Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Initial input from WP3 to the gap analysis and impacts document

A Horizon 2020 project; Grant agreement: 640276

Date: 30 June 2015

Lead Beneficiary: BIRA

Nature: R

Dissemination level: PU

Work-package	WP 3 (Comparison error budget closure – Quantifying metrology related uncertainties of data comparisons)
Deliverable	D3.1
Title	Initial input from WP3 to the gap analysis and impacts document
Nature	R
Dissemination	PU
Lead Beneficiary	Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
Date	30 june 2015
Status	Final, version 1
Editors	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB)
Authors	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB), D. Butterfield (NPL), T. Gardiner (NPL), R. Kivi (FMI-ARC), P. Kolmonen (FMI), G. de Leeuw (FMI), F. Madonna (CNR), L. Mona (CNR), A. Fassò (UniBergamo), K. Rannat (TUT)
Contacts	tijl.verhoelst@aeronomie.be, j-c.lambert@aeronomie.be
URL	http://www.gaia-clim.eu/

This document has been produced in the context of the GAIA-CLIM project. The research leading to these results has received funding from the European Union's Horizon 2020 Programme under grant agreement n° 640276. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

GAIA-CLIM: Gap Identification Template[1])

New gap types introduced here:

- **Measurement uncertainty (smoothing and sampling)**
- **Measurement uncertainty (representativeness)**
- **Comparison uncertainty (causes)**
- **Comparison uncertainty (mitigation)**
- **Comparison uncertainty (quantification)**

Gap Identifier G<wp>.<no>	**Gap Type** [2])	**Keywords** [3]) [Up to 10 (max)]	**ECV(s)** [Specify if not generic]	**Gap Description** (<100 characters)	**Trace** (both underlying WP deliverable(s) as well as external papers, reports etc)	**Gap Impacts** (Bulleted summary)	**Envisaged Remedy** (including timescale and cost estimate if possible)	**Remedy addressed in GAIA-CLIM** (Yes/No)
G3.1	**Comparison uncertainty (causes)**	**Error budget; Natural variability; Co-location criteria**	**all**	**Incomplete knowledge of spatio-temporal atmospheric variability at the scale of the inter-comparisons.**	**D3-1, including Annexes 1, 2 and 3**	• **Difficulty determining optimal co-location and coincidence criteria** • **Difficulty estimating sampling and smoothing difference errors**	• **Statistical analysis on existing co-located data sets.** • **More dedicated field campaigns** • **Future missions with high spatial resolution will provide further insight (e.g. Sentinels).**	**Partially**
G3.2	**Comparison uncertainty (mitigation)**	**Error budget; Natural variability; Co-location criteria**	**all**	**Limited quantification of the impact of co-location criteria.**	**D3-1, including Annexes 1, 2 and 3**	• **Difficulty to assess the importance of natural variability in the total error budget.**	**Generic approach to what a good co-location can be, followed by specific studies exploring different co-location criteria in a systematic way.**	**Yes**
G3.3	**Comparison uncertainty (mitigation)**	**Error budget; Natural variability; Co-location criteria; Validation protocol**	**all**	**Missing generic and specific standards for co-location criteria in validation work.**	**D3-1, including Annexes 1, 2 and 3**	• **Difficulty to compare different validation exercises.** • **Limits optimal use of the ground-based networks.**	**Publication of generic and detailed validation protocols, including metrology aspects of a data comparison and recommendations on optimal co-location criteria.**	**No**

G3.4	**Measurement uncertainty (smoothing and sampling)**	**Error budget; Natural variability; Co-location criteria**	**all**	**Limited characterization of the multi-dimensional (spatio-temporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties.**	**D3-1, including Annexes 1, 2 and 3**	• **Unknown contribution to the comparison error budget.** • **Limits interpretation in terms of data quality.**	**GAIA-CLIM WP3 will describe and quantify these uncertainties for selected ECVs and instruments.**	**Partially**
G3.5	**Measurement uncertainty (representative ness)**	**Error budget; Natural variability; Co-location criteria**	**all**	**Representativeness uncertainty missing for higher-level data based on averaging of individual measurements.**	**D3-1, including Annexes 1, 2 and 3**	• **Unknown contribution to the total uncertainty of the measurement, impacting both scientific use and validation work.**	**Studies quantifying the representativeness of averages, either using physical or statistical modelling tools.**	**No**
G3.6	**Comparison uncertainty (quantification)**	**Error budget; Natural variability; Co-location criteria**	**all**	**Missing comparison error budget decomposition including errors due to sampling and smoothing differences.**	**D3-1, including Annexes 1, 2 and 3**	• **Limits interpretation of comparisons in terms of data quality and fitness-for-purpose.**	**Studies quantifying the errors due to smoothing and sampling differences in actual comparisons, either using physical or statistical modelling tools.**	**Yes, for some ECVs and instrument combinations**

Notes

1) Gaps are assumed to be supported by **full text entries in the underlying WP deliverables** indicated in the column ‘Trace’ (a suggested full text format is provided below)
2) **Proposed Gap Types** (either scientific, technical, organizational), please complete the following list of gap types if necessary. Note: Gap type will be used to help organizing the table with collected gaps: coverage(horizontal); coverage (vertical); coverage (temporal) or ‘missing data’; resolution (vertical); uncertainty large (systematic); uncertainty large (random); uncertainty unknown (systematic); uncertainty unknown (random); …
3) **Proposed Keywords**, please complete the following list of keywords if necessary. Keywords will facilitate search tools for the gaps related to e.g. any networks, techniques: [measurement technique(s)], [network(s)], relative uncertainty, absolute uncertainty, error budget, smoothing error, retrieval, calibration, representativity, etc.

G3.1 Missing knowledge of spatio-temporal atmospheric variability at the scale of the inter-comparisons

Gap Type: *Comparison uncertainty (causes)*
Gap Keywords: *Error budget; Natural variability; Co-location criteria*
ECV(s): *All*
Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

When comparing two different measurements of an atmospheric variable, there almost inevitably exists a mismatch in measurement location, time, and smoothing properties (see e.g. Lambert et al. 2012 for the case of water vapour comparisons). As a result, spatio-temporal atmospheric variability and structures will impact these comparisons and introduce additional errors, not accounted for by the (instrumental and retrieval) uncertainties reported with the data. To be able to draw meaningful conclusions from the comparisons, these additional errors must either be minimized to well below the measurement uncertainties, or they must be reliably quantified. A key prerequisite for either approach is a quantified understanding of atmospheric variability at the scale of the comparisons.
While some literature exists on the representativeness of atmospheric measurements (e.g. Nappo et al., 1982; Kitchen et al. 1989; Pappalardo et al., 2010) and on the small-scale spatio-temporal variability of several key ECVs (e.g. Sparling et al., 2006; Sofieva et al. 2008), this work is rarely comprehensive, e.g. covering only a selected site or altitude range. The validation of satellite data records with pseudo global networks of ground-based reference instruments on the other hand requires a suitable quantification of atmospheric variability in very diverse conditions, covering all latitudes, altitudes, dynamical conditions, degrees of pollution,… This gap therefore concerns the need for a better, more comprehensive, understanding of atmospheric variability at the scales involved in typical satellite-ground comparisons.

Gap Impacts

Without a proper, quantified, understanding of atmospheric variability down to the scale of an individual measurement, it is impossible to determine optimal spatio-temporal co-location criteria. Co-location criteria are therefore often adopted from community practices (e.g., the classical radius of 50 to 500 km around a station) established principally to guarantee a sufficient amount of comparison pairs, but without accurate assessment of the impact of co-location mismatch and associated atmospheric variability on the comparisons. Moreover, without proper knowledge of the atmospheric variability, it is impossible to estimate a posteriori how much the co-location mismatch uncertainties contribute to the uncertainty budget of the data comparisons. Ultimately, this gap impacts strongly the potential for interpretation of the comparison results in terms of data quality.

Gap Remedy

Several approaches can be envisaged in order to improve knowledge on the spatio-temporal variability of the atmosphere at the scale of typical ground-satellite co-locations. First, existing data sets from both the ground networks and from past and current satellite instruments should be analyzed using statistical tools that allow to separate between 1D measurement uncertainties and errors due to co-location mismatches (e.g. Fasso et al. 2014). This approach will be explored within GAIA-CLIM's WP3 as part of a multi-method effort to estimate co-location mismatch errors in the ground-based validation of a few key ECVs. Second, dedicated field campaigns should be planned to sample the atmosphere at very high spatial and temporal resolution at selected sites of particular interest. These can either make use of airborne instruments on both manned and unmanned aircraft, or they should consist of an increased density of measurements with the reference instruments also used in the validation. Such campaigns are beyond the scope of GAIA-CLIM but results from WP3 could be instrumental in the design of future campaigns. Finally, the unprecedented spatial and temporal resolution of some upcoming satellite missions, such as the Sentinels, will allow a better quantification of atmospheric variability, spatially over the entire globe, and temporally within the field-of-view of the upcoming geo-stationary missions (Sentinel-4, TEMPO, GEMS). Nevertheless, their resolution will not match that of in-situ reference measurements such as those obtained from balloon-borne radiosondes and ozonesondes, and consequently they will not provide a comprehensive solution to this gap.

References (see also D3-1 annexes 1 through 3):

Fasso et al., “Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles”, AMT v7, 2014

Kitchen et al., “Representativeness errors for radiosonde observations”, Q. J. R. Meteorol. Soc. v115, p673-700, 1989

Lambert et al. “Ground-based remote sensing and in-situ methods for monitoring atmospheric water vapour – Chapter 9: Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues”, ch9, p177-199, ISSI, 2012

Nappo et al., “Workshop on the representativeness of meteorological obervations, June 1981, Boulder, Colorado” Bull. Am. Meteorol. Soc. v63, 1982

Pappalardo et al., “EARLINET correlative measurements for CALIPSO: First intercomparison results”, J.G.R.: Atmospheres v115, 2010

Sofieva et al., “On the variability of temperature profiles in the stratosphere: Implications for validation”, Geophys. Res. Lett. v35, 2008

Sparling et al., “Estimating the impact of small-scale variability in satellite measurement validation”, J.G.R.: Atmospheres v111, 2006

G3.2 Limited quantification of the impact of co-location criteria in validation work

Gap Type: *Comparison uncertainty (mitigation)*

Gap Keywords: *Error budget; Natural variability; Co-location criteria*

ECV(s): *All*

Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

The impact of a particular choice of co-location criterion, a choice which is most often not based on detailed information on the properties of atmospheric variability at the scale of the comparisons, is only rarely explored in the atmospheric validation literature. While some in-depth studies have already been performed (e.g. Guerlet et al. 2013, Van Malderen et al. 2014), work in which the influence of different criteria both on the number of co-located pairs and on the impact of natural variability is quantified, this is at the moment not common practice. In particular, when only a single criterion is adopted, atmospheric variability can only be suspected to impact the comparisons if for instance the standard deviation on the differences is larger than the combined measurement uncertainty (e.g. De Maziere et al. 2008).

