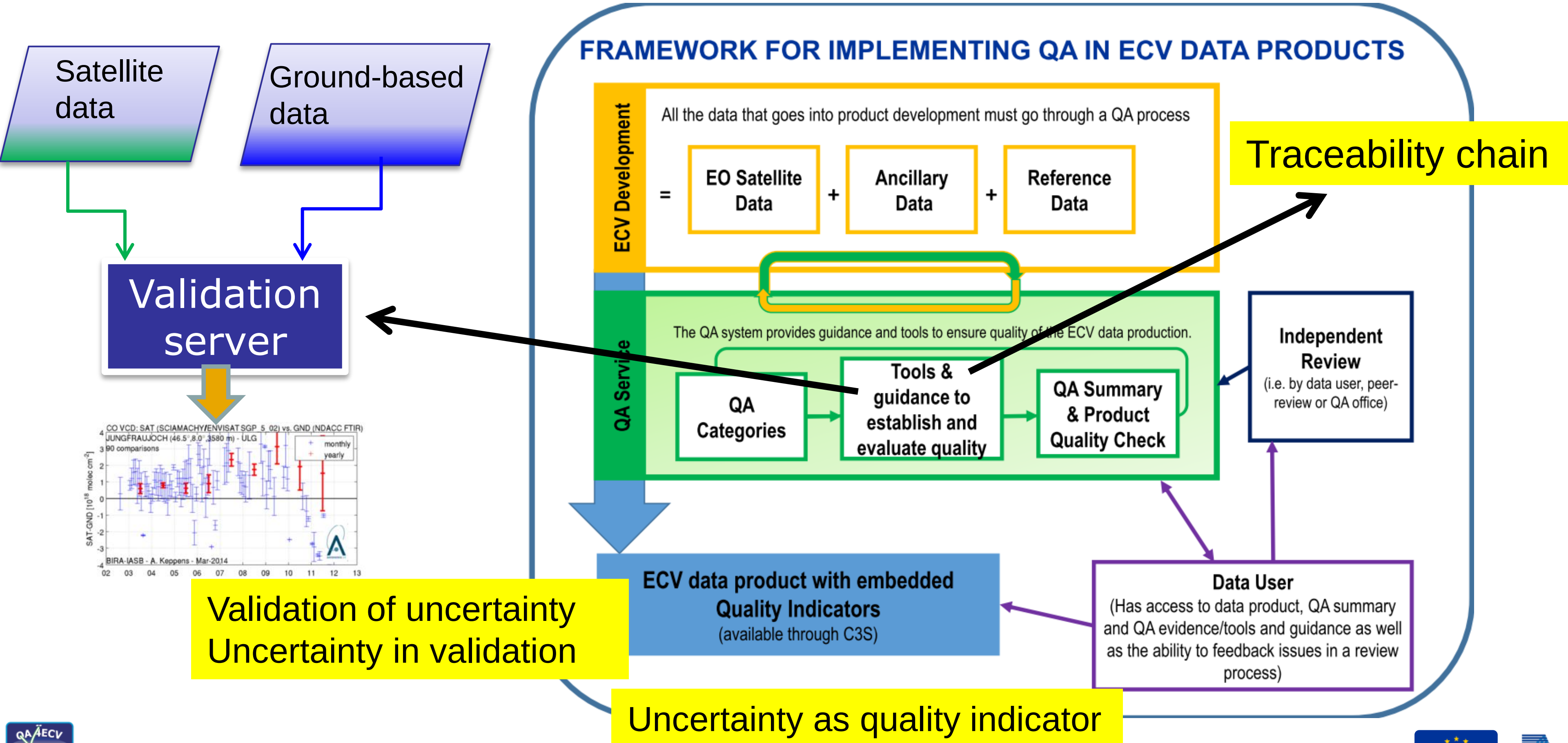
QA4ECV
esa
SEVENTH FRAMEWORK PROGRAMME
Copernicus – Europe's eyes on Earth

