GAIA-CLIM Report / Deliverable D3.8

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Final review of and update to the GAID from the perspective of WP3

A Horizon 2020 project; Grant agreement: 640276

Date: 18/12/2017

Beneficiary: BIRA

Nature: R

Dissemination level:

Work Package	WP 3
Deliverable	D3.8
Title	Final review of and update to the GAID from the perspective of WP3
Nature	R
Dissemination	Public
Beneficiary	BIRA-IASB
Date	18 December 2017
Status	Final
Authors	Tijl Verhoelst, Jean-Christopher Lambert, Peter Thorne
Contacts	tijl.verhoelst@aeronomie.be, jean-christopher.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.

Table of Contents

Introduction

The GAIA-CLIM project aims at assessing and improving global capabilities of ground-based, balloon-borne, and aircraft-based measurements (termed non-satellite measurements henceforth) for the characterisation of satellite measurements. The work under GAIA-CLIM encompasses the following tasks:

- Definition and mapping of existing non-satellite measurement capabilities;
- Improving the metrological characterisation of a subset of non-satellite (reference) observational techniques;
- Better accounting for co-location mismatches between satellite observations and non-satellite (reference) observations;
- Exploration of the role of data assimilation as an integrator of information;
- Creation of a 'Virtual Observatory' bringing together all comparison data, including their uncertainties, and providing public access to the information they contain;
- Identification and prioritization of gaps in knowledge and in capabilities. Under its work package 6, GAIA-CLIM performs an assessment of gaps in capabilities or knowledge relevant to the use of non-satellite data to characterise satellite measurements.

It is recognized that GAIA-CLIM shall provide progress in these application areas, but not necessarily close out all potential issues and challenges. Hence, in each of the project tasks outlined above, presently unfulfilled user needs ('gaps') have been identified through an iterative process throughout the project's lifetime. This gaps assessment exercise exclusively considers gaps identified as relevant to these GAIA-CLIM project aims. The identified key user communities for whom the impact of the identified gaps would be most relevant include:

- Service providers (e.g., ECMWF for NWP, CAMS and C3S)
- Providers and users of ECV climate data records (e.g., space agencies and satellite data user communities)
- Users of reference observations
- Users of baseline network observations
- Users of the 'Virtual Observatory'

The Gaps Assessment and Impacts Document (GAID) is a living document that summarises the outcome of this collection of gaps and their proposed remedies. It further describes the gap identification process, as well as the way these findings are presented and made accessible to users, stakeholders, and actors. The current set of gaps and remedies captured under the living GAID document v4 provides a firm basis for providing costed and prioritised recommendations for future work to improve our ability to use non-satellite data to characterise satellite measurements. The first draft of recommendations document[1] builds upon this careful and meticulous collection and cataloguing process to produce a set of eleven overarching recommendations for future work to close the most critical gaps identified through the life of the project

This document provides a snapshot of the gaps status as per December 2017 in relation to work package 3. It provides a third, and final, formal delivery of WP3 input to the process. The on-line 'Catalogue of Gaps' provides the latest version of the full content of the gaps and their proposed remedies. The catalogue is available from: http://www.gaia-clim.eu/page/gap-reference-list.

[1] http://www.gaia-clim.eu/page/recommendations

Input from external parties continues to be invited through the GAID website. A designated e-mail address[2] and a specific template for gap reporting are provided at the website. Further user engagement shall be achieved through a series of visits to key stakeholders through the end of 2017. This user feedback will be important in refining the GAID and ensuring its usefulness to the broader scientific and policymaker communities, as well as space agencies, international organisations and funding bodies.

[2] Email address for GAID feedback: gaid@gaia-clim.eu

1. Summary of existing gaps for WP3

Table 1.1. Overview of the gaps identified under work package 3 under GAID V4 and their identified remedies

Gap reference/ ownership	Gap title	Remedies
G3.01	Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location	• (R1) Improved high-resolution modelling to quantify mismatch effects • (R2) Use of statistical analysis techniques based upon available and targeted additional observations
G3.02	Missing standards for, and evaluation of, co-location criteria	• (R1) Systematic quantification of the impacts of different co-location criteria
G3.04	Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties	• (R1) Comprehensive modelling studies of measurement process • (R2) Empirical determination of true resolution by comparison with higher-resolution data
G3.05	Representativeness uncertainty assessment missing for higher-level data based on averaging of individual measurements	• (R1) Quantification of representativeness of averages using modelling, statistical and sub-sampling techniques
G3.06	Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences	• (R1) Use of Observing System Simulation Experiments (OSSEs) • (R2) Statistical estimation of typical co-location mismatch effects

2. Detailed update on traces for the gaps arising from WP3

[Please describe the changes you made to the v4 gap traces. In analogy to a review process when answering to the reviewers, you should explain **why** you made those changes.]

G3.01: "Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location"

No substantial changes besides some textual editing to make the gap trace easier to read.

G3.02: "Missing standards for, and evaluation of, co-location criteria

Added as a related gap G6.03 "Lack of sustained dedicated periodic observations to coincide with satellite overpasses to minimise co-location effects" since dedicated observations need to take into account what is learned when determining optimal co-location criteria.

G3.04: "Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties"

A 2nd remedy was introduced, in addition to the existing remedy of detailed modelling of the measurement process. The new remedy concerns the empirical determination of the true measurement resolution by intercomparison with higher-resolution data. In this approach, the true field-of-view and its sensitivity distribution are estimated via an optimization process on the differences with the high-resolution data. It is critically dependent upon the availability of a second high-quality measurement system that measures sufficiently similarly to the target measurement system.

Furthermore, a few textual improvements were made to improve readability and we also added as a related gap G6.03, since dedicated observations need to take into account the actual multi-dimensional smoothing and sampling properties of both measurement systems.

G3.05: "Representativeness uncertainty assessment missing for higher-level data based on averaging of individual measurements"

Only a few textual improvements were made, including a more explicit mention that this gap was outside the scope of GAIA-CLIM, the work in which focused on Level-2, and to some extent, Level-1b data.

G3.06: "Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences"

Only a few textual improvements were made.

3. Conclusions

The gaps identified from the perspective of WP3 remain largely unchanged during this final iteration step. A 2nd remedy was formulated for G3.04 and minor work was done to improve readability. Several gaps were linked to gap G6.03, which deals with the need for dedicated non-satellite measurements, coinciding as much as possible with satellite overpasses.