Gap Impacts

Without some quantification of the impact of the particular choice of co-location criterion that was adopted, it is virtually impossible to assess the contribution of natural variability to the total error budget of the comparisons. As such, this gap impacts significantly the potential interpretation of the comparison result in terms of data quality.

Gap Remedy

Potential remedies to this gap can (should) include both the testing of alternative criteria, e.g. tighter in maximum allowed spatio-temporal distance, and confrontation of the criteria with the known properties of the atmospheric variability at those scales (if this information exists, cfr. G3.1). It is recommended here that such tests be part of a generic validation protocol (cfr. G3.3).

References (see also D3-1 annexes 1 through 3):

De Maziere et al., “Validation of ACE-FTS v2.2 methane profiles from the upper troposphere to the lower mesosphere”, ACP v8, 2008

Guerlet et al., “Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements”, J.G.R. v118, 2013

Van Malderen, R. et al., “A multi-site intercomparison of integrated water vapour observations for climate change analysis”, AMT v7, 2014

G3.3 Missing generic and specific standards for co-location criteria in validation work

Gap Type: *Comparison uncertainty (mitigation)*
Gap Keywords: *Error budget; Natural variability; Co-location criteria*
ECV(s): *All*
Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

The number of different co-location criteria used in validation work on atmospheric ECVs is large, ranging from fixed maxima imposed on spatial and temporal distance (e.g. Ohyama et al. 2013, Dils et al. 2014, Hubert et al. 2015), over criteria based on the state/dynamics of the atmosphere (e.g. Wunch et al. 2011) or on representativeness area's derived from models (e.g. Oshchepkov et al. 2012), to airmass matching techniques that take into account the actual 3D/4D sensitivity of each measurement (e.g. Lambert et al. 1997,1999, Balis et al. 2007). These lead to significant differences between the different validation exercises, in particular regarding the contribution of atmospheric variability to the total error budget. To ensure reliable and traceable validation results, as required in operational validation work, community-agreed standards for co-location criteria should be developed, based on a detailed understanding of atmospheric variability, measurement properties, and user needs.

Gap Impacts

The lack of standard practices for co-location criteria with metrology grounds leads both to the use of sub-optimal criteria in many validation studies, and to difficulties when inter-comparing the results of validation work on related products, e.g. in the context of delta validations between different retrieval algorithms.

Gap Remedy

Co-location protocols, starting with a common, generic background compliant with metrology standards, and differentiated further according to ECV, instrument type, atmospheric regime, etc., should be established and used as a baseline in all operational atmospheric remote sensing validation using ground-based reference measurements.

References (see also D3-1 annexes 1 through 3):

Balis et al., "Ten years of GOME/ERS2 total ozone data – The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8", J.G.R. v112, 2007

Dils et al., "The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO2 and CH4 retrieval algorithm products with measurements from the TCCON", AMT v7, 2014

Hubert et al., "Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records", accepted for publication in AMTD, 2015

Lambert, J.-C., et al., Comparison of the GOME ozone and NO_2 total amounts at mid-latitude with ground-based zenith-sky measurements, in *Atmospheric Ozone - 18th Quad. Ozone Symp., L'Aquila, Italy, 1996*, R. Bojkov and G. Visconti (Eds.), Vol. I, pp. 301-304, 1997.

Lambert et al., "Investigation of Pole-to-Pole Performances of Spaceborne Atmospheric Chemistry Sensors with the NDSC", J. Atmos. Sci. v56, 1999

Ohyama et al., "Atmospheric Temperature and Water Vapour Retrievals from GOSAT Thermal Infrared Spectra and Initial Validation with Coincident Radiosonde Measurements", SOLA v9, p143-147, 2013

Oshchepkov et al., "Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space: Validation of PPDF-based CO2 retrievals from GOSAT", J.G.R. v117, 2012

Wunch et al., "A method for evaluating bias in global measurements of CO2 total columns from space", ACP v11, 2011

G3.4 Limited characterization of the spatio-temporal smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties

Gap Type: *Measurement uncertainty (smoothing and sampling)*
Gap Keywords: *Error budget; Natural variability; Co-location criteria; smoothing error*
ECV(s): *All*
Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

Remotely sensed data are often considered as column-like or point-like samples of an atmospheric variable, e.g., WP2 assumes column and vertical profile measurements of ozone, water vapor etc. at the vertical of the station. This is also the general assumption with satellite data, which are assumed to represent the column or profile at the vertical of the satellite field-of-view footprint in case of nadir sounders, and atmospheric concentrations along a vertical suite of successive tangent points in the case of limb and occultation sounders. In practice the quantities retrieved from a remote sensing measurement integrate atmospheric information over a tri-dimensional airmass and also over time. E.g., ground-based zenith-sky measurements of the scattered light at twilight integrate stratospheric UV-visible absorptions (by ozone, NO2, BrO etc.) over several hundreds of kilometers in the direction of the rising or setting Sun (Lambert et al., 1997). A satellite limb measurement will actually be sensitive to the atmosphere along the entire line-of-sight towards the photon source, depending on the specific emission, absorption, and scattering processes at play (e.g. Von Clarmann et al., 2009). Similarly, in-situ measurements of atmospheric profiles can not be associated with a single geo-location and time stamp, due for instance to balloon drift (e.g. Seidel et al. 2011). In a variable and inhomogeneous atmosphere, this leads to additional uncertainties not covered in the 1-dimensional uncertainties reported with the data (e.g. Lambert et al., 2011). A prerequisite for quantifying these additional uncertainties of multi-dimensional nature is not only a quantification of the atmospheric variability at the scale of the measurement (cfr. G3.1), but also a detailed understanding of the smoothing and sampling properties of the remote sensing system and associated retrieval scheme. Pioneering work on multi-dimensional characterization of smoothing and sampling properties of remote sensing systems and associated uncertainties was initiated during the last decade (e.g., in BELSPO/ProDEx projects SECPEA and A3C and in the EC FP6 GEOmon project), but in the context of integrated systems like Copernicus and GCOS, appropriate knowledge of smoothing and sampling uncertainties, still missing for several ECVs and remote sensing measurement types, has to be further developed and harmonized.

Gap Impacts

Without a detailed quantification of the uncertainties due to the 4-D smoothing and sampling properties of the measurement systems, it is impossible to gauge their importance in the total uncertainty of the measurement, and consequently also in the error budget of the comparison of two different measurements.

Gap Remedy

Smoothing and sampling properties can be quantified either from detailed modelling of the actual measurement process, or in a more pragmatic way based on the measurement principle and physical considerations (e.g. Lambert et al. 2011). Alternatively, they can in some cases be estimated empirically. Significant work on key ECVs is foreseen within GAIA-CLIM WP3.

References (see also D3-1 annexes 1 through 3):

Von Clarmann et al., “The horizontal resolution of MIPAS”, AMT v2, 2009

Lambert, J.-C., et al., Comparison of the GOME ozone and NO_2 total amounts at mid-latitude with ground-based zenith-sky measurements, in *Atmospheric Ozone - 18th Quad. Ozone Symp., L’Aquila, Italy, 1996*, R. Bojkov and G. Visconti (Eds.), Vol. I, pp. 301-304, 1997.

Lambert et al., “Multi-dimensional characterisation of remotely sensed data”, EC FP6 GEOmon Technical Notes, 2011

Seidel et al., “Global radiosonde balloon drift statistics”, J.G.R. v116, 2011

G3.5 Representativeness uncertainty missing for higher-level data based on averaging of individual measurements

Gap Type: *Measurement uncertainty (representativeness)*
Gap Keywords: *Error budget; Natural variability; Co-location criteria*
ECV(s): *All*
Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

For several purposes, individual measurements can be averaged into spatial and temporal means, e.g., monthly zonal means are widely used in climate research and monitoring. While some information is usually provided on the variance of the underlying measurements, the representativeness of these measurements is often not assessed, and not provided with the (level-3) data. Also in validation work, measurements after sometimes averaged after co-location (e.g. Valks et al., 2011; Schneising et al., 2012), but the representativeness issue is only rarely addressed (see Coldewey-Egbers et al., 2015 for a case where it is quantified).

Gap Impacts

The representativeness uncertainty can be larger than the formal uncertainty on the mean, and thus needs to be known, both for scientific use and for validation purposes.

Gap Remedy

Representativeness uncertainties can be estimated from by simulating the averaging process, with its particular sampling properties, on global gridded data, either from models or from well-sampled observational data sets. This work is not specifically foreseen within GAIA-CLIM, but it does constitute a straightforward application of the tools developed in WP3.

References (see also D3-1 annexes 1 through 3):

Coldewey-Egbers et al., "The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative", AMT v8, 2015

Schneising et al., "Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results", ACP v12, 2012

Valks et al., "Operational total and tropospheric NO2 column retrieval for GOME-2", AMT v4, 2011

G3.6 Comparison uncertainty budget decomposition including errors due to sampling and smoothing differences

Gap Type: *Comparison uncertainty (quantification)*
Gap Keywords: *Error budget; Natural variability; Co-location criteria*
ECV(s): *All*
Trace (external refs): *D3-1 Annex 1,2 and 3*

Gap Description

Only in a few particular cases is it possible to adopt co-location criteria that result in a sufficiently large number of co-located pairs, while at the same time keeping the impact of atmospheric variability on the comparisons (due to spatio-temporal mismatches) well below the measurement uncertainties. In all other cases, the discrepancy between two data sets will contain non-negligible terms arising from co-location and smoothing mismatch. If a reliable evaluation of the data quality is to be drawn from the comparisons, these so-called metrological errors have to be quantified so that the error budget can be decomposed into measurement and comparisons errors. This requires both a quantification of the atmospheric variability (cfr G3.1) and of the sampling and smoothing properties of the instruments that are being compared. Pioneering work on such error budget decomposition has been published in the past decade for temperature, humidity and ozone (e.g. Ridolfi et al., 2007; Cortesi et al., 2007; Lambert et al., 2011,2012; Fassò et al., 2014; Verhoelst et al., 2015), but such an approach is far from being a common practice.