*S[&]T project partner

SEVENTH FRAMEWORK PROGRAMME

FP7 QA4ECV project

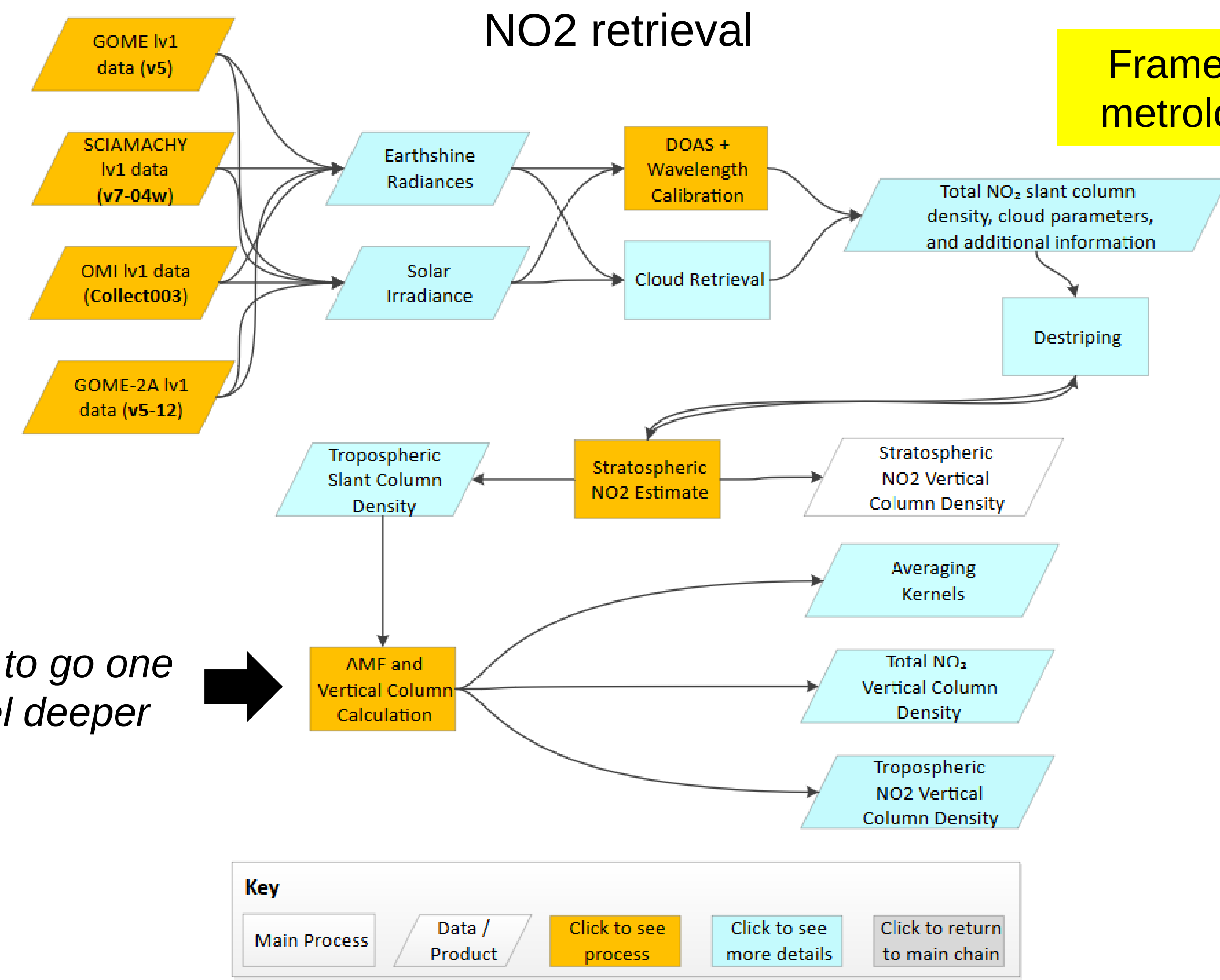
Objective: Develop pilot Quality Assurance framework for Essential Climate Variables
→ Producer can assure the quality of his/her data product
→ User can evaluate the fitness-for-purpose of the data product



Traceability chains/diagrams

http://www.qa4ecv.eu/ecvs Chains for HCHO, NO2, CO, Albedo, LAI, FAPAR

Framework to facilitate metrological traceability



Click to go one level deeper

Uncertainty as a Quality Indicator

Roles of uncertainty

Pixel-level

- Filter out bad observations
- Building derived products
  E.g., weighted average
- Input for climate model studies
- Evaluating comparisons, validations
- …

Dataset level

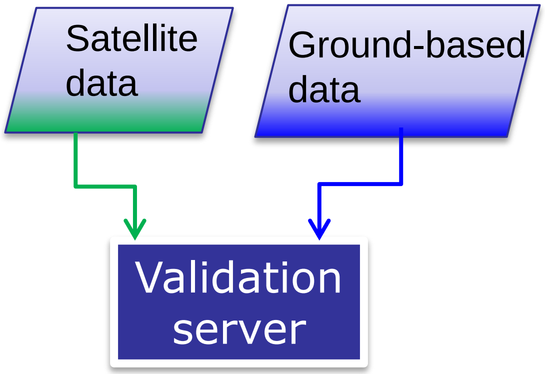
- data discovery
- quick verification of fitness-for-purpose of data product
- …

GEOSS/Copernicus perspective: QA4EO principles

Enabling harmonization and interoperability of data products

All Earth Observation (EO) data and derived products has associated with it a documented and fully traceable Quality Indicator (QI) allowing all users to evaluate the fitness-for-purpose of the data