4. Annex I Updated GAIA-CLIM Catalogue of gaps for WP3

Within this section, gaps that were detailed in section 1 are here expanded to give full trace of the current understanding of the gaps including a revision of its impacts and potential remedies

G3.01 Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location

Gap abstract:

The atmospheric concentration of nearly all ECVs varies in space and time at the scale of the individual measurements, and at the scale of their co-location in the context of data comparisons (e.g., for the purpose of satellite validation, data merging, and data assimilation). However, the amplitude and patterns of these variations are often unknown on such small scales. Consequently, it is impossible to quantify the uncertainties that result from sampling and smoothing properties of the measurements of the variable, structured atmospheric field. This gap thus concerns the need for a better quantification of atmospheric spatiotemporal variability at the small scales of individual measurements and co-locations.

Part I: Gap description

Primary gap type:

Uncertainty in relation to comparator

Secondary gap type:

- Knowledge of uncertainty budget and calibration
- Parameter (missing auxiliary data etc.)

ECVs impacted:

Temperature, Water vapour, Ozone, Aerosols, Carbon Dioxide, Methane

User category/Application area impacted:

- Operational services and service development (meteorological services, environmental services, Copernicus Climate Change Service (C3S) and Atmospheric Monitoring Service (CAMS), operational data assimilation development, etc.)
- International (collaborative) frameworks and bodies (SDGs, space agencies, EU institutions, WMO programmes/frameworks etc.)
- Climate research (research groups working on development, validation and improvement of ECV Climate Data Records)

Non-satellite instrument techniques involved:

Independent of instrument technique

Related gaps:

- G3.04 Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties
- G3.06 Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences

G3.04. To be addressed after G3.01

Argument: To estimate the additional uncertainties on a measurement that result from spatiotemporal atmospheric variability at the measurement sampling and smoothing scales, a quantification of that spatiotemporal variability is a prerequisite.

G3.06. To be addressed after G3.01

Argument: Understanding the uncertainty budget of a comparison (in a validation context) requires a quantification of the impact of co-location mismatch. This cannot be done without an estimate of the spatiotemporal variability of the ECV under study.

Detailed description:

Spatiotemporal variability of the atmosphere at the scale of the airmass being measured or - in the case of a measurement intercomparison - at the scale of the co-location, leads to additional uncertainties, not accounted for by the uncertainty budget reported with an individual measurement (Lambert et al., 2012). To quantify these additional uncertainties (cf. gaps G3.04 and G3.06), or to ensure that they remain negligible through the use of appropriate co-location criteria (cf. G3.02), a prerequisite is a proper understanding of atmospheric variability of the targeted ECV on those scales.

While scales above approx. 100km and 1h are relatively well captured for several GAIA-CLIM target ECVs in model or satellite gridded data (e.g., Verhoelst et al., 2015, for total ozone), information on smaller scales is most often restricted to results from dedicated campaigns or specific case studies, e.g., Sparling et al. (2006) for ozone profiles, Hewison (2013) for meteorological variables, and Pappalardo et al. (2010) for aerosols. Due to the exploratory nature of these studies, neither global nor complete vertical coverage is achieved. For instance, information on small-scale variability in the ozone field is limited to altitudes and regions probed with dedicated aircraft campaigns. The validation of satellite data records with pseudo global networks of ground-based reference instruments on the other hand requires an appropriate quantification of atmospheric variability in very diverse conditions, covering all latitudes, altitudes, dynamical conditions, degrees of pollution etc.

This gap therefore concerns the need for a better, more comprehensive, quantification of the spatiotemporal variability of the ECVs targeted by GAIA-CLIM.

Validation aspects addressed:

- Radiance (Level 1 product)
- Geophysical product (Level 2 product)
- Gridded product (Level 3)
- Assimilated product (Level 4)
- Time series and trends
- Representativity (spatial, temporal)
- Calibration (relative, absolute)
- Spectroscopy
- Auxiliary parameters (clouds, lightpath, surface albedo, emissivity)

Gap status after GAIA-CLIM:

GAIA-CLIM explored and demonstrated potential solutions to close this gap in the future:

Within GAIA-CLIM, a work package (WP3) was dedicated to research on co-location mismatch in an inhomogeneous and variable atmosphere. In the context of this work package, some studies were performed that quantified spatiotemporal variability for a few ECVs at a limited scale domain (e.g. temperature and water vapour temporal variability at 6-hour scale from radiosonde inter-comparisons, and aerosol optical depth variability at the scale of a satellite-ground co-location in the North-East US). Although this work was limited to a few ECVs, scales, geographical coverage etc. owing to the limited resources and data availability, GAIA-CLIM has demonstrated use cases / case studies which may permit a more exhaustive approach in future. Fully addressing this gap requires significant observational and modelling work, far beyond the scope of GAIA-CLIM, as described in detail in the remedies.

Part II: Benefits to resolution and risks to non-resolution

Identified Benefit	User category/Application area benefitted	Probability of benefit being realised	Impacts
Improved understanding of single measurement uncertainty, including the impact of the instrument smoothing and sampling properties	All users and application areas will benefit from it.	High	More reliable uncertainty estimates allow for more confidence in the data and optimized use in e.g. assimilation and other applications.
Improved definition of appropriate co-location criteria for validation work, minimizing errors due to co-location mismatch	All users and application areas will benefit from it.	High	Lower uncertainty due to co-location mismatch will result in tighter constraints on the products from validation work, supporting further instrument and algorithm development.
Improved interpretation of comparison results because co-location mismatch errors can be quantified.	All users and application areas will benefit from it.	High	Improved quantification of the uncertainty due to co-location mismatch will allow more stringent tests of the reported measurement uncertainties, supporting further instrument and algorithm development.
Identified risk	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**
Incomplete uncertainty budget for single measurements and derived products	All users and application areas will suffer from it.	High	Poor confidence in data and services; potential over-interpretation; difficult/unreliable generation of higher level data products (through data assimilation and/or merging).

Incomplete uncertainty budget for data comparisons	All users and application areas will suffer from it.	High	Sub-optimal feedback from data comparisons, in particular in the context of satellite validation. Potential of both EO and ground segments not fully realized.

