Validation practices for satellite-based Earth observation data across communities

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

(1) Royal Belgian Institute for Space Aeronomy, Brussels, Belgium (tijl.verhoelst@aeronomie.be), (2) European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK, (3) Research Institute for Geo-Hydrological Protection - National Research Council, Perugia, Italy, (4) Department of Meteorology, University of Reading, Reading, UK, (5) Meteorological Institute, Universität Hamburg, Hamburg, Germany, (6) Spanish Meteorological Agency, AEMET, Madrid, Spain, (7) Magdeburg-Stendal University of Applied Sciences, Magdeburg, Germany, (8) University of Leicester, Leicester, UK, (9) KU Leuven, Leuven, Belgium, (10) The Inversion Lab, Hamburg, Germany, (11) Max Planck Institute for Meteorology, Hamburg, Germany, (12) Deutscher Wetterdienst, Offenbach, Germany, (13) Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland, (14) Ludwig-Maximilians-Universität München (LMU), Munich, Germany, deceased 2 July 2017

INTERNATIONAL SPACE SCIENCE INSTITUTE

belspo.be

Context & Aims

The validation of satellite-based measurements and their uncertainties is a challenge common to all Earth Observation (EO) communities, each of which attempts to assess the fitness-for-purpose of the satellite dataset for specific scientific or public-service applications, and to ensure the traceability of the measurements to fundamental standards. We report here on the activities and outcomes of an International Space Science Institute (ISSI) team which brought together land, ocean, and atmosphere validation experts, with the aim to share expertise and tools across EO communities. This work lead to the publication of a review paper (Loew et al., Reviews of Geophysics v55, 2017), accessible through the QR code above, or at *https://doi.org/10.1002/2017RG000562*.

Terminology

A first challenge within and across communities is the language used to report on validation methods and results. Within several domains, an effort is ongoing to adopt the nomenclature used in the metrology (i.e. measurement science) community, which is described in detail in the Vocabulaire International de Métrologie and the Guide to the expression of Uncertainty in Measurement (VIM & GUM, JCGM, 2011,2012). See also the CEOS Terms, Definitions and Cal/Val Best Practices and the WMO Quality Management Framework.

Methodology: the baseline

* Pair-wise comparisons after co-location and homogenization/harmonization;
* Consistency check w.r.t. to ex-ante uncertainties, ideally including mismatch/representativeness uncertainties;
* Scalar metrics, differentiating between systematic and random effects;
* Interpretation as a function of measurement influence quantities and geophysical phenomena.

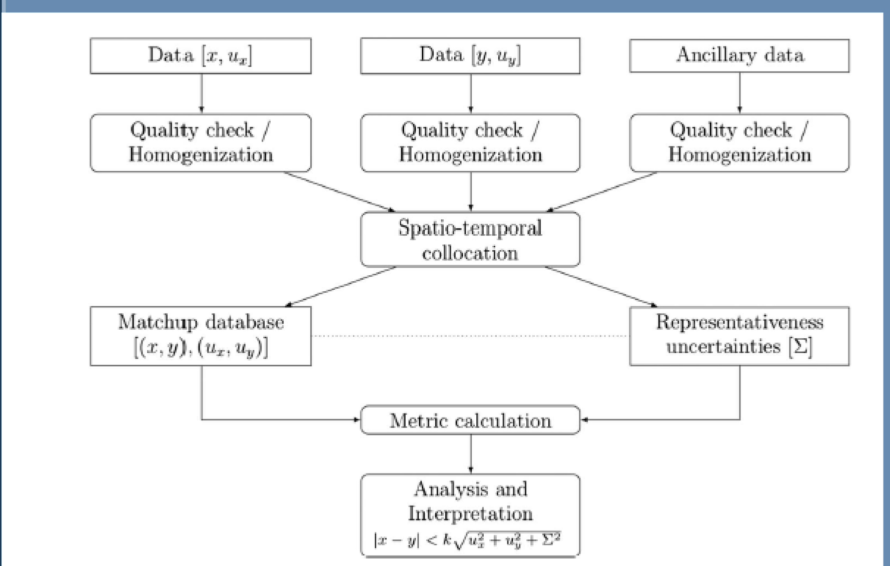


Figure 1. Schematic overview of the general validation process.