Gap Impacts

A reliable interpretation of comparisons between satellite and ground-based reference measurements requires a decomposition of the error budget in both measurement and comparisons errors, otherwise there exists the risk that either some data quality issues are not revealed or some discrepancies are erroneously attributed to one of the data sources, rather than to the impact of differences in sensing atmospheric inhomogeneities.

Gap Remedy

Depending on the impact of G3.1 and G3.4, error budget decomposition in principle only requires a computational effort to be taken into account when defining a validation exercise. This work will be performed within WP3 of GAIA-CLIM for key ECV and instrument combinations.

References (see also D3-1 annexes 1 through 3):

Ridolfi et al., “Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements”, ACP v7, 2007

Cortesi et al., “Geophysical validation of MIPAS-ENVISAT operational ozone data”, ACP v7, 2007

Fassò et al., “Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles”, AMT v7, 2014

Lambert et al., “Multi-dimensional characterisation of remotely sensed data”, EC FP6 GEOmon Technical Notes, 2011

Lambert et al. “Ground-based remote sensing and in-situ methods for monitoring atmospheric water vapour – Chapter 9: Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues”, ch9, p177-199, ISSI, 2012

Verhoelst et al., “Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons”, accepted for publication in AMTD, 2015

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Initial input from WP3 to the gap analysis and impacts document, Annex 1:

Literature review on temperature and humidity data validation

A Horizon 2020 project; Grant agreement: 640276

Date: 30 June 2015

Lead Beneficiary: BIRA

Nature: R

Dissemination level: PU

Work-package	WP 3 (Comparison error budget closure – Quantifying metrology related uncertainties of data comparisons)
Deliverable	D3.1
Title	Initial input from WP3 to the gap analysis and impacts document, Annex 1: Literature review on temperature and humidity data validation
Nature	R
Dissemination	PU
Lead Beneficiary	Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
Date	30 june 2015
Status	Final, version 1
Editor	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB)
Authors	D. Butterfield (NPL), T. Gardiner (NPL)
Contacts	Tom.gardiner@npl.co.uk
URL	http://www.gaia-clim.eu/

This document has been produced in the context of the GAIA-CLIM project. The research leading to these results has received funding from the European Union's Horizon 2020 Programme under grant agreement n° 640276. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
1	WV	Continuous quality assessment of atmospheric water vapour measurement techniques: FTIR, Cimel, MFRSR, GPS, and Vaisala RS92	Parallel measurements for FTIR, Cimel & MFRSR. Spatial coincidence <50m 1 hour for Radiosonde 15 minute for GPS	Δt assessed and 1h found to be the optimum comparison with radiosondes	FTIR / RS92 scatter 19% -29% dependant on altitude Cimel / MFRSR up to 60% GPS underestimates compared to FTIR
2	WV	Collocating satellite-based radar and radiometer measurements – methodology and usage examples	Collocation finding procedure: 1. Orbits (granules) with time overlap are selected. 2. Orbit sections are selected according to a rough temporal criterion. 3. Measurements possibly fulfilling the spatial criterion are selected. 4. The temporal criterion is applied to the selected measurements.	Some discussion on the quality of the collocation.	Sampling effects taken into account due to differences in footprint sizes. Collocations studied: MHS-NOAA18, CPR-CloudSat, Artificial Neural Network to develop new MHS based product
3	T	An assessment of differences in lower stratospheric temperature records from (A)MSU, radiosondes, and GPS radio occultation.	Input profiles interpolated to 43 vertical levels. Results binned into horizontal and temporal resolution of 2.5^o x 2.5^o (monthly means). Latitudinal bands created and then aggregated into larger bands, weighted for surface area.	No assessment of effect or development of coincidence criteria.	Sampling errors estimated due to sample density (area covered). (A)MSU data do not need sampling error correction because they provide very dense horizontal sampling. The consistency of radiosondes and RO was improved substantially by subtracting their respective sampling errors. Poor vertical (A)MSU, resolution may miss important features of the vertical atmospheric structure. This points to the advantage of homogeneously distributed measurements with high vertical resolution.
4	WV	ARIS-Campaign: intercomparison of three	MLS profile: ±2^o (220km) in latitude and ±5^o in longitude	No assessment of effect or development of coincidence	24 hour averaging used for profile comparisons.

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
		ground based 22 GHz radiometers for middle atmospheric water vapor at the Zugspitze in winter 2009	To retrieve a water vapor profile from the spectrum of a ground based 22 GHz radiometer, hours or even days of measurements need to be integrated to achieve a sufficient signal to noise ratio (SNR). MLS profile with measurement site:	criteria.	The direct intercomparison of the three radiometers indicated some suboptimal configurations which could be improved during or after the campaign.
5	WV	Ground-based water vapor raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring	2hr lidar measurement compared to RS92 wvp launched from measurement site. Launch time coincides with 1st hour of lidar measurement. . Calibrations performed against RS92 and GPS + 22-GHz microwave radiometer.	4 coincidence methods and 3 scaling methods tested to give a total of 12 cases being evaluated	The performance figures show that, with the present target of routinely running lidar two hours per night, 4 nights per week, it can achieve measurements with a precision in the UTLS equivalent to that achieved by launching one CFH per month. “Hybrid calibration” method, now used to minimize the cost of launching radiosondes and increase the accuracy and stability of the absolute calibration.
6	WV	Integrated water vapour above Ny Ålesund, Spitsbergen: a multi-sensor intercomparison	Solar FTIR within ±2h of radiosonde launch Lunar FTIR within ±12h of radiosonde launch RAM within ±0.5h of radiosonde launch GPS within ±5min of radiosonde launch SCIAMACHY & GOME Ny Ålesund within ground pixel and time ±2h of radiosonde launch AMSU-B if 5 ground pixels within 50km radius of Ny Ålesund.	Radiosonde, FTIR, RAM and GPS measurements are conducted at the same place, however, only the RAM is measuring the same point every time. Limited information of assessment of coincidence criteria.	Radiosonde humidity taken as the base measurement for comparisons.
7	T	Operational profiling of	Collocated radiosonde launched 1	No assessment of effect or	Radiosonde would be at ~20km in

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
		temperature using ground-based microwave radiometry at Payerne: prospects and challenges	hour before radiometer measurement. Tower based temperature measurements 200m away (2, 10 &30m heights). Radiometer temperature retrievals every 12-13 minutes	development of coincidence criteria.	the 1st hour before radiometer measurements started, however detailed comparisons are made 0 – 4km altitude between the radiometer and radiosonde. The radiometer measurements are biased compared to the radiosondes! Most likely the difference is down to time differences in measurements in addition to the reasons given in the paper.
8	WV	Intercomparison of atmospheric water vapour soundings from the differential absorption lidar (DIAL) and the solar FTIR system on Mt. Zugspitze	Spatial separation 680m. 289m altitude difference. Δt = 22min	The interval length of 22 min yields the smallest standard deviation of differences between DIAL and FTIR IWV values. 22 minute coincidence interval dominated by volume mismatch.	FTIR taken as reference measurement. Using the diurnally changing spatial overlaps between the solar FTIR and the DIAL it was found that a spatial matching on the 100m scale is required to derive <0.1 mm precision for the state-of-the-art IWV sounders. This complements the finding by Sussmann et al (NPL Ref 9) that a temporal matching in the order of 10min or better is required for the same purpose.
9	WV	Technical Note: Harmonized retrieval of column-integrated atmospheric water vapour from the FTIR network – first examples for long-term records and station trends	1 min ≤ Δt ≤3.75 min	A FTIR-FTIR side-by-side intercomparison reveals a strong exponential increase in stdv as a function of increasing temporal mismatch starting at Δt =1 min due to atmospheric water variability. An upper limit of 3.75 min is derived	Integrated water vapour column comparisons between FTIR and radiosonde virtually impossible to the required precision, due to ascent time of radiosonde is greater than the water content time stability of the atmosphere.

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
				to give a statistically significant sample.	
10	T	An investigation of atmospheric temperature profiles in the Australian region using collocated GPS radio occultation and radiosonde data	Three temporal (1, 2 &3h) and three spatial colocations (100, 200 & 300km) tested.	There is no statistical difference detected at the 3 temporal criteria or the 3 spatial criteria.	38 radiosonde launch sites. The RO temperature profiles were interpolated at 16 pressure levels to match the radiosonde profiles.
11	Both	Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE-2009	Multiple transducers on single radiosonde. Small number of collocated Raman and FTIR ground based water vapour measurements. Small timestamp mismatches corrected for by using temperature profile data from each transducer	Collocation irrelevant for radiosonde measurements as collocated on single balloon. FTIR and Raman results not discussed.	"Differences between paired RS92 sondes exceeded the manufacturer-quoted reproducibility limits only 11, 28 and 5% of the time for P, T and RH, respectively. Exclusion of the anomalous ΔT and ΔRH profiles reduces these fractions to 3% (T) and 0% (RH). RS92-iMet P and T differences exceeded their combined measurement uncertainty limits 23 and 31% of the time (anomalies removed), respectively, much more frequently than the 5% expected for 2σ limits."
12	WV	First airborne water vapour lidar measurements in the tropical upper troposphere and mid-latitudes lower stratosphere: accuracy evaluation and intercomparisons with other instruments	Coincidence discussed but no assessment of effect.	Coincidence discussed but no assessment of effect.	
13	WV	Water vapour profiles from SCIAMACHY solar occultation measurements	SCIAMACHY versus ECMWF: 0.72^{o} x 0.045^{o} field of view vs 1.5^{o} x 1.5^{o} latitude / longitude grid, nearest	No assessment of effect or development of coincidence criteria	ECMWF water vapour densities are generally lower than both ACE-FTS and SCIAMACHY data at