Part III: Gap remedies

Gap remedies:

Remedy 1: Improved high-resolution modelling to quantify mismatch effects

Primary gap remedy type:

Research

Secondary gap remedy type:

Technical

Proposed remedy description:

A first remedy to gain better insight in the small-scale spatiotemporal variability of atmospheric ECVs is by high-resolution modelling studies at the global scale, resulting in comprehensive data sets of atmospheric fields, at high horizontal, vertical, and temporal resolution, based not solely on higher-resolution grids but also including the relevant physics and (photo) chemistry at those scales.

Improved spatiotemporal resolution in atmosphere models is a much broader scientific goal, with great computational and theoretical (e.g. convection and turbulence treatment) challenges. As such, this remedy probably requires a level of effort and resources beyond what can be justified solely by the need for satellite data validation. The technological/ organizational viability is therefore considered medium and the cost estimate high.

Relevance:

If successful, this remedy would largely close the gap, and it would facilitate remedies for most other gaps related to comparator uncertainties through the use of OSSEs (Observing System Simulation Experiments) based on these modelled fields.

Measurable outcome of success:

The quality of the model output at its finest resolution can be estimated by comparison with high-resolution measurement data sets, preferably those with limited horizontal, vertical, and temporal smoothing effects, e.g.

from balloon-borne sondes. Ideally, an agreement is found within the combined model and measurement uncertainties.

Expected viability for the outcome of success:

Medium

Scale of work:

Single Institution, Consortium

Time bound to remedy:

Less than 10 years

Indicative cost estimate (investment):

High cost (> 5 million)

Indicative cost estimate (exploitation):

Non-applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services

Remedy 2: Use of statistical analysis techniques based upon available and targeted additional observations

Primary gap remedy type:

Research

Secondary gap remedy type:

Technical

Proposed remedy description:

This remedy concerns the statistical analysis of existing and future satellite and non-satellite high-resolution data sets, which allows us to separate the contribution of atmospheric variability from the total uncertainty budget of a data comparison, e.g. using so-called ‘structure functions’ or heteroskedastic functional regression. Within the geographical and temporal coverage of the data set, these methods produce an estimate of the variability (or auto-correlation) of the field. Note that, as for Remedy G3.01(1), the scientific interest for higher resolution in the data sets is much broader than only the validation needs, e.g. for the identification of emission sources in an urban environment.

The technological and organizational effort required to make step changes in the spatiotemporal resolution of the observational data sets is in general very large, and comes with a large financial cost (more than 5M euro), in particular if global coverage is aimed for. Hence, such developments need a much larger user base and the use proposed here should be considered secondary to the scientific objectives of such new missions. Nevertheless, smaller dedicated campaigns with for instance aircraft or Unmanned Aerial Vehicles (UAVs) can offer great insight at particularly interesting sites (e.g. at ground stations with a multitude of instruments observing a particular ECV), and this at medium cost (between 1M and 5M euro).

Relevance:

This remedy directly addresses the gap, as already illustrated for instance with aircraft data for ozone by Sparling et al. (2006).

Measurable outcome of success:

The primary outcome would be publications describing for the different ECVs and various atmospheric regimes, locations and altitude ranges the atmospheric variability at scales ranging from those of in-situ measurements (e.g. 10s of meters for balloon sonde measurements) to that of a satellite pixel (several 10s to 100s of kilometres). These can be based either on existing data sets, or represent an exploitation of newly designed campaigns and missions.

Expected viability for the outcome of success:

High

Scale of work:

Single institution

Time bound to remedy:

Less than 5 years

Indicative cost estimate (investment):

Low cost (< 1 million)

Indicative cost estimate (exploitation):

Non-applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological services

References:

- Butterfield et al.:" Determining the temporal variability in atmospheric temperature profiles measured using radiosondes and assessment of correction factors for different launch schedules", AMT, v8, 2015
- Lambert, J.-C., et al., "Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues", in "Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods", N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 pp., ISBN: 978-1-4614-3908-0, DOI 10.1007/978-1-4614-3909-7_2, © Springer New York 2012
- Pappalardo et al., "EARLINET correlative measurements for CALIPSO: First intercomparison results", J.G.R.: Atmospheres v115, 2010
- Sparling et al., "Estimating the impact of small-scale variability in satellite measurement validation", J.G.R.: Atmospheres v111, 2006
- Verhoelst et al., "Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons", AMT v8, 2015

G3.02 Missing standards for, and evaluation of, co-location criteria

Gap abstract:

The impact of a particular choice of co-location criterion is only rarely studied in the scientific literature reporting on satellite validation results. However, without some quantification of the impact of the co-location criterion that was adopted, it is virtually impossible to assess the contribution of natural variability to the total error budget of the data comparisons. As such, this gap impacts significantly the potential interpretation of the data comparison result in terms of data quality. Some in-depth studies do exist, but testing multiple criteria, or using criteria based on the latest results of such exploratory work, is far from common (indeed, often arbitrary) practice(s). This gap thus concerns the need for more awareness among validation teams, for more detailed studies comparing the (dis-)advantages of various co-location criteria, and for community-agreed standards on co-location criteria that are broadly adopted in the context of operational services.

Part I: Gap description

Primary gap type:

Uncertainty in relation to comparator

Secondary gap type:

Governance (missing documentation, cooperation etc.)

ECVs impacted:

Temperature, Water vapour, Ozone, Aerosols, Carbon Dioxide, Methane

User category/Application area impacted:

- Operational services and service development (meteorological services, environmental services, Copernicus services (C3S) and Atmospheric Monitoring Service (CAMS), operational data assimilation development, etc.)
- International (collaborative) frameworks and bodies (SDGs, space agency, EU institutions, WMO programmes/frameworks etc.)
- Climate research (research groups working on development, validation and improvement of ECV Climate Data Records)

Non-satellite instrument techniques involved:

Independent of instrument technique

Related gaps:

- G3.04 Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties
- G3.06 Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences
- G6.03 Lack of sustained dedicated periodic observations to coincide with satellite overpasses to minimise co-location effects

G3.04. To be addressed before G3.02

Argument: Ideally, co-location criteria take into account the smoothing and sampling properties of the measurements. Consequently, studies on co-location criteria can benefit from a proper characterization of these smoothing and sampling properties.

G3.06. To be addressed before G3.02

Argument: The merit of certain co-location criteria can best be assessed when the uncertainty budget of the comparisons is decomposed in measurement and co-location mismatch uncertainties.

G6.03. To be addressed after G3.02

Argument: Deciding on the best time and location for targeted reference observations should be informed by information on the optimal co-location criteria.