Table 4. Summary of Common Metrics Definition As Well As Their Inherent Assumptions and Sensitivities[a]

Metric	Definition	Assumptions	Sensitive to
Measures of systematic differences			
Bias	$\hat{\mu}_x - \hat{\mu}_y$	S, G	sm, sr
Median difference	$p_x^{50} - p_y^{50}$	S	sm, sr
Measures of statistical spread			
RMSD	$\sqrt{E[(x-y)^2]}$	A, S, G, L, IE, O	sm, sr, rm, rr
cRMSD	$\sqrt{E[((x-\hat{\mu}_x)-(y-\hat{\mu}_y))^2]}$	A, S, G, L, IE, O	sm, sr, rm, rr
Triple collocation measures			
R^t	$\sqrt{\frac{\hat{\sigma}_{xy}\hat{\sigma}_{xz}}{\hat{\sigma}_x^2\hat{\sigma}_{yz}}}$	A, S, G, L, IE, O	rm, rr
RMSE	$\sqrt{\hat{\sigma}_x^2 - \frac{\hat{\sigma}_{xy}\hat{\sigma}_{xz}}{\hat{\sigma}_{yz}}}$	A, S, G, L, IE, O	rm, rr
SNR (dB)	$-10 \log\left(\frac{\hat{\sigma}_x^2\hat{\sigma}_{yz}}{\hat{\sigma}_{xy}\hat{\sigma}_{xz}} - 1\right)$	A, S, G, L, IE, O	rm, rr
Statistical dependency measures			
Pearson's R	$\frac{\hat{\sigma}_{xy}}{\sqrt{\hat{\sigma}_x^2\hat{\sigma}_y^2}}$	A, S, G, L, IE, O	rm, rr
Spearman's ρ	$\frac{\hat{\sigma}_{r_{xy}}}{\sqrt{\hat{\sigma}_{r_x}^2\hat{\sigma}_{r_y}^2}}$	A, S, IE, O	rm, rr
Kendall's τ	$\frac{n_c - n_d}{n_0}$	A, S, IE, O	rm, rr
Mutual information, I	$\iint f_{x,y}(x,y)\log\left(\frac{f_{x,y}(x,y)}{f_x(x)f_y(y)}\right)dxdy$	A, S, IE, O	rm, rr
Temporal stability measures			
Absolute temporal stability		L	

[a]Abbreviations: S, stationarity; G, Gaussianity; L, linearity; IE, independence of error terms; O, orthogonality; A, additive error model; sm, systematic measurement uncertainties; sr, systematic representativeness differences; rm, random measurement uncertainties; rr, random representativeness differences.

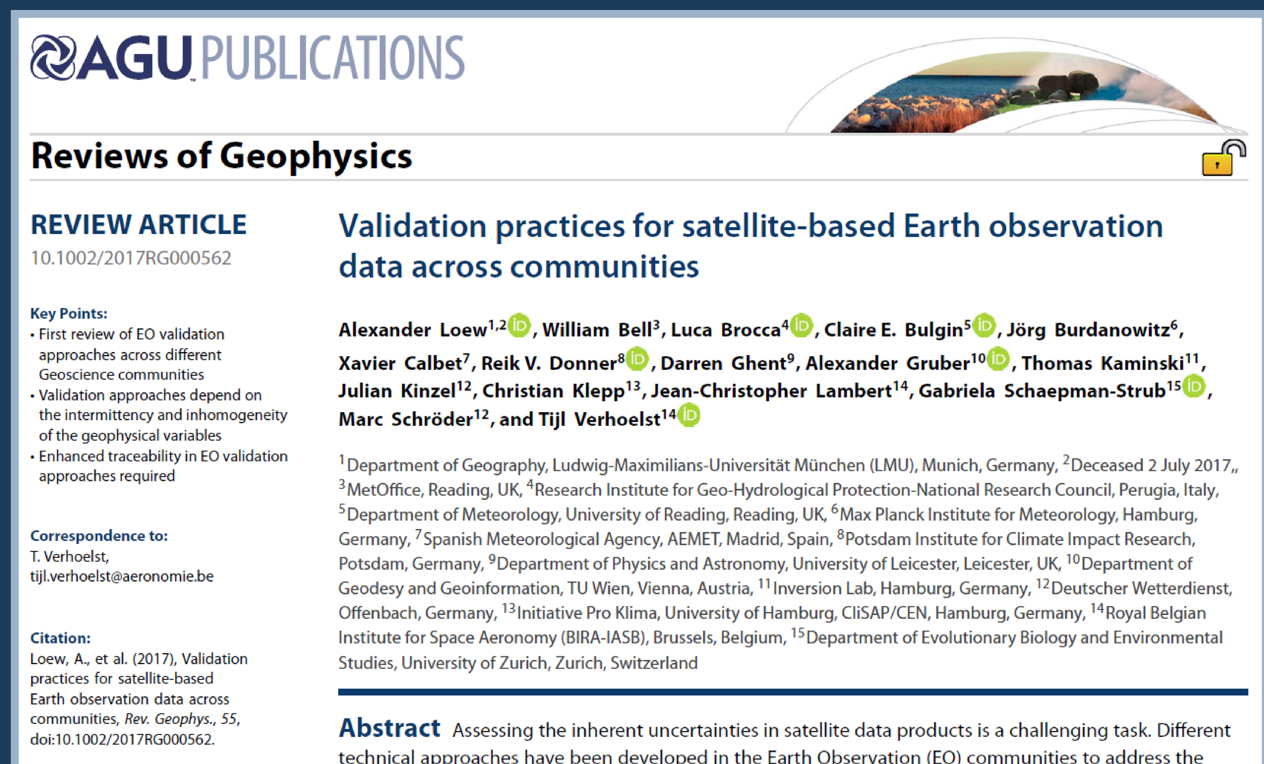
Methodology: advanced techniques

* Triple co-location and multiple triple co-location analysis (TCA)
* Spectral methods (Fourier and Wavelet)
* Field intercomparison and Functional Network Analysis
* Consistency through process models
* Indirect validation
* Data assimilation
* Self co-location
* ...