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
		derived with an onion peeling approach	measurement to 6h ECMWF forecast. SCIAMACHY versus ACE-FTS: Same day, maximum tangent point distance of 500km.		all heights. SCIAMACHY water vapour data tend to be systematically higher than ACE-FTS.
14	Both	Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology	Detailed discussions on the methodology for dealing with coincidence error in time and space	Detailed discussions on the methodology for dealing with coincidence error in time and space	Also detailed discussions on: χ^2 test smoothing bias determination precision validation
15	WV	Intercomparisons of Stratospheric Water Vapour Sensors: FLASH-B and NOAA/CMDL Frost-Point Hygrometer	Collocated transducers on a single balloon. Small timestamp mismatches corrected for by using temperature profile data from each transducer	Not required to collocated instruments making parallel measurements	The simultaneous measurements show good agreement between both instruments, with a mean deviation of -2.4% ± 3.1% (1sd) for data between 15 and 25 km. Small wet bias od NOAA/CDML. 5s time lag, between NOAA/CMDL and FLASH-B
16	WV	A layer-averaged relative humidity profile retrieval for microwave observations: design and results for the Megha-Tropiques payload	Δt ≤ 45min, Δx ≤ 50 km, over ocean only SAPHIR (10 x 10) km^2 to 14:5 x 22:7) km^2 at edge MADRAS (67.25 x 40) km^2 to (10.1 x 6) km^2 depending on channel	No assessment of effect or development of coincidence criteria	SAPHIR sounder and the MADRAS imager (microwave payload of the Megha-Tropiques platform) Comparison with radiosonde measurements performed during the CINDY/DYNAMO/AMIE (winter 2011–2012) campaign. Mathematical model used to fill in gaps between radiosonde measurements, clear sky only.
17	T	A methodology for the validation of temperature profiles from hyperspectral infrared sounders using GPS radio occultation: Experience with AIRS and	AIRS retrieval field of view at nadir is 45 km with a vertical resolution of 1 to 5km. GPS RO profile occurs within 1 h of AIRS granule time and is within granule latitude / longitude box.	Effect of time coincidence between 0.5h – 3h investigated.	Validation of vertical temperature profile retrievals from infrared and microwave sounders by intercomparison with Global Positioning System (GPS) radio occultation (RO). The

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
		COSMIC	3 methods tested for closest spatial coincidence		bias and RMS error profiles are shown to depend strongly on the vertical averaging applied to the difference profiles but are relatively insensitive to horizontal and temporal mismatch.
18	WV	A multi-instrument comparison of integrated water vapour measurements at a high latitude site	The sampling area of the different datasets ranges from practically point measurements (for the radiosonde) to quite large averaging areas (50 km radius circle for AMSU-B, 83 × 31 km^2 for ERA-Interim). Δt ≤1h used for all measurement types. This ensures that the typical air mass displacement due to the wind is only a few kilometres. Different temporal sampling durations for each measurement method Lower integration altitude limit of 470m for IWV column	A good temporal matching criterion is one that is consistent with these spatial scales. The link between the spatial and temporal scales is the wind speed. There is also a horizontal sampling uncertainty, even for the radiosondes, as they drift horizontally during their ascent. For the ground-based remote sensing instruments (GPS, FTIR, and microwave) there is a similar horizontal sampling uncertainty, since they measure along slant paths of varying elevation angle and horizontal direction.	Radiosondes, Global Positioning System (GPS), ground-based Fourier-transform infrared (FTIR) spectrometer, ground based microwave radiometer, and satellite-based microwave radiometer (AMSU-B). Additionally, ERA-Interim model reanalysis data GPS data set taken as reference
19	WV	A multi-site intercomparison of integrated water vapour observations for climate change analysis	Spatial requirement: ≤50 km IGS, AERONET sun photometer and IGRA radiosonde sites Inclusion of the IGS station in the satellite ground pixel GOME/SCIAMACHY/GOME-2 and AIRS	The smaller the spatial distance, the lower the SD and the higher the correlation coefficient. Limiting the distance between the satellite ground pixel centre and the GPS station is less crucial for	Intercomparison of IWV measurements from satellite devices (in the visible, GOME/SCIAMACHY/GOME-2, and in the thermal infrared, AIRS), in situ measurements (radiosondes) and ground-based instruments (GPS, sun photometer), to assess

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
			Corrections applied to radiosonde data to start IWV column at equal starting pressures Δt ≤10 min for CIMEL Δt ≤30 min for radiosondes, GOME(2)/SCIAMACHY and AIRS	a good GPS–satellite IWV agreement.	their use in water vapour trends analysis. GPS as reference measurement. 28 locations, all Northern Hemisphere. The best IWV agreement was obtained between the ground-based and in situ instruments, especially GPS and radiosondes, both all-weather devices
20	WV	Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP	All Lidars, GPS and radiosonde launch pad are collocated within a few tens of meters. All radiosonde types launched together on a single balloon. No coincidence criteria	No coincidence criteria or discussion on the effect of coincidence.	Two Raman lidars (IGN mobile lidar and OHP NDACC lidar), a stellar spectrometer (SOPHIE), a differential absorption spectrometer (SAOZ), a sun photometer (AERONET), 5 GPS receivers and 4 types of radiosondes, plus 2 10m fixed masts.
21	Boundary layer detection	An evaluation of COSMIC radio occultation data in the lower atmosphere over the Southern Ocean	A direct comparison of temporally and spatially co-located COSMIC profiles	Collocated instruments. No mention of coincidence when comparisons are made with radiosondes.	Two COSMIC products used: “wetPrf”, which is based on 1-D variational analysis from ECMWF, and the “atmPrf” product, which contains the raw measurements from COSMIC
22	Both	An Intercomparison of GPS RO Retrievals with Collocated Analysis and In Situ Observations within Tropical Cyclones	600km radios Δt ≤ 3h Comparisons with radiosondes meeting coincidence criteria	No discussions on the effect of coincidence	four Global Position System (GPS) Radio Occultation (RO) missions: Global Positioning System/Meteorology, CHAallengingMinisatellite Payload, Satellite de Aplicaciones Cientificas-C, and Constellation Observing System for Meteorology, Ionosphere and Climate and Taiwan’s FORMOsa

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
					SATellite Mission #3
23	WV	Analysis of water vapour over Nigeria using radiosonde and satellite data	Monthly average data for radiosondes CM-SAF results averaged into monthly results for comparison with radiosondes	No coincidence criteria or discussion on the effect of coincidence, apart from saying that the 3 radiosonde locations are representative of the whole measurement area.	CM-SAF water vapour products. 3 radiosonde sites
24	Both	Atmospheric Temperature and Water Vapour Retrievals from GOSAT Thermal Infrared Spectra and Initial Validation with Coincident Radiosonde Measurements	Δt ≤ 1h Δx ≤ 110km	No discussions on the effect of coincidence	Comparison of GOSAT with radiosondes over land and oceans
25	WV	Atmospheric water vapour measurements by using ground and satellite based instrumentation and radiosonde	Ground based measurements and radiosonde launch location are collocated. 30 minute Lidar averages coincide with radiosonde launched 10 minutes into measurement period	Effect of balloon drift due to wind discussed. Effect considered to be small due to balloon flight overlapping ground based measurement period and collocated launch position	Intercomparison between ground based lidar, radiosonde and satellite atmospheric water vapour measurements.
26	Both	The NOAA Products Validation System (NPROVS)	Single 'closest' sonde sounding to sateliite within 6 h and 250 km.	No quantification of collocation uncertainty.	'Closest' from a combination of time and distance with a factor to convert t to d based on particular satellite and measurement characteristics. Collocation time starts 45 mins after sonde launch.
27	Both	Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013,	GPSRO profiles within 3 h and 300 km of sonde launch.	No direct quantification of collocation uncertainty but state that 'mean differences are unaffected ….since …errors from the space/time mismatches are basically random'.	GPS satellite occulatation data compared to Vaisala sondes. Sonde data from global network from 2002-2013 and GRUAN data from 2009.

Paper No.	WV / T Profile	Title	Coincidence criteria	Assessment of uncertainty due to coincidence	Extra info
28	Both	Validation of satellite sounder environmental data records : Application to the Cross-track Infrared Microwave Sounder Suite	'Conventional RAOB Matchup' of 6 h and 250 km. 'Dedicated RAOB Matchup' typical launched 15-60min ahead of overpass.	No quantification of mis-match uncertainty, but acknowledge that 'Conventional matchup' requirements could lead to significant uncertainties.	Generalised methodology for validating satellite environmental data records with high-resolution in-situ data, demonstrated for CrIMSS.
29	WV	Water vapour observations up to the lower stratosphere through the Raman lidar during the Maido Lidar Calibration Campaign	Lidar profiles integrated for 1 h from dedicated sonde launch at same location.	No quantification of collocation uncertainties.	Hybrid lidar calibration using system monitoring and GPS-IWV or sonde WV data.
30	Both	Validation of Atmospheric Infrared Sounder temperature and water vapour retrievals with matched radiosonde measurements and forecasts	All data sets available within 3h, 100 km window. Tighter window of 1h, 50 km reduced dataset by 60%.	No quantification of collocation uncertainties.	Combined analysis of AIRS, RAOB, ATOVS and NCEP_GFS data. Collocations also screened by RAOB instrument type.
31	WV	Validation of upper-tropospheric humidity from SAPHIR on board Megha-Tropiques using tropical soundings	A 'restrictive' collocation requirement of 45 min and 50 km.	No quantification of collocation uncertainties.	Comparison between satellite microwave radiometer and sondes. Analysis using conventional and GRUAN-analysed sonde data.