Detailed description:

Co-location criteria should represent an optimal compromise between the obtained number of co-located measurements (as large as possible to have robust statistical results) and the impact of natural variability on the comparisons (as low as possible to allow a meaningful comparison between measured differences and reported measurement uncertainties). Hitherto, only a few ground-based satellite validation studies explored the impact of the adopted co-location criteria on the comparison results (e.g. Wunch et al., 2011, and Dils et al., 2014, for CO_2 , Verhoelst et al., 2015, for O_3, Pappalardo et al., 2010, for aerosols, Lambert et al. 2012, for water vapour, Van Malderen et al. 2014, for integrated water vapour). Still, atmospheric variability is often assumed –or even known- to impact the comparisons, but without detailed testing of several co-location criteria (or by extensive model-based simulations), this impact is hard to quantify. Besides the need for dedicated studies, this gap also concerns the "community practices" regarding co-location approaches, which are neither consistent across different studies, nor optimal as they often rely on historical co-location criteria, which are not necessarily fit-for-purpose for the accuracy and spatiotemporal sampling properties of current measurement systems. Consequently, to ensure reliable and traceable validation results, as required in an operational context, community-agreed standards for co-location criteria should be developed, published, and adopted.

Operational space missions or space instruments impacted:

Independent of specific space mission or space instruments

Validation aspects addressed:

- Radiance (Level 1 product)
- Geophysical product (Level 2 product)
- Gridded product (Level 3)
- Assimilated product (Level 4)

- Time series and trends
- Representativity (spatial, temporal)
- Calibration (relative, absolute)
- Spectroscopy
- Auxiliary parameters (clouds, lightpath, surface albedo, emissivity)

Gap status after GAIA-CLIM:

GAIA-CLIM explored and demonstrated potential solutions to close this gap:

Two activities within GAIA-CLIM targeted this gap to some extent:

Within GAIA-CLIM, a dedicated task (T3.2 in WP3) dealt with data intercomparison studies, focusing on a closure of the comparison uncertainty budget and including an exploration of different co-location criteria, see for instance the results on total ozone columns published by Verhoelst et al. (2015, their Fig. 11).

The Virtual Observatory developed within GAIA-CLIM offers the user the possibility to adjust co-location criteria and to visualize the resulting impact on the comparison results.

However, no attempt has been made within GAIA-CLIM to produce an authoritative analysis and resulting documentation on this matter.

Part II: Benefits to resolution and risks to non-resolution

Identified Benefit	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**
Greater awareness of the impact of natural variability on the comparison results;	All users and application areas will benefit from it	High	More reliable feedback on data quality, in particular on the reported uncertainties. This in turn increases confidence in the data for the end user and allows more meaningful use in a variety of applications.
Improved definition of appropriate co-location criteria for validation work, minimizing errors due to co-location mismatch.	All users and application areas will benefit from it	High	Lower uncertainty due to co-location mismatch will result in tighter constraints on the products from validation work, supporting further instrument and algorithm development.
Facilitates intercomparison of different validation studies	All users and application areas will benefit from it	High	More reliable comparisons between different products (each having its own validation report) to better assess their fitness-for-purpose for a specific user application.
Identified risk	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**

Poor feedback on data quality (in particular on the reported uncertainties) from validation studies due to unknown/unquantified influence of atmospheric variability.	All users and application areas will suffer from it.	High	Poor confidence in data and services; potential over-interpretation; difficult/unreliable assimilation; Potential of both EO and ground segments not fully realized.
Difficulty to compare validation results on similar products performed by different teams	All users and application areas will suffer from it.	High	Sub-optimal choice of data product for a given application.

Part III: Gap remedies

Gap remedies:

Remedy 1: Systematic quantification of the impacts of different co-location criteria

Primary gap remedy type:

Research

Secondary gap remedy type:

Governance

Proposed remedy description:

Dedicated studies are required which explore in detail the advantages and disadvantages of several co-location methods and criteria. Dedicated working groups or activities could/should be set up within the framework of the ground-based observing networks, as already initiated for meteorological variables at a GRUAN-GSICS-GNSSRO WIGOS workshop on Upper-Air Observing System Integration and Application, hosted by WMO in Geneva in May 2014. Dissemination among, and acceptance by, the key stakeholders may be challenging and can probably best be achieved in the context of overarching frameworks such as the CEOS Working Group on Calibration & Validation (WGCV). The financial cost should be very low. Also, the space agencies and service providers could/should insist on sufficient attention for (and analysis of) the adopted co-location criteria in the validation protocols followed by their validation teams.

Relevance:

These studies and the proposed associated governance support target this gap directly. They will provide stakeholders with a traceable, authoritative reference on which to base their validation requirements and protocols regarding co-location criteria. It will also facilitate meta-analysis of different validation studies without the need to take into account differences in results due to differences in the impact of co-location mismatch on the results.

Measurable outcome of success:

Peer-reviewed publications or widely distributed technical notes on the subject, from an authoritative body; Explicit inclusion of requirements on the co-location methodology and criteria in validation protocols.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 3 years

Indicative cost estimate (investment):

Low cost (< 1 million)

Indicative cost estimate (exploitation):

No

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- WMO
- ESA, EUMETSAT or other Space agency
- Academia, individual research institutes

References:

- 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
- Lambert, J.-C., et al., "Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues", in "Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods", N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 pp., ISBN: 978-1-4614-3908-0, DOI 10.1007/978-1-4614-3909-7_2, © Springer New York 2012

- Pappalardo et al., “EARLINET correlative measurements for CALIPSO: First intercomparison results”, J.G.R.: Atmospheres v115, 2010
- Van Malderen, R. et al., “A multi-site intercomparison of integrated water vapour observations for climate change analysis”, AMT v7, 2014
- Verhoelst et al., “Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons”, AMT v8, 2015
- Wunch et al., “A method for evaluating bias in global measurements of CO2 total columns from space”, ACP v11, 2011

.

G3.04 Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties

Gap abstract:

This gap concerns the need for a more detailed characterisation of the actual spatiotemporal smoothing and sampling properties of both satellite-based EO measurements and ground-based in-situ or remote-sensing measurements. Indeed, EO measurements are most often associated with single locations, or at best pixel footprints, while in fact the actual measurement sensitivity covers a larger spatiotemporal extent, due for instance to the radiative transfer determining the measured quantities, or the actual measurement geometry (choice of line-of-sight, trajectory of a weather balloon, etc.). In an inhomogeneous and variable atmosphere, this leads to additional errors and uncertainties that are not part of the reported measurement uncertainties, but still need to be quantified, in particular when performing comparisons with other types of measurements, with different smoothing and sampling characteristics. For several ECVs and measurement techniques, significant work is needed to (1) determine/model the actual spatiotemporal smoothing and sampling properties, and (2) quantify the resulting uncertainties on the measurements of the variable atmosphere.