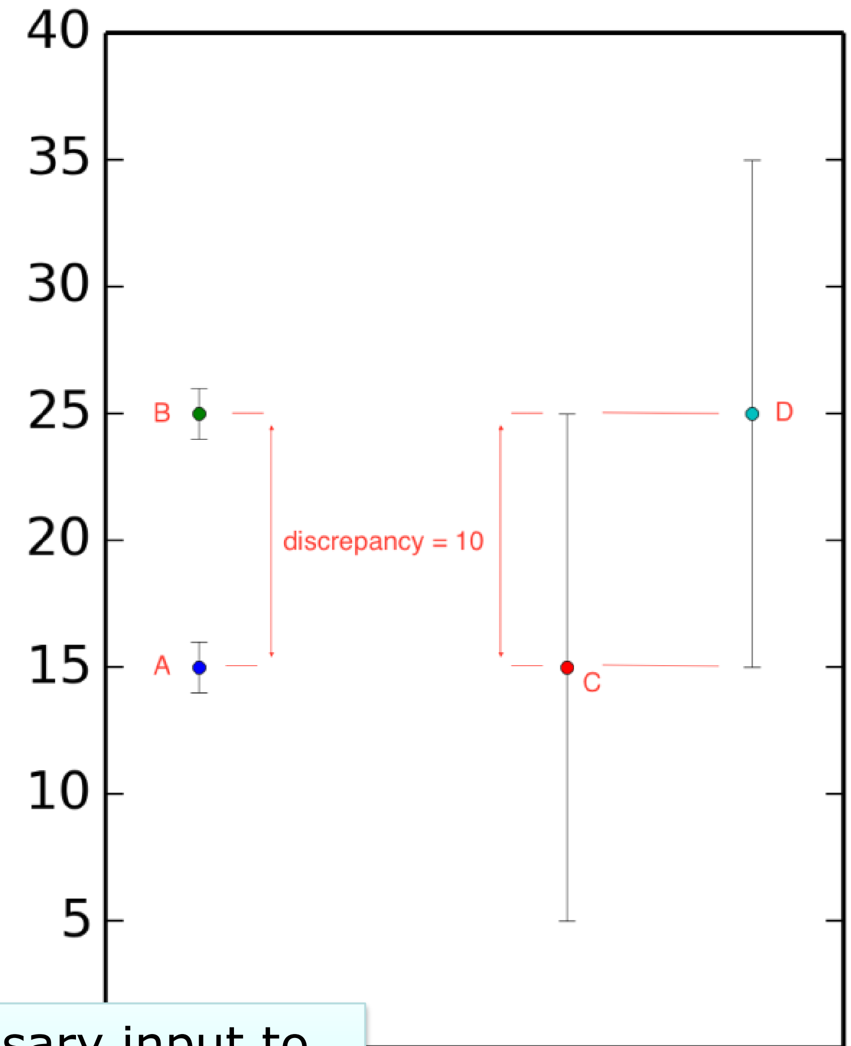
Application of uncertainty in geophysical validation



Suppose A and C are reliable reference measures.

B has the same **discrepancy** with A, as D has with C.

But B has a **bias** (significant discrepancy), while D has no proven bias (no significant discrepancy). There is a problem with the data of B!



Uncertainty is necessary input to track inconsistencies!

Consistency in uncertainty

- Naming convention
- Content
- Methods of calculation

Data products have to fulfil user requirements…

GCOS (2011) → Horizontal & temporal & vertical resolution
Accuracy
Stability

What do these terms mean exactly?
Are we being consistent with other projects?

Look in the dictionary!

Terms and definitions in FP7/H2020 projects

Authorative metrology source

JCGM: ISO, BIPM, IUPAC,…

- International vocabulary of metrology (VIM)
- Guide to uncertainty in measurement (GUM)

Approaches and scopes

- QA4ECV Atmosphere list
  - **metrology**, remote sensing, monitoring, modelling aspects
  - Implemented in QA/validation protocols for MACC-I/II, PASODOBLE, Multi-TASTE, CCI
- FIDUCEO vocabulary
  - Metrology, data correlation, data harmonization, geo-rectification, …
- GAIA-CLIM
  - Guidance note (NPL)
  - Metrology of data comparisons
- CLIPC (http://www.clipc.eu/glossary)
  - Drawn from IPCC Data distribution centre, EUPORIAS project, IS-ENES project (Climate4Impact)

Terms and definitions in projects: accuracy

closeness of agreement between a measured

VIM/GUM

Inconsistent definitions!
1-2-3 (qualitative term) vs 4 (numerical)
1-2 (any deviation) vs 3-4 (systematic effect only)

1

accuracy

2

closeness of the agreement between measurement result and true value. (Accuracy is a qualitative term only.)

Guidance note

3

A qualitative term describing the (lack of) systematic uncertainties. A measurement said to be "higher accuracy" would have smaller uncertainties associated with systematic effects. Note that it is possible to have a high accuracy measurement in the presence of large random effects.

4

the average distance between a set of measurements and the 'true' value of the object being measured. For climate prediction this can be defined as the average distance between a set of forecasts and an estimate of the observational reference. [Climate4Impact]

Accuracy in GCOS?

GCOS accuracy
~ VIM's bias ?

GCOS
(Guideline for the generation of datasets and products meeting GCOS Requirements, 2010)

Measured by the bias or systematic error of the data, i.e. the difference between the short-term average measured value of a variable and the true value. The short-term average is the average of a sufficient number of successive measurements of the variable under identical conditions, such that the random error is negligible relative to the systematic error. The latter can be introduced by instrument biases or through the choice of remote sensing retrieval schemes.

GCOS accuracy
~ VIM/GUM's total uncertainty ?

GCOS
(Systematic observation requirements for satellite-based data products for climate, 2011 update)

The user requirement for accuracy is a requirement for **closeness of agreement between product values and true values**. As true values are unknown, users are provided in practice with product values that are estimates of true values, and producers may also provide **estimates of the uncertainties** of their product values. Product uncertainty may also be assessed by users' own validation activities or by independent evaluation of available products. Each approach has its merits and each is encouraged.