Paper References

1. Continuous quality assessment of atmospheric water vapour measurement techniques: FTIR, Cimel, MFRSR, GPS, and Vaisala RS92, M. Schneider, P. M. Romero, F. Hase, T. Blumenstock, E. Cuevas, and R. Ramos, Atmos. Meas. Tech., 3, 323–338, 2010.
2. Collocating satellite-based radar and radiometer measurements – methodology and usage examples, G. Holl, S. A. Buehler, B. Rydberg, and C. Jim´enez, Atmos. Meas. Tech., 3, 693–708, 2010.
3. An assessment of differences in lower stratospheric temperature records from (A)MSU, radiosondes, and GPS radio occultation, F. Ladstädter, A. K. Steiner, U. Foelsche, L. Haimberger, C. Tavolato, and G. Kirchengast, Atmos. Meas. Tech., 4, 1965–1977, 2011.
4. ARIS-Campaign: intercomparison of three ground based 22 GHz radiometers for middle atmospheric water vapor at the Zugspitze in winter 2009, C. Straub, A. Murk, N. Kämpfer, S. H. W. Golchert, G. Hochschild, K. Hallgren, and P. Hartogh, Atmos. Meas. Tech., 4, 1979–1994, 2011.
5. Ground-based water vapor raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring, T. Leblanc, I. S. McDermid, and T. D. Walsh, Atmos. Meas. Tech., 5, 17–36, 2012.
6. Integrated water vapour above Ny Ålesund, Spitsbergen: a multi-sensor intercomparison, M. Pałm, C. Melsheimer, S. Nöel, S. Heise, J. Notholt, J. Burrows, and O. Schrems, Atmos. Chem. Phys., 10, 1215–1226, 2010.
7. Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges, U. Löhnert and O. Maier, Atmos. Meas. Tech., 5, 1121–1134, 2012.
8. Intercomparison of atmospheric water vapour soundings from the differential absorption lidar (DIAL) and the solar FTIR system on Mt. Zugspitze, H. Vogelmann, R. Sussmann, T. Trickl, and T. Borsdorff, Atmos. Meas. Tech., 4, 835–841, 2011.
9. Technical Note: Harmonized retrieval of column-integrated atmospheric water vapour from the FTIR network – first examples for long-term records and station trends, R. Sussmann, T. Borsdorff, M. Rettinger, C. Camy-Peyret, P. Demoulin, P. Duchatelet, E. Mahieu, and C. Servais, Atmos. Chem. Phys., 9, 8987–8999, 2009.
10. An investigation of atmospheric temperature profiles in the Australian region using collocated GPS radio occultation and radiosonde data, K. Zhang, E. Fu, D. Silcock, Y. Wang, and Y. Kuleshov, Atmos. Meas. Tech., 4, 2087–2092, 2011.
11. Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE-2009, D. F. Hurst, E. G. Hall, A. F. Jordan, L. M. Miloshevich, D. N. Whiteman, T. Leblanc, D. Walsh, H. Vömel, and S. J. Oltmans, Atmos. Meas. Tech., 4, 2777–2793, 2011.
12. First airborne water vapour lidar measurements in the tropical upper troposphere and mid-latitudes lower stratosphere: accuracy evaluation and intercomparisons with other instruments, C. Kiemle, M. Wirth, A. Fix, G. Ehret, U. Schumann, T. Gardiner, C. Schiller, N. Sitnikov, and G. Stiller, Atmos. Chem. Phys., 8, 5245–5261, 2008.
13. Water vapour profiles from SCIAMACHY solar occultation measurements derived with an onion peeling approach, S. Nöel, K. Bramstedt, A. Rozanov, H. Bovensmann, and J. P. Burrows, Atmos. Meas. Tech., 3, 523–535, 2010.
14. Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, T. von Clarmann, Atmos. Chem. Phys., 6, 4311–4320, 2006.
15. Intercomparisons of Stratospheric Water Vapor Sensors: FLASH-B and NOAA/CMDL Frost-Point Hygrometer,H VÖMEL V. YUSHKOV, S. KHAYKIN, L. KORSHUNOV, E. KYRÖ AND R. KIVI, JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY, Vol 24, June 2007, 941-952
16. A layer-averaged relative humidity profile retrieval for microwave observations: design and results for the Megha-Tropiques payload, R. G. Sivira, H. Brogniez, C. Mallet, and Y. Oussar, Atmos. Meas. Tech., 8, 1055–1071, 2015
17. A methodology for the validation of temperature profiles from hyperspectral infrared sounders using GPS radio occultation: Experience with AIRS and COSMIC, Michelle L. Feltz, Robert O. Knuteson, Henry E. Revercomb, and David C. Tobin, Journal of Geophysical Research: Atmospheres, J. Geophys. Res. Atmos., 119, 1680–1691

18. A multi-instrument comparison of integrated water vapour measurements at a high latitude site, S. A. Buehler, S. Östman, C. Melsheimer, G. Holl, S. Eliasson, V. O. John, T. Blumenstock, F. Hase, G. Elgered, U. Raffalski, T. Nasuno, M. Satoh, M. Milz, and J. Mendrok, Atmos. Chem. Phys., 12, 10925–10943, 2012.
19. A multi-site intercomparison of integrated water vapour observations for climate change analysis, R. Van Malderen, H. Brenot, E. Pottiaux, S. Beirle, C. Hermans, M. De Mazière, T. Wagner, H. De Backer, and C. Bruyninx, Atmos. Meas. Tech., 7, 2487–2512, 2014
20. Accuracy assessment of water vapour measurements from in situ and remote sensing techniques during the DEMEVAP 2011 campaign at OHP, O. Bock, P. Bosser, T. Bourcy, L. David, F. Goutail, C. Hoareau, P. Keckhut, D. Legain, A. Pazmino, J. Pelon, K. Pipis, G. Poujol, A. Sarkissian, C. Thom, G. Tournois, and D. Tzanos, Atmos. Meas. Tech., 6, 2777–2802, 2013
21. An evaluation of COSMIC radio occultation data in the lower atmosphere over the Southern Ocean, L. B. Hande, S. T. Siems, M. J. Manton, and D. H. Lenschow, Atmos. Meas. Tech., 8, 97–107, 2015
22. An Intercomparison of GPS RO Retrievals with Collocated Analysis and In Situ Observations within Tropical Cyclones, Henry R.Winterbottom and Qingnong Xiao, Advances in Meteorology, Volume 2010, Article ID 715749, 10 pages.
23. Analysis of water vapour over Nigeria using radiosonde and satellite data, B. Adeyemi & S Joerg, Journal of applied meteorology and climatology, Vol 51, Oct 2012, 1855-1866
24. Atmospheric Temperature and Water Vapour Retrievals from GOSAT Thermal Infrared Spectra and Initial Validation with Coincident Radiosonde Measurements, Hirofumi Ohyama, Shuji Kawakami, Kei Shiomi, Isamu Morino, and Osamu Uchino, SOLA, 2013, Vol. 9, 143−147
25. Atmospheric water vapour measurements by using ground and satellite based instrumentation and radiosonde, V. Cuomo, P.F. Ambrico, A. Amodeo, A. Boselli, P. Di Girolamo, V. Lanorte, M. Pandolfi, G. Pappalardo*, N. Pergola*, C. Pietrapertosa, F. Romano, V. Tramutoli, N. Spinelli, Proceedings of SPIE - The International Society, Vol 4070, 2000, 73-80.
26. The NOAA Products Validation System (NPROVS), T Reale, B Sun, F Tilley, N Pettey, Journal of Atmospheric and Oceanic Technology, Vol 29, May 2012, 629-644.
27. Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013, F. Ladstädter, A. K. Steiner, M. Schwärz, and G. Kirchengast, Atmos. Meas. Tech., 8, 1819–1834, 2015
28. Validation of satellite sounder environmental data records : Application to the Cross-track Infrared Microwave Sounder Suite, N Nalli, C Barnet, A Reale, D. Tobin, A Gambacorta et al, Journal of Geophysical Research : Atmospheres, Vol 188, 13,628-13,643, 2013
29. Water vapour observations up to the lower stratosphere through the Raman lidar during the Maido Lidar Calibration Campaign, D. Dionisi, P Keckhut, Y Courcoux, A Hauchecorne et al, Atmos. Meas. Tech., 9, 1425-1445, 2015
30. Validation of Atmospheric Infrared Sounder temperature and water vapour retrievals with matched radiosonde measurements and forecasts, M Divakarla, C Barnet, M Goldberg, L McMillin, E maddy, W Wolf, L Zhou, X Liu, Journal of Geophysical Research, Vol. 111, D09S15, 2006
31. Validation of upper-tropospheric humidity from SAPHIR on board Megha-Tropiques using tropical soundings, H Brogniez, G Clain, Journal of Applied Meteorology and Climate, Vol. 54, 896-908, 2015

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Initial input from WP3 to the gap analysis and impacts document, Annex 2:

Literature review on atmospheric composition data validation

A Horizon 2020 project; Grant agreement: 640276

Date: 30 June 2015

Lead Beneficiary: BIRA

Nature: R

Dissemination level: PU

Work-package	WP 3 (Comparison error budget closure – Quantifying metrology related uncertainties of data comparisons)
Deliverable	D3.1
Title	Initial input from WP3 to the gap analysis and impacts document Annex 2: Literature review on atmospheric composition data validation
Nature	R
Dissemination	PU
Lead Beneficiary	Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
Date	30 june 2015
Status	Final, version 1
Editor	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB)
Authors	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB)
Contacts	tijl.verhoelst@aeronomie.be, j-c.lambert@aeronomie.be
URL	http://www.gaia-clim.eu/

This document has been produced in the context of the GAIA-CLIM project. The research leading to these results has received funding from the European Union's Horizon 2020 Programme under grant agreement n° 640276. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

Paper no.	ECV	Short reference	Instruments	Coincidence criteria	Assessment of uncertainty due to metrological differences	Extra info
1	CO2, CH4	Buchwitz et al. 2015	SCIAMACHY and TANSO vs. TCCON	500km, 2h, following Dils et al. 2014	None in terms of coincidence criteria or smoothing differences. Reference to Dils et al. 2014.	Systematic error quantified as inter-station bias stdev. Random error from stdev on differences without accounting for natural variability.
2	CO2, CH4	Dils et al. 2014	SCIAMACHY and TANSO vs. TCCON	500km, 2h	Other co-location criteria explored, but tighter means too few co-locations, larger implies greater variability. Co-location noise remains below user requirements.	Atmospheric variability at the spatio-temporal scale of the coincidences quantified from FTS intercomparisons. Suggestion of dynamic co-location criteria,cfr. Guerlet et al. 2013.
3	CO2	Guerlet et al. 2013	TANSO vs. TCCON	5 degrees, 2h OR TM5 model representative area (limited by 7.5x22.5 degrees), 2h, similar to Oshchepkov et al. 2012	Results based on both co-location criteria are compared, but no quantification of co-location error.	Refers to Oshchepkov et al. 2012, and to Wunch et al., 2011 for model-based coincidence criteria.
4	CO2	Oshchepkov et al. 2012	GOSAT vs. TCCON	NIES TM representative area, limited by 15x45 degrees, 1h	Coincidence criteria designed to limit co-location error to 1ppm, which is the retrieval error.	The NIES TM is also compared to the GOSAT observations as monthly means. Refers to Wunch et al. 2011.