Part I: Gap description

Primary gap type:

Knowledge of uncertainty budget and calibration

Secondary gap type:

Uncertainty in relation to comparator measures

ECVs impacted:

Temperature, Water vapour, Ozone, Aerosols, Carbon Dioxide, Methane

User category/application area impacted:

- Operational services and service development (meteorological services, environmental services, Copernicus services (C3S) and Atmospheric Monitoring Service (CAMS), operational data assimilation development, etc.)
- International (collaborative) frameworks and bodies (SDGs, space agency, EU institutions, WMO programmes/frameworks etc.)
- Climate research (research groups working on development, validation and improvement of ECV Climate Data Records)

Non-satellite instrument techniques involved:

- Radiosonde
- Ozonesonde
- Lidar
- FPH/CFH
- Microwave Radiometer
- FTIR
- Brewer/Dobson
- UV/VIS zenith DOAS
- UV/VIS MAXDOAS

Related gaps:

- G3.01 Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location
- G3.06 Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences
- G6.03 Lack of sustained dedicated periodic observations to coincide with satellite overpasses to minimise co-location effects

G3.01. To be addressed before G3.04
Argument: A quantification of the uncertainties that result from the specific sampling and smoothing properties of an instrument requires information on the spatiotemporal variability of the atmospheric field.

G3.06. To be addressed after G3.04
Argument: Error/uncertainty budget decomposition of a comparison requires a proper understanding of the smoothing and sampling properties of the instruments involved, i.e. requires G3.04 to be remedied.

G6.03. To be addressed after G3.02
Argument: Deciding on the best time and location for targeted reference observations should be informed by information on the actual sampling and smoothing properties of the measurement systems.

Detailed description:

Remotely sensed data are often considered as column-like or point-like samples of an atmospheric variable, associated for instance with the location of a ground-based instrument. This is also the general assumption for satellite data, which are assumed to represent the column or profile above the satellite field-of-view footprint in case of nadir sounders, and atmospheric concentrations along a vertical set 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 O_3, NO2, BrO etc.) over several hundreds of kilometres 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 cannot be associated with a single geo-location and time stamp, due for instance to balloon drift for ozone- and radiosondes. 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, 2012).

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 (c.f. G3.01), 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 the EC FP6 GEOmon project and in the current EC H2020 GAIA-CLIM project), but in the context of integrated systems like Copernicus and GCOS, appropriate knowledge of smoothing and sampling uncertainties, which is still missing for several ECVs and remote sensing measurement types, has to be further developed and harmonized.

Operational space missions or space instruments impacted:

Independent of specific space mission or space instruments

Validation aspects addressed:

- Radiance (Level 1 product)
- Geophysical product (Level 2 product)
- Gridded product (Level 3)
- Assimilated product (Level 4)
- Time series and trends
- Representativity (spatial, temporal)
- Calibration (relative, absolute)
- Spectroscopy
- Auxiliary parameters (clouds, lightpath, surface albedo, emissivity)

Gap status after GAIA-CLIM:

GAIA-CLIM explored and demonstrated potential solutions to close this gap

Addressing this gap was a major objective of GAIA-CLIM, within which specific tasks were dedicated to the characterisation of smoothing and sampling properties of selected instruments and for selected ECVs. Results have been obtained for total ozone columns, for ozone, temperature, and humidity profiles, and for aerosol columns and profiles from a diverse set of ground-based instruments. Regarding satellite data, only a selection of current missions were explored. Results were made available in technical notes, namely D3.4 ("Report on measurement mismatch studies and their impact on data comparisons") and D3.6 ("Library of (1) smoothing/sampling error estimates for key atmospheric composition measurement systems and (2) smoothing/sampling error estimates for key data comparisons"), and through the 'Virtual Observatory'. In the long term, this gap will require continued efforts to fully characterize the spatiotemporal smoothing and sampling properties of both new ground-based instruments and upcoming satellite sensors. Hence the gap requires constant re-evaluation as technology and observing programs evolve.

Part II: Benefits to resolution and risks to non-resolution

Identified Benefit	User category/Application area benefitted	Probability of benefit being realised	Impacts
More complete assessment of the impact of natural variability on the measurements;	All users and application areas will benefit from it	High	Better uncertainty characterization. This in turn increases confidence in the data for the end user and allows more meaningful use in a variety of applications.

Improved definition of appropriate co-location criteria for validation work, taking into account the actual sampling and smoothing properties, and ultimately minimizing errors due to co-location mismatch.	All users and application areas will benefit from it	High	Lower uncertainty due to co-location mismatch will result in tighter constraints on the products from validation work, supporting further instrument and algorithm development.
Identified risk	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**
Incomplete total uncertainty budget for a single measurements.	All users and application areas will suffer from it.	High	Incomplete data characterization and potentially limited or flawed interpretation, whatever the use type.
Incomplete uncertainty budget for measurement comparisons, e.g. for validation.	All users and application areas will suffer from it.	High	Flawed validation results: missing uncertainty components lead to failed consistency checks, and a less performant validation system.

Part III: Gap remedies

Gap remedies:

Remedy 1: Comprehensive modelling studies of measurement process.

Primary gap remedy type:

Research

Proposed remedy description:

Detailed modelling of the measurement process, including multi-dimensional radiative transfer if applicable, to quantify the 4-D measurement sensitivity. An example are multi-D averaging kernels for retrieval-type measurements. This work requires a significant effort from the instrument teams, for which dedicated, though still relatively low (per instrument), resources are required, in particular for code modifications and additions. If appropriate, the results from these detailed calculations can be parametrized for easy and efficient use when calculating the resulting errors and uncertainties for large amounts of data. This uncertainty calculation is done by combining the quantification of the measurement sensitivity with knowledge on the spatiotemporal variability of the atmospheric field (cf. G3.01). When these detailed modelling studies are out of reach, a similar estimate of the multi-D measurement sensitivity can be made in a more pragmatic way based on the measurement principle and physical considerations (e.g. Lambert et al. 2011), or it can in some cases be estimated with empirical methods by comparing data sets with differing resolution. Note that an essential prerequisite is the availability of all required metadata with the measurements, such as viewing angles or GPS trajectories.