Recommendation

- Regarding metrological terms, we consider the **VIM/GUM as highest authority**.
  - It has been established by JCGM, which involves BIPM, ISO, NIST, IUPAC, etc.
  - It provides a consistent metrology 'universe' to build on

→Therefore: use VIM/GUM compliant terms to describe the metrology

Uncertainty vs. Error

Found in articles, reports…

Table X presents ~~the error budget…~~

Uncertainty budget

Field X contains the ~~relative error~~ of the data product…

Relative standard uncertainty

~~The uncertainty in A causes an error in B…~~

Errors cause uncertainty, not the other way around!

Error: measured quantity value minus a reference quantity value [VIM]

Uncertainty: non-negative parameter characterizing the *dispersion of the quantity values* being attributed to a measurand, based on the information used [VIM]

Accuracy / trueness / precision

- Three very different concepts, but often confused
- Accuracy and trueness should not be given numerical values
- If you want to express things numerically, use 'total uncertainty', 'uncertainty due to random effects', 'uncertainty due to systematic effects', etc.

Measurement accuracy
closeness of agreement between a measured quantity value and a true quantity value of a measurand. [VIM]

Measurement trueness
closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value. [VIM]

Measurement precision
closeness of agreement between indications or measured quantity values obtained by replicate measurements. [VIM]

Structural uncertainty ~ Method uncertainty

"structural uncertainty"
Is there a corresponding term in VIM/GUM?

GUM

Structural uncertainty

Uncertainty arising through the choice of approach
[Thorne, Bull. Am. Meteorol. Soc., 2005]

$\cong$

Uncertainty of (measurement/retrieval) method

Uncertainty associated with the method of measurement, as there can be other methods, some of them as yet unknown or in some way impractical, that would give systematically different results of apparently equal validity [GUM, F2.5].

Practical: is this systematic uncertainty component accounted for in the data product?

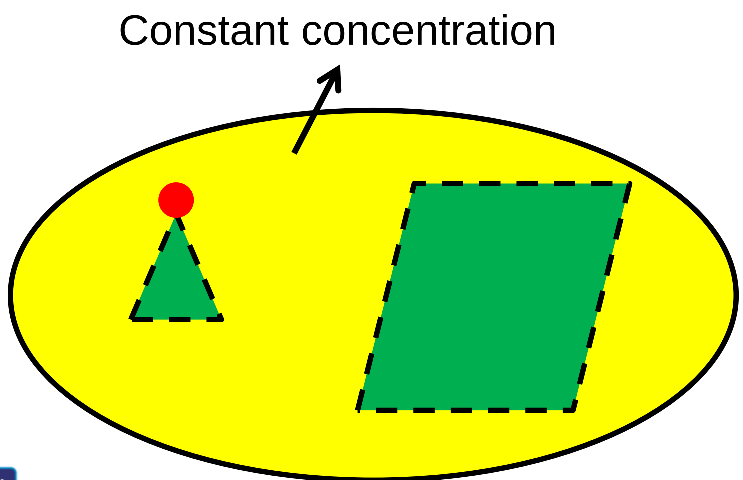
Representativeness vs. Sensitivity

"…the horizontal ~~representativeness~~ of MAX-DOAS observations…"

Sensitivity

From the context, it is clear that it is about *sensitivity.*

Representativeness
the extent to which a set of measurements taken in a given space-time domain reflect the actual conditions in the same **or different** space-time domain taken on a scale appropriate for a specific application [QA4ECV list ← Nappo]



Sensitivity
quotient of the change in an indication of a measuring system and the corresponding change in a value of a quantity being measured [VIM]

Example:
Over a flat field ground-based measurement and satellite data have non-overlapping horizontal sensitivity ranges, but they have the same representativeness.

Consistency in uncertainty

- Naming convention
- Content
- Methods of calculation

Differences among atmospheric data products:

- Uncertainty on **auxiliary parameters** included/not included
- Distinction between **random and systematic** uncertainty made/not made
- Uncertainty due to **method approximation** included/not included
- …

Consistency in uncertainty

- Naming convention
- Content
- Methods of calculation

1. Analytical (linear) uncertainty propagation
2. Analytical "posterior" uncertainties
3. Ensemble approaches
4. Validation, i.e. comparison with independent reference data
5. EO-measurement or model intercomparisons
6. Expert judgment on specific diagnostics
7. Sensitivity experiments, idealized experiments, OSSEs

cfr. the break-out group discussion on "Methods To Assess Uncertainty" (CiCS FP7 uncertainty workshop, February 2016)
cfr. BAMS meeting summary report (Otto et al., in press)