Paper no.	ECV	Short reference	Instruments	Coincidence criteria	Assessment of uncertainty due to metrological differences	Extra info
5	CO2	Wunch et al. 2011	GOSAT vs. TCCON	SH: no constraints NH: 10 days, 10x30 degrees, and +/- 2K T_700 potential temperature difference.	Impact of choice of coincidence criteria is analysed in detail.	Southern hemisphere used as a whole for bias determination (almost no natural variability).
6	CO2, CH4	Schneising et al. 2012	SCIAMACHY vs. TCCON and CarbonTracker (model)	Monthly means within 500km	No analysis at all.	
7	CH4, N2O	Payan et al. 2009	MIPAS profiles vs. various sub-orbital instruments	Default: 300km,3h except: IBEX: back trajectories; MIPAS-B: 460km; BONBON, SPIRALE: 5-day back trajectory;ASUR: 1000km, 12h; FTIR: 300km/3h or 400km/4h	No assessment of co-location criterium impact done. No attempt to quantify metrological errors.	Large set of different correlative data.
8	CH4	De Mazière et al. 2008	ACE-FTS CH4 profiles vs. FTIR and SPIRALE	FTIR: 1000km,24h except polar: PV diff. limit and 500km,12h SPIRALE: 13h,413km (1 co-location)	No quantification, but atmospheric variability is suspected to play a role as observed stdev of differences is larger than meas. uncertainty.	
9	NO2 (total and tropo column)	Valks et al. 2011	GOME-2 vs. MAXDOAS at OHP	100km, MAXDOAS interpolated in time to GOME-2 overpass time + averaging to monthly means.	No analysis of sampling and smoothing errors. Refers to Lambert et al. O3MSAF tech note on NO2 end-to-end validation protocol (2008)	Refers to the DANDELIONS and CINDI campaigns for validation strategies.

Paper no.	ECV	Short reference	Instruments	Coincidence criteria	Assessment of uncertainty due to metrological differences	Extra info
10	NO2	Celarier et al. 2008	OMI vs. ZSL-DOAS, MAXDOAS and direct sun (Pandora, FTUVS)	Ground DOAS sit e within OMI pixel. Brewers: daily means required for robust statistics.	Empirical determination of atmospheric variability and inhomogeneity from MAXDOAS measurements. Comparisons are found to be affected by differences in smoothing .	Diurnal cycle correction based on SLIMCAT model.
11	O3, CH2, NO2,…	Von Clarmann et al. 2009	MIPAS	NA	Determination of the horizontal resolution of MIPAS limb retrievals from horizontal averaging kernels.	Order of magnitude is a few 100km.
12	O3 (total column)	Balis et al. 2007	GOME vs. Brewers, Dobsons and SAOZ instruments	Brewers & Dobsons: 150km, 3h SAOZ: intersection between GOME footprint and estimated SAOZ airmass	Some discussion on smoothing and sampling difference errors, leading to optimized co-location criteria, but no quantification of the errors in the comparisons.	SAOZ observation operator presented.
13	O3	Cortesi et al. 2007	MIPAS vs. O3 sondes, lidars and MWR	300km, 3h or if necessary 500km, 10h	Detailed error budget decomposition using an OSSE approach.	Sampling and smoothing difference errors (both horizontal and vertical) are estimated.
14	O3	Sparling et al. 2006	Aircraft	NA	Detailed assessment of small-scale variability of the ozone field in the UT/LS. Recommendations on co-location criteria as a function of measurement uncertainty are formulated.	Some discussion on the effect of different resolutions.
15	O3, NO2	Lambert et al. 2011	Limb and nadir sounders, ground-	NA	Detailed assessment of the horizontal smoothing	Includes simulation of smoothing errors by

Paper no.	ECV	Short reference	Instruments	Coincidence criteria	Assessment of uncertainty due to metrological differences	Extra info
			based direct-sun and ZSL-DOAS instruments		properties of different types of satellite and ground-based measurement systems, including the construction of pragmatic observation operators.	application of the observation operators on modelled fields, i.e. an OSSE.
16	O3	Hubert et al. 2015	14 limb sounders vs. O3 sondes and lidars	Closest within 500km, 6h or 12h	No investigations done, but values from Cortesi et al. 2007 are used for qualitative interpretation.	This is actually a drift and mutual consistency analysis.
17	O3 total column	Verhoelst et al. 2015	GOME, SCIAMACHY, GOME-2A vs. Brewers, Dobsons, and ZSL-DOAS	Brewers and Dobsons: 150km, 3h; ZSL-DOAS: airmass intersection as in Balis et al. 2007	Full quantification of sampling and smoothing errors using an OSSE (OSSSMOSE) based on MERRA and IFS-MOZART fields and pragmatic observation operators.	Error budget closure achieved for most instrument combinations.

1. “The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO2 and CH4 global data sets”, Buchwitz et al., RSE v162, p344, 2015

2. “The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO2 and CH4 retrieval algorithm products with measurements from the TCCON”, Dils et al. AMT v7, 2014

3. “Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements”, Guerlet et al., JGR v118, 2013

4. “Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space: Validation of PPDF-based CO2 retrievals from GOSAT”, Oshchepkov et al., JGR v117, 2012

5. “A method for evaluating bias in global measurements of CO2 total columns from space”, Wunch et al. ACP v11, 2011

6. “Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results”, Schneising et al. ACP v12, 2012

7. "Validation of version-4.61 methane and nitrous oxide observed by MIPAS", Payan et al. ACP v9, 2009

8. "Validation of ACE-FTS v2.2 methane profiles from the upper troposphere to the lower mesosphere", De Mazière et al., ACP v8, 2008

9. "Operational total and tropospheric NO2 column retrieval for GOME-2", Valks et al., AMT v4, 2011

10. "Validation of Ozone Monitoring Instrument nitrogen dioxide columns", Celarier et al., JGR v113, 2008

11. "The horizontal resolution of MIPAS", von Clarmann et al., AMT v2, 2009

12. "Ten years of GOME/ERS2 total ozone data – The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8", Balis et al. JGR v112, 2007

13. "Geophysical validation of MIPAS-ENVISAT operational ozone data", Cortesi et al., ACP v7, 2007

14. "Estimating the impact of small-scale variability in satellite measurement validation", Sparling et al., JGR v111, 2006

15. "Multi-dimensional characterisation of remotely sensed data", Lambert et al., EC FP6 GEOmon Technical notes, 2011

16. "Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records", Hubert et al., accepted for publication in AMTD, 2015

17. "Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons", Verhoelst et al., accepted for publication in AMTD, 2015

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Initial input from WP3 to the gap analysis and impacts document, Annex 3:

Literature review on T, q, and aerosol data validation

A Horizon 2020 project; Grant agreement: 640276

Date: 30 June 2015

Lead Beneficiary: BIRA

Nature: R

Dissemination level: PU

Work-package	WP 3 (Comparison error budget closure – Quantifying metrology related uncertainties of data comparisons)
Deliverable	D3.1
Title	Initial input from WP3 to the gap analysis and impacts document, Annex 3: literature review on T, q, and aerosol data validation
Nature	R
Dissemination	PU
Lead Beneficiary	Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
Date	30 june 2015
Status	Final, version 1
Editors	T. Verhoelst (BIRA-IASB), J.-C. Lambert (BIRA-IASB)
Authors	P. Kolmonen (FMI), G. de Leeuw (FMI), F. Madonna (CNR), L. Mona (CNR)
Contacts	madonna@imaa.cnr.it, gerrit.leeuw@fmi.fi
URL	http://www.gaia-clim.eu/

This document has been produced in the context of the GAIA-CLIM project. The research leading to these results has received funding from the European Union's Horizon 2020 Programme under grant agreement n° 640276. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

1. Validation papers on aerosol and meteorological variables, with the identification of missing comparison error assessment.

Instrumental contribution to the total uncertainty budget (random and systematic uncertainties) has been investigated for various sensors, e.g. Raman lidars (e.g. Whiteman et al., 2001), radiosondes (Immler et al., 2010) or weather radars (e.g. O'Connor et al., 2005). For other sensors several studies are ongoing, like for the GPS/GNSS in the frame of GRUAN, or like the MWR in the frame of TOPROF EU cost action and within GAIA-CLIM WP2. On the other hand, one of the critical contributions to the uncertainty budget is related to the collocation mismatch in space and time among pairs of sensors. Although these different measurements (of the same atmospheric parameter) are often assumed to be nominally collocated, there is a real physical separation between their actual measurement locations and timing. This assumption is generally true for ground-based observation, but when one is ground-based and another satellite-based a certain spatial and temporal mismatch needs to be allowed for given the very short overpass time of a satellite.

The satellite validation community considers, as a priority, the availability of robust collocation criteria that would increase the matches by a significant amount at an affordable cost due to data synergy. Appropriate collocation criteria are strongly required to combine different measurements, to potentially reduce the overall uncertainty in the atmospheric column or profile measurement.

In the following, a list of peer-reviewed papers is reported dealing with the validation of aerosol, temperature and humidity satellite products using ground-based or other satellite based measurements, or dealing with approaches aiming at the quantification of the co-location uncertainty (in time and space) comparing ground based measurements only.

Each paper is briefly summarized including the minimum information about the the identification of missing comparison error assessment.

1.1 Temperature and humidity: Reference of methods and definition

Kitchen, M., 1989. Representativeness errors for radiosonde observations. Q. J. R. Meteorol. Soc. 115, 673-700.

In this paper, the total observational error associated with the use of radiosonde is decomposed in terms of measurements errors, effect of the atmospheric variability, the representativeness error and co-location error. The last is investigated in terms of the difference associated with the separation in time of the observations, with the horizontal separation between the observations. The latter also includes the cases of the comparison betweeen points observations and the case of point and area-averaged observations.
The paper concludes that in the design of satellite/radiosonde collocation experiments, the errors associated with the comparison can be properly assessed only from a knowledge of atmospheric variability on the appropriate scales. For typical collocation criteria and modern radiosonde designs, the radiosonde measurement errors make only a small contribution to the total error of comparison.