Relevance:

This remedy will provide a description for every instrument and measurement type of the full 4-D measurement sensitivity, and the errors and uncertainties resulting from the assumption that a measurement can be associated with a nominal geo-location and time.

Measurable outcome of success:

Publications and technical notes describing for every instrument and measurement type the full 4-D measurement sensitivity, and the errors and uncertainties resulting from the assumption that a measurement can be associated with a nominal geo-location and time.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 5 years

Indicative cost estimate (investment):

Low cost (< 1 million)

Indicative cost estimate (exploitation):

No

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- ESA, EUMETSAT or other Space agency
- Academia, individual research institutes

Remedy 2: Empirical determination of true resolution by comparison with high-resolution data.

Primary gap remedy type:

Research

Proposed remedy description:

If temporally coinciding data with higher spatial resolution are available, the true horizontal resolution of a measurement system can be determined empirically by comparing the measurements of the two instruments as obtained on the same scene. This approach was for instance demonstrated by Sihler et al. (2017) for satellite and ground-based DOAS-type measurements. It is empirical in the sense that it does not require extensive modelling of the measurement process. Rather, it requires some basic assumptions on the actual footprint and the sensitivity therein of each measurement, which is then further optimized by comparison with the high-resolution data set, if necessary over a large set of diverse scenes. This approach was also explored within GAIA-CLIM, where it was used to estimate the true vertical resolution and weighting function of temperature and humidity soundings, as described in D3.4.

Relevance:

This remedy addresses the gap partially (since it only deals with the resolution aspects) and it requires an independent, high-resolution data set of sufficient quality. As such, it is not universally applicable, but it does provide a valuable resolution estimate, independent of any classical metrological modelling

Measurable outcome of success:

Publications and technical notes describing for every instrument and measurement type the full 4-D measurement sensitivity, and the errors and uncertainties resulting from the assumption that a measurement can be associated with a nominal geo-location and time.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 5 years

Indicative cost estimate (investment):

Low cost (< 1 million)

Indicative cost estimate (exploitation):

Non- applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- ESA, EUMETSAT or other space agency
- Academia, individual research institutes

References:

- Lambert, J.-C., et al., Comparison of the GOME ozone and NO2 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
- Lambert, J.-C., et al., "Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues", in "Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods", N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 pp., ISBN: 978-1-4614-3908-0, DOI 10.1007/978-1-4614-3909-7_2, © Springer New York 2012
- Von Clarmann et al., "The horizontal resolution of MIPAS", AMT v2, 2009
- Seidel et al., "Global radiosonde balloon drift statistics", J.G.R. v116, 2011
- Sihler, H., Lübcke, P., Lang, R., Beirle, S., de Graaf, M., Hörmann, C., Lampel, J., Penning de Vries, M., Remmers, J., Trollope, E., Wang, Y., and Wagner, T.: In-operation field-of-view retrieval (IFR) for satellite and ground-based DOAS-type instruments applying coincident high-resolution imager data, Atmos. Meas. Tech., 10, 881-903, https://doi.org/10.5194/amt-10-881-2017, 2017.

G3.05 Representativeness uncertainty assessment missing for higher-level data based on averaging of individual measurements

Gap abstract:

Level-3 data are, by definition, constructed by averaging asynoptic level-2 data over certain space-time intervals, so as to arrive at a (regularly) gridded data product. However, the (global) sampling pattern of the sounder(s) that produced the original level-2 data is never perfectly uniform, nor are revisit times short enough to guarantee dense and homogeneous temporal sampling of e.g. a monthly mean at high horizontal resolution. Consequently, the averages may deviate substantially from the true average field that would be obtained if complete spatiotemporal coverage were possible. These so-called representativeness errors are only rarely investigated, and almost never provided with a product, in spite of their importance in interpreting the data.

Part I: Gap description

Primary gap type:

Knowledge of uncertainty budget and calibration

Secondary gap type:

- Uncertainty in relation to comparator measures
- Governance (missing documentation, cooperation etc.)

ECVs impacted:

Temperature, Water vapour, Ozone, Aerosols, Carbon Dioxide, Methane

User category/application area impacted:

- Operational services and service development (meteorological services, environmental services, Copernicus services (C3S) and Atmospheric Monitoring Service (CAMS), operational data assimilation development, etc.)
- International (collaborative) frameworks and bodies (SDGs, space agency, EU institutions, WMO programmes/frameworks etc.)
- Climate research (research groups working on development, validation and improvement of ECV Climate Data Records)

Non-satellite instrument techniques involved:

- Radiosonde
- Ozonesonde
- Lidar
- FPH/CFH

- Microwave Radiometer
- FTIR
- Brewer/Dobson
- UV/VIS zenith DOAS
- UV/VIS MAXDOAS
- Pandora
- GNSS-PW

Related gaps:

- G3.01 Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location

G 3.01. To be addressed before G3.05

Argument: A quantification of representativeness uncertainties requires an adequate representation of the atmosphere at the scale of the measurements.

Detailed description:

The creation of level-3 data by averaging non-uniformly distributed level-2 measurements inevitably leads to representativeness errors, see e.g. Coldewey-Egbers et al., (2015) for the case of a level-3 (gridded monthly means) total ozone data set. The resulting representativeness uncertainty can be larger than the formal uncertainty on the mean. In the best case this would represent an additional random uncertainty term. If the sampling pattern of the sounder changes in time, this may give rise to systematic, time-dependent representativeness errors that affect for example trend analyses for climate research (see e.g. Damadeo et al., 2014). However, estimates of these representativeness uncertainties are rarely included with the data product. Also, the representativeness of the ground-based network should be taken into account when validating such data sets, i.e. the sparse spatial and temporal sampling of the ground network leads to significant representativeness uncertainties in for instance derived monthly (zonal) means.

Note that also in the context of validation of level-2 data, measurements are sometimes averaged after co-location (e.g. Valks et al., 2011; Schneising et al.,2012) without explicit calculation of the representativeness errors and resulting uncertainty.

Operational space missions or space instruments impacted:

Independent of specific space mission or space instruments

Validation aspects addressed:

- Gridded product (Level 3)
- Time series and trends
- Representativity (spatial, temporal)

Gap status after GAIA-CLIM:

After GAIA-CLIM this gap will remain as it was not addressed within the project (level-3 and level-4 data were in general not addressed within the project).