Examples for each are available in the Appendix of the review paper.

User requirements

* The validation question is essentially: Are the data fit-for-purpose, do they meet the user requirements?
* In all communities, a disconnect was found between the detailed output of the advanced validation methodologies and the often simplistic and/or ambiguous user requirements.
* User requirements need to be specific (e.g. about spatial and temporal domain), differentiated (e.g. between systematic and random effects) and traceable, i.e. with a well-documented origin.



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.

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

Reference measurements

The backbone of any validation exercise is the availability of reference measurements, where "reference" implies:

* Traceability, i.e. they can be linked through an unbroken chain of calibrations and comparisons to a metrological reference (Système International or community-agreed);
* Full uncertainty characterization;
* Availability;
* At the network level: Sufficient coverage of the potential parameter space.

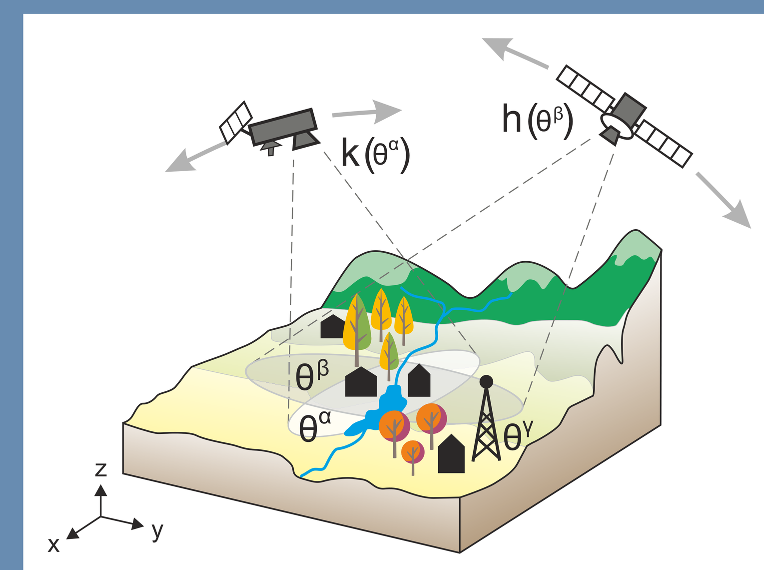
While the maturity of reference measurements varies across communities, their importance is realized in all, as is evident from the increased interest from (space) agencies in supporting in-situ and ground-based networks.

Co-location mismatch

Spatiotemporal co-location is one of the more challenging aspects of a validation exercise. Concerns to be addressed are:

* The need to have co-located measurement close to each other relative to the spatiotemporal scale on which the variability of the geophysical field becomes comparable to the measurement uncertainties.
* If possible, differences in spatiotemporal resolution should be minimized.
* The co-location criteria must take into account the need for a sufficient amount of co-located pairs for robust statistical analysis. This is often at odds with the 1st requirement and a compromise must be made.

Within the different communities, a wealth of techniques have been developed to optimize the co-location, either on the original measurement grids or using interpolation and up– and downscaling techniques. Nevertheless, some co-location mismatch usually remains, and various approaches exist to assess representativeness errors and co-location uncertainty. A proper consistency check between satellite and reference data takes into account not only the uncertainties of each data set, but also the uncertainties inherent to the imperfect co-location.



Recommendations

* Refine user requirements: They need to be specific, sufficiently differentiated, traceable and fully documented.
* Follow metrological (i.e. measurement science) terminology and uncertainty expression and assessment methods. Reference works are the Vocabulaire International de Métrologie (VIM) and the Guide to the expression of Uncertainty in Measurement (GUM).
* Ensure the traceability of reference measurements, including a detailed uncertainty budget.
* Ensure the sustained availability of fiducial reference measurements from in-situ and ground-based networks. This should be a continuous focal point for (space) agencies and service providers.
* Go beyond scalar pair-wise comparison metrics and include the satellite and reference measurement uncertainties.
* Be aware of scale mismatch: optimize co-location criteria and quantify any remaining co-location uncertainty
* Establish and publish best practices, and share traceable, open-source validation software tools.