McGrath, R., Tido Semmler, Conor Sweeney, and Shiyu Wang, 2006: Impact of Balloon Drift Errors in Radiosonde Data on Climate Statistics. J. Climate, 19, 3430–3442. doi: http://dx.doi.org/10.1175/JCLI3804.1

In this paper, simulated radiosonde observations have been used to evaluate the impact of ignoring the spatial and temporal spread of data inherent in radiosonde observations. Simulations were generated for the period 1960–2002 for 87 stations from the CARDS archive using ERA-40 data to derive ascent profiles that mimic actual ascents and also reflect local practices. A quantification in terms of RMS deviation is provided. Measurements uncertainty are completely neglected and only the difference between the profiles is considered.

Fassò, A., Ignaccolo, R., Madonna, F., Demoz, B. B., and Franco-Villoria, M.: Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles, Atmos. Meas. Tech., 7, 1803-1816, doi:10.5194/amt-7-1803-2014, 2014.

In this paper, a statistical modelling approach capable of explaining the relationship between collocation uncertainty and a set of environmental factors, height and distance between imperfectly collocated trajectories is presented. The new statistical approach is based on the heteroskedastic functional regression (HFR) model which extends the standard functional regression approach and allows a natural definition of uncertainty profiles. Along this line, a five-fold decomposition of the total collocation uncertainty is proposed, giving both a profile budget and an integrated column budget. HFR is a data-driven approach valid for any atmospheric parameter, which can be assumed smooth.
In the paper the decomposition of the uncertainty budget is obtained using the statistical evaluation of a large dataset of radiosonde pair but the radiosonde measurements errors are not considered.

McDonald, A. J., B. Tan, and X. Chu (2010), Role of gravity waves in the spatial and temporal variability of stratospheric temperature measured by COSMIC/FORMOSAT-3 and Rayleigh lidar observations, J. Geophys. Res., 115, D19128, doi:10.1029/2009JD013658.

This study utilizes COSMIC satellite and lidar observations to examine the spatial and temporal variability of stratospheric temperature at a number of scales. The geographic variation of the RMS temperature difference between pairs of COSMIC profiles shows a strong correspondence to previous climatologies of gravity wave activity. In addition, the co-location uncertainty is evaluated using second order structure functions directly related to the horizontal wave number power spectrum. Examination of temperature variability as a function of spatial and temporal separation indicates that gravity wave activity dominates stratospheric temperature variability, and this has impacts on validation study site selection.

Sofieva, V. F., F. Dalaudier, R. Kivi, and E. Kyro (2008), On the variability of temperature profiles in the stratosphere: Implications for validation, Geophys. Res. Lett., 35, L23808, doi:10.1029/2008GL035539.

In this paper, the variability of the small-scale structure of temperature fields in the stratosphere using temperature profiles from radio-soundings is analyzed with a small time difference between sonde launches. The set of the collocated temperature profiles allows obtaining experimental estimates of the horizontal structure function of temperature fluctuations. Implications of these results for validation of high-resolution profiles are discussed. Radiosonde measurements errors are not considered.

Calbet, X., Kivi, R., Tjemkes, S., Montagner, F., and Stuhlmann, R.: Matching radiative transfer models and radiosonde data from the EPS/Metop Sodankylä campaign to IASI measurements, Atmos. Meas. Tech., 4, 1177-1189, doi:10.5194/amt-4-1177-2011, 2011.

The study consider the closest IASI field of view to the Sodankyla observatory, selected for comparison purposes, which has a footprint that usually encloses the observatory location. The study shows that the radiosondes do not drift very far away from the launch location. The paper concludes that on the experience gathered, spatial collocation does not seem to play a big role in the radiance matching, but temporal co-location and time interpolation are critical to achieve these results. Radiosonde measurement errors are not considered.

Frehlich, R. and Sharman, R.: Estimates of turbulence from numerical weather prediction model output with applications to turbulence diagnosis and data assimilation, Mon. Weather Rev., 132, 2308-2324, 2004.

The paper deal with the estimates of small-scale turbulence from numerical model output are produced from local estimates of the spatial structure functions of model variables such as the velocity and temperature. It is determined that the total observation error for typical rawinsonde measurements of velocity are dominated by the sampling error or ‘‘error of representativeness’’ resulting from the effects of small scale turbulence.
Rawinsonde measurement uncertainty budget is not considered but somehow determined using a few assumptions.

Mears, C. A., J. Wang, D. Smith, and F. J. Wentz (2015), Intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and ground-based GPS instruments, J. Geophys. Res. Atmos., 120, doi:10.1002/2014JD022694.

The paper deal with the intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and a network of ground-based GPS instruments. In the assembly of co-located datasets, missing data due the retrieval unavailability force the use of measurements farther from the station, thereby increasing the differences between satellite and GPS measurements. We can reduce these differences by accounting for the local gradient in IWV in the region surrounding the station. Local satellite measurements are fitted using a bilinear fit to obtain the best fit plane through the data points.

Pougatchev, N., August, T., Calbet, X., Hultberg, T., Oduleye, O., Schlüssel, P., Stiller, B., Germain, K. St., and Bingham, G.: IASI temperature and water vapor retrievals - error assessment and validation, Atmos. Chem. Phys., 9, 6453-6458, doi:10.5194/acp-9-6453-2009, 2009.

The study deals with the validation of IASI temperature and humidity profiles. It introduces an approach developed by Pougatchev (2008), which does not require other measurements besides the correlative data per se. The best estimate of the true atmospheric state and corresponding nominal satellite measurement are provided by the linear statistical Validation Assessment Model (VAM). For this particular study, the VAM uses correlative radiosonde profiles as input and returns the optimal estimate of the nominal IASI retrieval by utilizing IASI averaging kernels and statistical characteristics of the ensembles of the reference radiosondes. Temporal non-coincidence errors and associated correlative matrices are derived from statistical analysis of the radiosonde profiles (Pougatchev, 2008). Spatial non-coincidence error and retrieval noise are inferred from actual IASI retrievals.
Radiosonde measurement errors are not considered.

Seidel, D. J., Sun, B., Pettey, M., and Reale, A.: Global radiosonde balloon drift statistics; J. Geophys. Res., 116, 1-8, doi:10.1029/2010JD014891, 2011.

This study deal with the quantification of the spatial sampling for in situ GRUAN observations using balloon drift statistics (e.g. probability distribution functions, percentile values). Radiosonde balloon drift distance and elapsed time
climatologies have been developed from global data from 419 stations for the period July 2007 through June 2009.
Radiosonde measurement errors are not considered.

Sun, B., Reale, A., Seidel, D. J., and Hunt, D. C.: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics, J. Geophys. Res., 115, D23104, doi:10.1029/2010JD014457, 2010.

The study aims to quantify the error characteristics of 12 radiosonde types flown in the global operational network, as a function of height and for both day and nighttime observations, for each of the three variables. Moreover, the effects of imperfect temporal and spatial collocation on the radiosonde ‐ COSMIC differences is determined for application to the general problem of satellite calibration and validation using in situ sounding data. A quantification of the co-location uncertainty is provided in terms of RMS and standard deviation with respect to time and spatial mismatches.

Tobin, D. C., Revercomb, H. E., Knuteson, R. O., Lesht, B. M., Strow, L. L., Hannon, S. E., Feltz, W. F., Moy, L. A., Fetzer, E. J., and Cress, T. S.: Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation, J. Geophys. Res., 111, D09S14, doi:10.1029/2005JD006103, 2006.

The paper deal with the elaboration of an atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation. It concludes that spatial collocation mismatch does not seem to play a big role in the radiance matching, due to the large footprint characterizing these measurements. On the contrary, temporal collocation and time interpolation are critical to achieving these results due to the related vertical thermodynamic factors (Tobin et al., 2006). Conclusion are similar to Calbet et al.

1.2 Aerosol: Reference of methods and definition

By nature, ground-based in situ and remote sensing aerosol measurements characterize a particular volume of air at a particular point in time and space. As a consequence, measurements can only characterize the aerosol content at the particular location where the sampling takes place as a function of time. Consequently, the choice of location and whether and when measurements are performed in order to be appropriate to answer particular questions is critical. This leads to the question of representativeness of observations.

In order to link data from different stations and/or instruments different aspects need to be addressed:

- detected or retrieved properties should be compliant to preset accuracy requirements
- local aspects (like pollution or orography) should be identified to avoid the application of the local data to its surrounding region.
- quantitative measures based on objective applications is needed to assess data quality and the regional influence of any monitoring site

The comparison between independent ground-based and satellite observations of the same observable could be particularly suited for this kind of study.

One approach used for regional assessment was based on monthly and seasonal evaluation of average relative deviance and average relative bias. The rate of exceedance for error and bias give an indication of the representativeness of sampling.

Another suitable approach was proposed by Nappo et al. 1982 for the view of point-to-area representativeness. This method defines a measurement q as representative for the average Q in a larger volume if the probability that the squared difference between point and volume is smaller than a certain threshold δ for more than 90% of the time. The acceptable threshold d has to be defined for each investigated case:

- it should not be smaller than the uncertainty of the measurement
- It should be set in according to the specific problem to be addressed (for example in terms of AOD the acceptance could be higher for air quality issues and lower (more accuracy needed) for climate research problems

Comparison with models could be also very useful in this sense but shortcomings of both methods should be taken into account. Where the in situ measurements have a limited validity as discussed above, satellite measurements may be subject to retrieval errors while models depend on the accuracy of the descriptions of the various relevant processes and the availability of accurate input data, in particular emissions of aerosol particles and aerosol precursors.

Attempts to investigate the representativeness of aerosol measurement are available in literature [Kaufmann et al., 2000; Anderson et al., J2003; Omar et al 2005; Kahn et al2007; Holzer-Popp et al 2008], but a precise and quantitative methodology is still missing. Moreover, one aspect was overlooked in this context: the vertical variability of the aerosol. This was mainly due to the un-availability of long-term and distributed measurements of aerosol properties vertical profiles. The relatively recent availability of long-term distributed lidar measurements of aerosol vertical profiles performed by quality assured ground-based networks such as EARLINET (European Aerosol research Lidar NETwork; ref) or the satellite record of lidar observations from CALIOP (ref) and the future Earthcare mission could fill this gap.