Part II: Benefits to resolution and risks to non-resolution

Identified Benefit	User category/Application area benefitted	Probability of benefit being realised	Impacts
More complete uncertainty quantification on the reported data.	All users and application areas will benefit from it	High	Better uncertainty characterization. This in turn increases confidence in the data for the end user and allows more meaningful use in a variety of applications.
Identified risk	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**
Underestimated uncertainty on the reported data (averages)	All users and application areas will suffer from it.	High	Incomplete data characterization and potentially limited or flawed interpretation, whatever the use type.

Part III: Gap remedies

Gap remedies:

Remedy1: Quantification of representativeness of averages using modelling, statistical and sub-sampling techniques

Primary gap remedy type:

Research

Secondary gap remedy type:

Governance

Proposed remedy description:

Studies are required quantifying the representativeness of averages, e.g. by model-based simulations of averages based on either the limited real sampling or on an ideal, complete sampling. This approach was followed for instance by Coldewey-Egbers (2015) for a total ozone L3 product. More pragmatically, representativeness uncertainties can also be computed as a function of parametrized measurement inhomogeneity and climatological field variability (for instance Sofieva et al., 2014). Note that the demand for such studies is also a governance issue: service providers and overarching frameworks should insist that any L3 data set comes with such a quantification of representativeness uncertainties.

The effort required to address this gap depends on the particular product and on whether atmospheric variability is well understood for that ECV (c.f. gap G3.01). For most of the ECVs targeted by GAIA-CLIM, an estimate of the representativeness uncertainty should be achievable at a low cost. The additional validation required to assess the quality of this representativeness uncertainty estimate may –in absence of existing reference data sets at sufficiently high spatial and temporal sampling– require a more significant investment, e.g. to conduct intensive field campaigns.

Relevance:

This remedy directly addresses and fills the gap.

Measurable outcome of success:

Success is achieved when level-3 data sets include not only the formal uncertainty on the mean and the variance around that mean, but also an estimate of the representativeness uncertainty on that mean. The reliability of this reported representativeness uncertainty must than also be validated or verified.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 3 years

Indicative cost estimate (investment):

Low cost (< 1 million)

Indicative cost estimate (exploitation):

Non-applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- ESA, EUMETSAT or other space agency
- Academia, individual research institutes

References:

- 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
- Damadeo et al.,: ”Reevaluation of stratospheric ozone trends from SAGE II data using a simultaneous temporal and spatial analysis”, Atmos. Chem. Phys., 14, 2014
- Lambert, J.-C., et al., “Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues”, in “Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods”, N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 pp., ISBN: 978-1-4614-3908-0, DOI 10.1007/978-1-4614-3909-7_2, © Springer New York 2012
- 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

G3.06 Missing comparison (validation) uncertainty budget decomposition including uncertainty due to sampling and smoothing differences

Gap abstract:

A data validation study is meant to check the consistency of a given dataset with respect to a reference dataset within their reported uncertainties. As such, the uncertainty budget of the data comparison is crucial. Besides the measurement uncertainties on both data sets, the discrepancy between the two datasets will be increased by uncertainties associated with data harmonization manipulations (e.g. unit conversions requiring auxiliary data, interpolations for altitude regridding) and with co-location mismatch, i.e. differences in sampling and smoothing of the structured and variable atmospheric field. In particular, the latter term is hard to quantify and often missing in validation studies, resulting in incomplete uncertainty budgets and improper consistency checks.

Part I: Gap description

Primary gap type:

Uncertainty in relation to comparator

Secondary gap type:

- Knowledge of uncertainty budget and calibration
- Governance (missing documentation, cooperation etc.)

ECVs impacted:

Temperature, Water vapour, Ozone, Aerosols, Carbon Dioxide, Methane

User category/application area impacted:

- Operational services and service development (meteorological services, environmental services, Copernicus services (C3S) and Atmospheric Monitoring Service (CAMS), operational data assimilation development, etc.)
- International (collaborative) frameworks and bodies (SDGs, space agency, EU institutions, WMO programmes/frameworks etc.)
- Climate research (research groups working on development, validation and improvement of ECV Climate Data Records)

Non-satellite instrument techniques involved:

- Radiosonde
- Ozonesonde
- Lidar
- FPH/CFH

- Microwave Radiometer
- FTIR
- Brewer/Dobson
- UV/VIS zenith DOAS
- UV/VIS MAXDOAS
- Pandora
- GNSS-PW

Related gaps:

- G3.01 Incomplete knowledge of spatiotemporal atmospheric variability at the scale of the measurements and of their co-location
- G3.04 Limited characterization of the multi-dimensional (spatiotemporal) smoothing and sampling properties of atmospheric remote sensing systems, and of the resulting uncertainties

G3.01. To be addressed before G3.06
Argument: To quantify the additional errors and uncertainties in a comparison due to co-location mismatch, it is advantageous to have external information on the atmospheric variability on the scale of the co-location mismatch.
G3.04. To be addressed before G3.06
Argument: To quantify the additional errors and uncertainties in a comparison due to co-location mismatch, it is important to know the smoothing and sampling properties of the individual instruments

Detailed description:

Ideally, every validation study based on comparisons with ground-based reference data should investigate whether the comparison statistics (bias or mean difference, spread on the differences, drift, etc.) are compatible with the reported random and systematic measurement uncertainties, while taking into account the additional uncertainties due to spatiotemporal sampling and smoothing differences, i.e. non-perfect co-location of the airmasses sensed by both instruments. Indeed, it is only in a few particular cases 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 spatiotemporal mismatches) well below the measurement uncertainties. In all other cases, the discrepancy between two data sets will contain non-negligible terms arising from sampling and smoothing differences, which need to be taken into account. In fact, such an analysis is essential to fully assess the data quality and its fitness-for-purpose, but in practice, it is rarely performed, as this co-location mismatch is hard to quantify reliably. Some pioneering work was published by Cortesi et al. (2007) on uncertainty budget closure for MIPAS/ENVISAT ozone profile validation, by Ridolfi et al. (2007) for the case of MIPAS/ENVISAT temperature profiles validation, by Fasso et al. (2013) in the context of radiosonde intercomparisons, by Lambert et al. (2012) on water vapour comparisons, and by Verhoelst et al. (2015) for GOME-2/MetOp-A total ozone column validation. However, no such studies have hitherto been performed for most other ECVs and/or instruments. This gap therefore concerns the need for (1) further research dealing with methods to quantify co-location mismatch, and (2) governance initiatives to include in the common practices among validation teams dedicated efforts to construct full uncertainty budgets, and use these in the consistency checks.