Discussion broader than measurement uncertainty

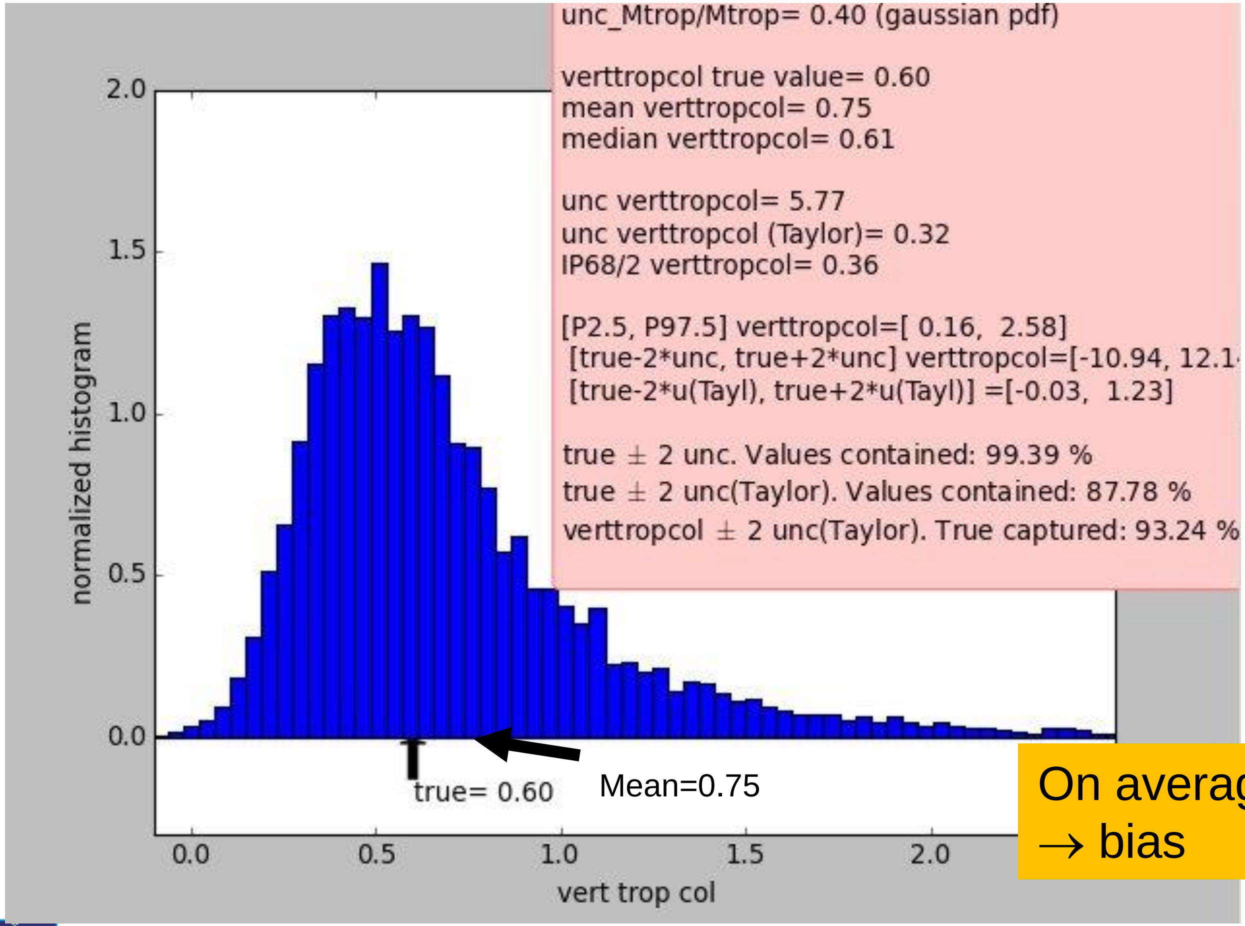
Nonlinearity and uncertainty propagation

NO2 retrieval

$$\hat{\mathrm{x}}_{\mathrm{tr}} = \frac{N_s(\mathbf{y}) - N_{s,st}}{M_{tr}(x_{a,tr}, \mathbf{b})}$$

Nonlinear measurement equation
Large input uncertainties

- **Breakdown** linear uncertainty propagation
- Introduction of a **bias**



On average, we are too high!
→ bias

VCD [molec/cm^2]

Way forward

- Circulated questionnaire on uncertainties for QA4ECV data producers
  - Terminology (is GUM/VIM followed?)
  - Format expressing uncertainty (standard deviation, 95% confidence interval,…)
  - Content (what is taken into account?)
  - Calculation method
- Input from GAIA-CLIM
  - Improve on mismatch uncertainty satellite vs. ground-based
  - Add multidimensional (spatiotemporal) aspect (see presentation Thorne-Verhoelst)

Feedback on Fiduceo

- Metrological traceability is an essential requirement of the QA4ECV QA system.
  - Fiduceo feedback on QA4ECV practices (e.g., on correlations) will be highly appreciated.
- Common terms and definitions as well as harmonised calculation approaches needed.
  - Harmonisation within C3S as targeted by Fiduceo highly desired.
- What if meas. eq. strongly nonlinear and large input uncertainties?
- From uncertainty to 95% intervals. Interpretation?
  - Bayesian credible interval. 95% chance that truth in interval, given prior probability.
  - Frequentist confidence interval. Truth captured 95% of the time. Is truth captured for 95% of the pixel data?

The QA4ECV project has received funding from the European Union's Seventh Framework Programme under grant agreement n° 607405.

Thank you for your attention!

- Questions?

The QA4ECV project has received funding from the European Union's Seventh Framework Programme under grant agreement n° 607405.