Comparing EARLINET and CALIPSO aerosol vertical profiles, a first representativeness study for aerosol vertical profiles was carried out [Pappalardo et al., 2010]. Comparing aerosol backscatter at 532nm as most abundant parameter available from the 2 platforms, data are compared for altitude layers where the 2 instruments report backscatter layers. The comparisons of CALIPSO and EARLINET backscatter measurements at 532 nm are performed a) for a fixed maximum horizontal distance of 100 km and different time shifts, and b) for a fixed time shift of 10 min and different horizontal distances. Mean differences profiles are investigated for discovering potential local effects and disregarding them in the successive steps. Then, the count distributions of CALIPSO and EARLINET backscatter-coefficient measurements are compared for classes corresponding to different time shifts (case a) or different space distance(b). Finally, the correlation coefficient between CALIPSO and EARLINET backscatter counts distributions is investigated as a function of the time (a) or of the distance (b). Also the correlation study could be performed at different altitude ranges.

Validation of AOD

Satellite retrieved aerosol optical depth (AOD) is usually validated with independent ground based spectral AOD measurements available from sun photometers (direct sun measurements, combination with angular scattering measurements provides information on additional aerosol properties such as particle size distribution and refractive index). The most widely used ground-based AOD measurements are those from the AERONET (AErosol RObotic NETwork) (Holben et al., 1998) over land and coastal regions and the Marine Aerosol Network (MAN) over ocean (Smirnov et al., 2009). Other networks are suitable networks are, e.g. the WMP-GAW PFR (Precision Filter Radiometer) network providing only AOD. AERONET uses CIMEL sun photometers of several types which may be different as regards the availability of polarization measurements and the number of wavebands, but all provide a set of basic parameters. MAN is based on ships of opportunity and scientific cruises where hand-held sun photometers are used to provide AOD. Both networks follow a well-defined calibration protocol. The power of the AERONET and MAN is the availability and quality assurance of the data which can be downloaded freely from the AERONET site. The AERONET direct algorithm provides AOD with an accuracy of about 0.01-0.02 (Eck et al., 1999) or 5-10 % for AOD smaller than 0.2. The inversion algorithm employs the almucantar measurements to determine various atmospheric aerosol properties such as single scattering albedo and size distribution. The different levels of the AERONET data describe the quality of the measured aerosol property. Level 1.0 data is the raw output that an AERONET instrument measures every 15 minutes. The data at level 1.5 is filtered automatically for clouds and instrument errors. The level 2.0 data which is recommended for use in validation has been checked by a human observer.

There are several issues to be taken into account when satellite retrieved aerosol properties are validated by comparison with the AERONET data. The most important ones are the spatial and temporal collocations. The spatial collocation has been discussed for MODIS validation by Ichoku et al. (2002). They argue that it is unlikely that strict satellite point measurements would give a representative validation. In principle, it would be ideal to compare statistical values between the AOD from satellite and ground based retrieval. For the satellite AOD this means that an average AOD, together with standard deviation, for an area around an AERONET site should be determined. This gives also an additional benefit since there might not be a retrieved AOD directly on top of an AERONET site because of clouds but the surrounding area can be clear providing AOD for the AERONET-satellite comparison. It must be noted that the averaging will decrease the effect of random errors in the retrieved AOD (Breon et al., 2011). The main question in the averaging is how large the area around an AERONET should be. On one hand, too big an area for the satellite AOD could lead to poor presentation of the AERONET site due to large variance in aerosol conditions. On the other hand, too small area does not provide enough information for the computation of the standard deviation. Ichiko et al. varied the size of the area from 30 x 30 km^2 to 70 x 70 km^2 at several AERONET sites and did not generally see large differences in the averaged AOD or the standard deviation. They ended up using a 50 x 50 km^2 area for reasonable statistics for variance determination as the ground pixel size for which the MODIS AOD is provided is 10 x 10 km^2 (in the most recent MODIS Collection6 a 3x3 km^2 product will be available; Remer et al. 2013). Breon et al. (2011) chose a 50 km radius around an AERONET site claiming that this area gives the best comparison between satellite and AERONET AOD.

The ground pixel size effects the spatial collocation in validation. The most common satellite AOD products have level 2 and level 3 data. The level 2 data has a ground pixel size of 10 x 10 km^2 for MODIS (e.g. Ichoku et al, 2002), and for a number of ATSR2/AATSR algorithms (de Leeuw et al., 2015). The level 3 data has most often 1 x 1 degrees area. The above mentioned averaging is suitable for the level 2 data but not for the level 3 data. For level 3 a single area containing an AERONET site should be used as even the neighboring areas do not most certainly represent the aerosol conditions at the site because they are too far from it. The variance in the level 3 AOD is usually included in the data along with possibly some other statistical values.

The temporal collocation concerns the AERONET AOD when compared with the satellite AOD. The exact collocation between the satellite and AERONET data is here very difficult to find. The time window for collocation has to be carefully chosen because too big a window will lead to non-representative collocation when large temporal changes in aerosol conditions occur during the time window. Breon et al. (2011) found out that if the window is more than 30 minutes around the satellite retrieval the comparison between the AERONET and satellite AOD starts to worsen. They do not recommend longer time than the mentioned 30 minutes.

Sometimes the daily average values of AERONET AOD are used for the satellite validation. The diurnal variability can be from 10 % to 40 % (Smirnov et al., 2002) and must be taken into account for these kind of comparisons.

AOD uncertainties

In the satellite aerosol retrieval community two measures for the uncertainty are generally used. The first one is the ‘envelope method’ based on the comparison to AERONET AOD and determine the standard deviation from a least squares fit. These methods are applied to, e.g. the MODIS validation (Remer et al. 2005; Levy et al 2013). Also the Aerosol_cci uses this method as a metric for the validation of L2 and L3 aerosol products while in addition a scoring method is used based on the combination of spatial and temporal correlation of satellite and AERONET retrieved AOD (de Leeuw et al. 2015). The second, more rigorous method is the introduction of a per pixel uncertainty characterization based on rigorous error propagation as discussed in the ESA Climate Change Initiative (CCI) (Hollman et al. 2013) with specific application to aerosol retrieval products developed as part of the Aerosol_cci project (Holzer-Popp et al, 2013) during workshop and validation exercises (see http://www.esa-aerosol-cci.org/ for reports on the subject and Povey and Grainger 2015 for an introduction) and discussion in the AERO-SAT meetings (http://www.aero-sat.org/).

Aerosol References of methods and definition

- Anderson et al., J. Geophys. Res., 108 , 8647, 2003;
- Henne, et al., Atm. Chem. Phys. 8 , 3119-3139, 2008;
- Holzer-Popp et al., Atmos. Chem. Phys., 8, 7651–7672, 2008
- Larssen, et al., European Environment Agency, Copenhagen, Denmark 56 pp., 1999
- Nappo, et al., Bull. Amer. Meteorol. Soc., 63, 761-764.1982
- Povey and Granger, in preparation for submission (2015)

Aerosol References on validation studies

- Breon F.-M., Vermeulen, A., and Descloitres, J., 2011, An evaluation of satellite aerosol products against sunphotometer measurements, Remote Sens. Environ., 115, 3102-3111.
- de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W., Descloitres, J., Grainger, R.G., Griesfeller, J. et al., 2013, Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis, Remote Sensing of Environment 162 (2015) 295–315. DOI information: 10.1016/j.rse.2013.04.023.
- Holzer-Popp, T., de Leeuw, G., Martynenko, D., Kluser, L., Bevan, S., Davies, W., Ducos, F. et al., 2013, Aerosol retrieval experiments in the ESA Aerosol cci project, Atmos. Meas. Tech., 6, 1919-1957.
- Kahn, et al, J. Geophys. Res.110, D10S04, doi:10.1029/2004JD004706, 2005.
- Kahn et al., J. Geophys. Res. 112 (2007), D18205.
- Kaufmann et al., Geophys. Res. Lett. 27 (2000), 3861-3864.
- Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, doi:10.5194/amt-6-2989-2013, 2013.
- Omar et al., J. Geophys. Res. 110 (2005), D10S14
- Pappalardo, G., et al. J. Geophys. Res., 115, D00H19, doi:10.1029/2009JD012147, 2010.
- Remer, L. A., Kaufman, Y. J., Tanr´e, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R. R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS aerosol algorithm, products and validation, J. Atmos. Sci., 62, 947–973, 2005.
- Sayer, A. M., et al,, J. Geophys. Res. Atmos., 118, 7864–7872, doi:10.1002/jgrd.50600., 2013
- Schuster et al., Atmos. Chem. Phys., 12, 7431–7452,, 2012

Aerosol References on reference networks:

- Eck, T. F., et al. (2009), Optical properties of boreal region biomass burning aerosols in central Alaska and seasonal variation of aerosol optical depth at an Arctic coastal site, J. Geophys. Res., 114, D11201, doi:10.1029/2008JD010870,2009.
- Holben B.N., T.F.Eck, I.Slutsker, D.Tanre, J.P.Buis, A.Setzer, E.Vermote, J.A.Reagan, Y.Kaufman, T.Nakajima, F.Lavenu, I.Jankowiak, and A.Smirnov, 1998, AERONET - A federated instrument network and data archive for aerosol characterization, Rem. Sens. Environ., 66, 1-16.
- Smirnov, A., Holben, B.N., Slutsker, I., Giles, D.M., McClain, C.R., Eck, T., Sakerin, S.M. et al. 2009. Maritime Aerosol Network as a component of Aerosol Robotic Network. *J. Geophys. Res.* 114 D6.

Aerosol General references

- Smirnov, A., B. N. Holben, T. F. Eck, I. Slutsker, B. Chatenet, and R. T. Pinker, 2002, Diurnal variability of aerosol optical depth observed at AERONET (Aerosol Robotic Network) sites, Geophys. Res. Lett. , 29 (23), 2115, doi:10.1029/2002GL016305.
- Hollmann, R., C. Merchant, R. Saunders, C. Downy, M. Buchwitz, A. Cazenave, E. Chuvieco, P. Defourny, G. de Leeuw, R. Forsberg, T. Holzer-Popp, F. Paul, S. Sandven, S., Sathyendranath, M. van Roozendael, W. Wagner (2013).The ESA climate change initiative: satellite data records for essential climate variables. Bulletin of the American Meteorological Society Oct. 2013, 1541-1552. ; e-View: doi: http://dx.doi.org/10.1175/BAMS-D-11-00254.1.
- Remer et al. Atmos. Meas. Tech., 6, 1829–1844, 2013