Operational space missions or space instruments impacted:

Independent of specific space mission or space instruments

Validation aspects addressed:

- Radiance (Level 1 product)

- Geophysical product (Level 2 product)
- Gridded product (Level 3)
- Assimilated product (Level 4)
- Time series and trends
- Representativity (spatial, temporal)
- Calibration (relative, absolute)
- Spectroscopy
- Auxiliary parameters (clouds, lightpath, surface albedo, emissivity)

Expected gap status after GAIA-CLIM:

GAIA-CLIM explored and demonstrated potential solutions to close this gap

Dedicated studies within GAIA-CLIM aimed for full error (or uncertainty) budget decomposition for representative comparison exercises, involving all non-satellite measurement types targeted by GAIA-CLIM and several current satellite sounders. Moreover, some of these results were transferred into the Virtual Observatory to allow end users to also decompose the uncertainty budget of their comparisons. Nevertheless, further work is required to quantify comparison error budgets in many cases, to operationalise comparison error budget calculations in operational satellite validation and production of higher level services, and to increase awareness in the community of the need for comparison error budget closure.

Part II: Benefits to resolution and risks to non-resolution

Identified Benefit	User category/Application area benefitted	Probability of benefit being realised	Impacts
Improved feedback on data quality from the validation work, including on the reported uncertainties.	All users and application areas will benefit from it	High	Optimized use of the data, avoiding over-interpretation but potentially also allowing greater detail to be extracted.
Tighter constraints from validation work support product development	All users and application areas will benefit from it	High	Shortcomings in products are more easily identified, driving further development and ultimately ensuring better, more reliable data products.
Identified risk	**User category/Application area benefitted**	**Probability of benefit being realised**	**Impacts**
Incomplete –or even incorrect- feedback from a validation exercise on the data quality.	All users and application areas will suffer from it.	High	Poorly quantified data quality, affecting all use types. Sub-optimal feedback to data providers slows product development. The potential of the EO system is not fully realized.

Part III: Gap remedies

Gap remedies:

Remedy 1: Use of Observing System Simulation Experiments (OSSEs)

Primary gap remedy type:

Research

Proposed remedy description:

This remedy concerns Observing System Simulation Experiments (OSSEs), such as those performed with the OSSSMOSE system by Verhoelst et al. (2015) on total ozone column comparisons. These are based on a quantification of the atmospheric field and its variability (c.f. gap G3.01), e.g. in the shape of reanalysis fields, and on a detailed description of the sampling and smoothing properties of the instruments that are being compared (c.f. gap G3.04). The aim is to calculate the error due to spatiotemporal mismatch for each comparison pair, and to derive the mismatch uncertainties from these, so that they can be added to the measurement uncertainties to derive the full uncertainty budget.

The technological and organizational challenges are mostly related to the underlying gaps G3.01 and G3.04. When these are properly addressed, the calculation of the full uncertainty budget of a comparison exercise requires only a low investment in time and resources. Integrating this into an operational validation context does constitute an additional challenge requiring dedicated effort and funding.

Relevance:

This remedy addresses directly the gap.

Measurable outcome of success:

At a high level, success is achieved when validation (and other comparison) results are published including a full uncertainty budget decomposition, taking into account spatiotemporal mismatch uncertainties. Or when they include a convincing demonstration that mismatch uncertainties are well below the measurement uncertainties and are negligible.

At a lower level, success is achieved if the OSSE allows one to close the uncertainty budget, i.e. the measured differences (or their statistics) are compatible with the sum of all uncertainty sources. Note that this requires reliable measurement uncertainties as well.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 3 years

Indicative cost estimate (investment):

Medium cost (< 5 million)

Indicative cost estimate (exploitation):

Non-applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- ESA, EUMETSAT or other space agency
- Academia, individual research institutes

Remedy 2: Statistical estimation of typical co-location mismatch effects

Primary gap remedy type:

Research

Proposed remedy description:

An alternative to estimating co-location mismatch (the main missing term in the uncertainty budget decomposition of a comparison) from model simulations, is to employ statistical modelling on the differences, for instance with a heteroskedastic functional regression approach, (as implemented for instance in the STAT4COLL software package). In certain applications, this approach also allows one to disentangle measurement uncertainties from co-location mismatch, at least for the random components. GAIA-CLIM will have employed such an approach for a subset of specific cases (spatial domains and ECVs / measurement techniques). Further efforts are required to generalise the approach and tools to enable broader exploitation, including integration into an operational validation context.

Relevance:

Employ statistical modelling on the differences, for instance with a heteroskedastic functional regression approach. Efforts are required to generalise the GAIA-CLIM approach and tools to enable broader exploitation.

Measurable outcome of success:

At a high level, success is achieved when validation (and other comparison) results are published including a full uncertainty budget decomposition, taking into account spatiotemporal mismatch uncertainties. Or when they include a convincing demonstration that mismatch uncertainties are well below the measurement uncertainties and are therefore negligible.

At a lower level, success is achieved if the statistical modelling allows one to close the uncertainty budget, i.e. the measured differences (or their statistics) are compatible with the sum of all uncertainty sources. Note that this requires reliable measurement uncertainties as well.

Expected viability for the outcome of success:

High

Scale of work:

- Single institution
- Consortium

Time bound to remedy:

Less than 3 years

Indicative cost estimate (investment):

Medium cost (< 5 million)

Indicative cost estimate (exploitation):

Non-applicable

Potential actors:

- EU H2020 funding
- Copernicus funding
- National funding agencies
- National Meteorological Services
- ESA, EUMETSAT or other Space agency
- Academia, individual research institutes

References:

- 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, J.-C., et al., “Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues”, in “Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods”, N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 pp., ISBN: 978-1-4614-3908-0, DOI 10.1007/978-1-4614-3909-7_2, © Springer New York 2012
- Ridolfi et al., “Geophysical validation of temperature retrieved by the ESA processor from MIPAS/ENVISAT atmospheric limb-emission measurements”, ACP v7, 2007
- Verhoelst et al., “Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons”, AMT v8, 2015