Extra material

Beyond linear uncertainty propagation

Uncertainty propagation is usually first order.
But satellite retrievals

- Use often **nonlinear** equations
- Can have **large uncertainties** on the input quantities

$$\hat{\mathrm{x}}_{\mathrm{tr}} = \frac{N_s(\mathbf{y}) - N_{s,st}}{M_{tr}\left(x_{a,tr}, \hat{\mathbf{b}}\right)}$$

Boersma (2004)
NO2 retrieval

Small simulation (values from TEMIS DOMINO v2 file):

Slant column N_s=5.0 E15 molec/cm2, $u(N_s)$=1.0E15 molec/cm2
Slant stratospheric column $N_{s,st}$ = 2.0E15 molec/cm2,$u(N_{s,st})$ = 0.25 molec/cm2
Tropospheric AMF M_{tr} = 5.0.
Uncertainty on AMF: varies between 0 and 55% in Boersma et al. (2004) (fig 10)

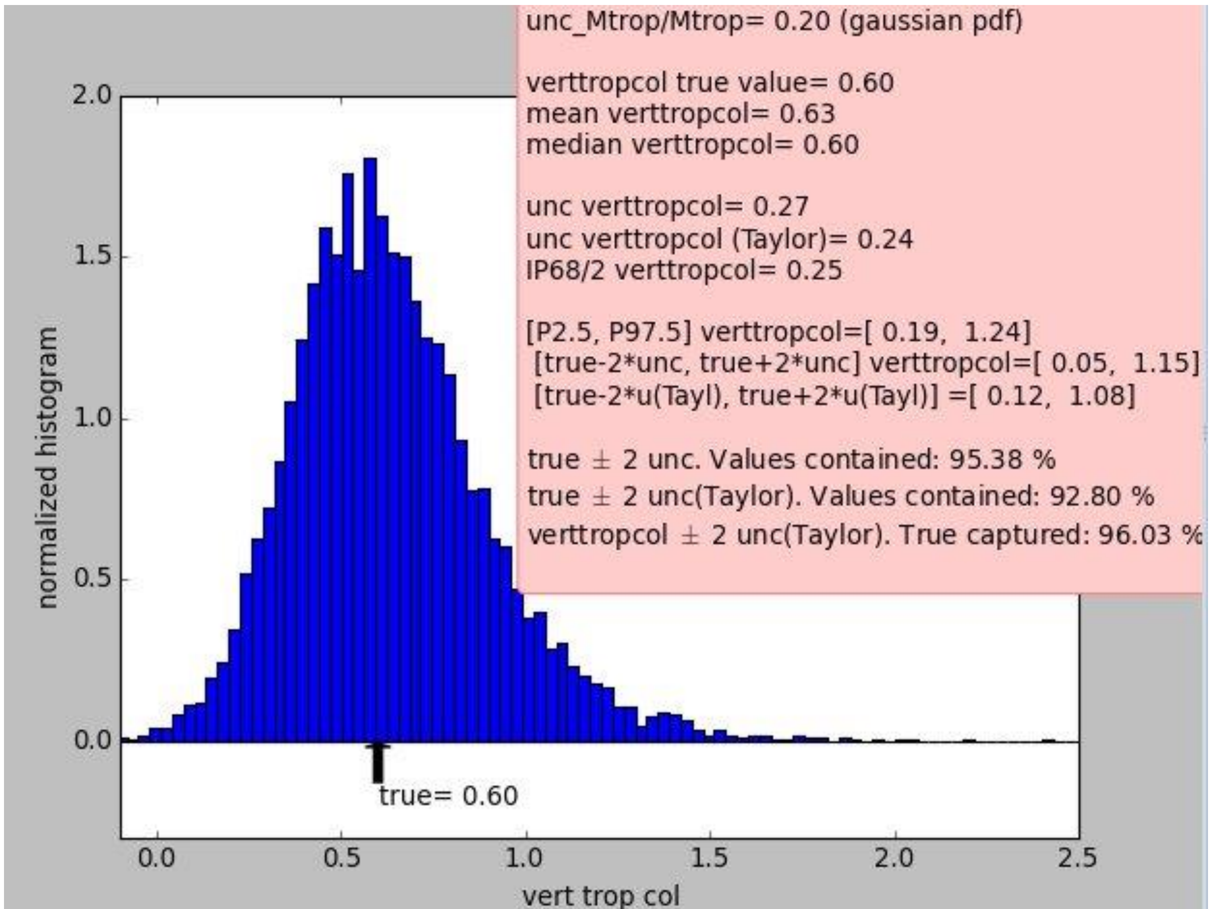
Take these as 'true' values for input quantities
Random variation on input quantities (10000 runs)
Generate 10000 values for $\hat{x}_{tr}$

"Given true values, what could have been measured?"

A small simulation

$$\hat{x}_{tr} = \frac{N_s(\mathbf{y}) - N_{s,st}}{M_{tr}(x_{a,tr}, \hat{\mathbf{b}})}$$

Take 'true' values for input quantities
Generate pdf, from Gaussian pdf on input quantities



$$\frac{u_{M_{tr}}}{M_{tr}} = 0.2$$

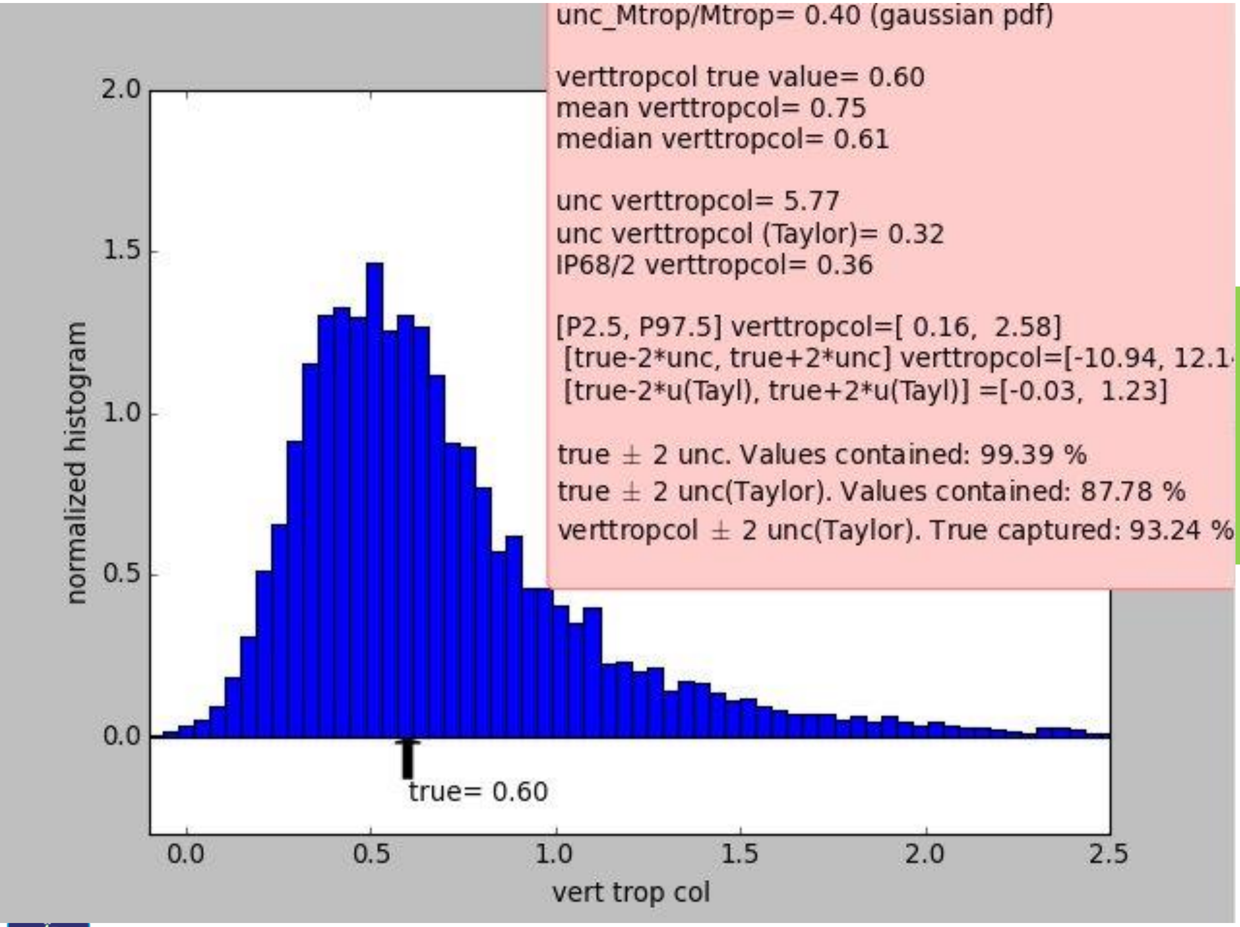
Mean close to true value
Taylor approximation reasonable
Uncertainty intervals:
close to nominal 95%

VCD [molec/cm^2]

A small simulation (c'td)

$$\hat{x}_{tr} = \frac{N_s(\mathbf{y}) - N_{s,st}}{M_{tr}(x_{a,tr}, \hat{\mathbf{b}})}$$

Take 'true' values for input quantities
Generate pdf, from Gaussian pdf on input quantities



VCD [molec/cm^2]

$$\frac{u_{M_{tr}}}{M_{tr}} = 0.4$$

Mean far from true value
Taylor approximation bad
Uncertainty intervals:

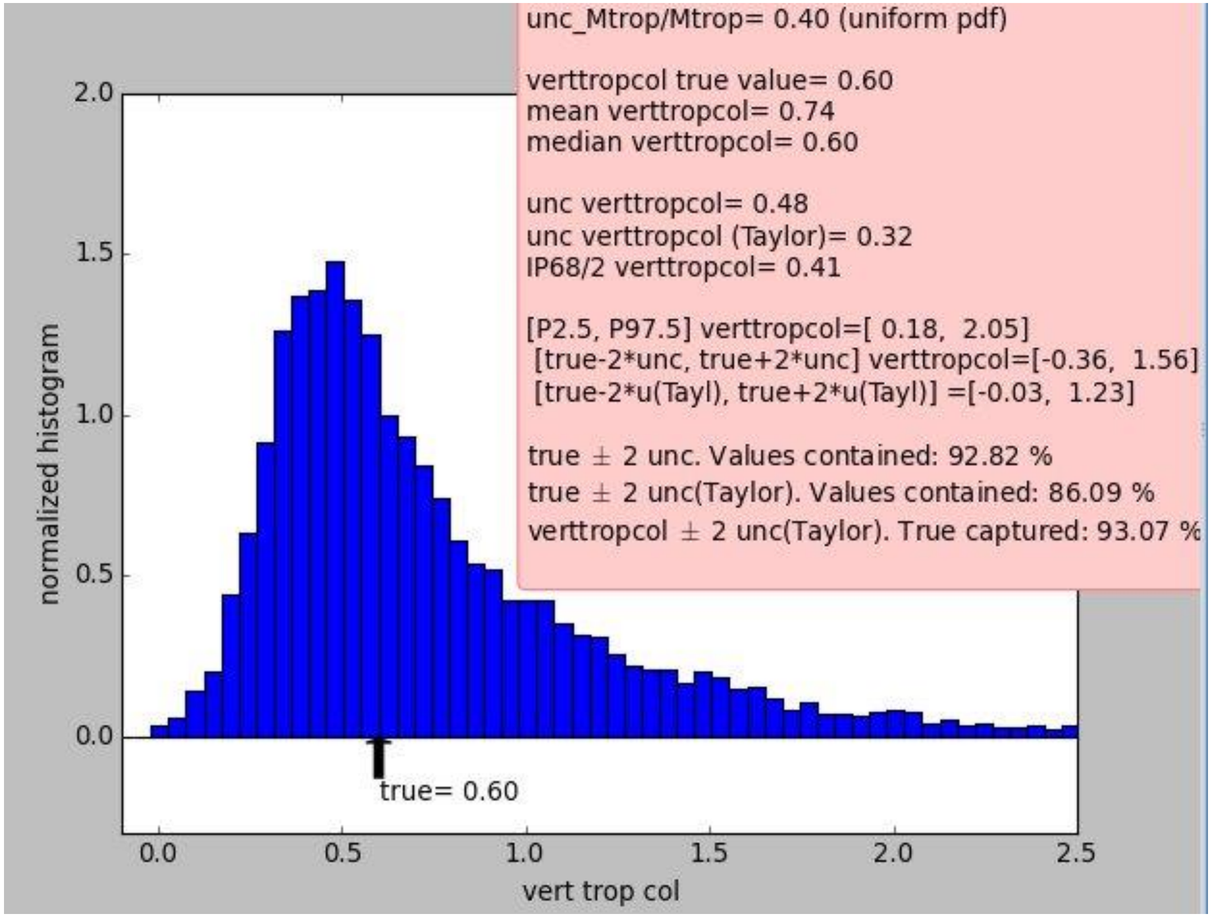
- Can deviate from 95% nominal
- Can be very wide

On average, we are too high!

A small simulation (c'td)

$$\hat{x}_{tr} = \frac{N_s(\mathbf{y}) - N_{s,st}}{M_{tr}(x_{a,tr}, \hat{\mathbf{b}})}$$

Take 'true' values for input quantities
Generate pdf. Uniform pdf for M_{tr} (negative values cannot occur)



VCD [molec/cm^2]

$$\frac{u_{M_{tr}}}{M_{tr}} = 0.4$$

u within [0.4, 8.7]

Mean far from true value
Taylor approximation bad
Uncertainty intervals:
- Can deviate from 95% nominal
- Can be quite wide

On average, we are too high!

For a better assessment, a realistic pdf on AMF is needed!

More nonlinear problems

Metrology literature

$$Y = X^2$$

Uncertainty on X **creates a bias** in Y, unless an unbiased estimator is used for Y

(Attivissimo, 2011), (Bich, metrologia, 2006), (Macii, IEEE Trans. Instrum. Meas., 2011)

$$Y = \sqrt{X_1^2 + X_2^2}$$

Even using GUM supplement 1 (Monte Carlo), one can **systematically fail to capture the truth** in the coverage interval

(Hall, Metrologia, 2008), (Willink, Metrologia, 2009)

Relevance within validation server

$$x_{ND,i} = (T_i/p_i)x_{VMR,i}$$

Conversion from VMR to ND
What if uncertainty on p is high?

$$\Delta_r = \frac{c_{sat} - c_{gr}}{c_{gr}}$$

Relative difference
What if uncertainty on c_{gr} is high?

About the Guide to uncertainty of measurement

- GUM is the metrologists' bible…
  - … with devote followers, sharp theological disputes and outright heretics (-:
- It contains 'frequentist' (type A) and 'Bayesian' (type B) aspects
- First-order Taylor series uncertainty propagation, approximate probability intervals. But to accommodate for this:
- GUM-S1: propagation of probability density functions (Monte Carlo)

Interesting literature:
- Kacker (Metrologia, 2003), on a fully Bayesian GUM.
- Willink (Metrologia, 2009). Critic on Bayesian school in general, and GUM-S1 specific.
- New GUM committee draft: *www.affidabilita.eu/pdfeventi/JCGM_100_201X_CD.pdf*
  - (Strong) critic: White (Metrologia, 2016)