Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

WP5: Creation of a "Virtual Observatory" visualisation and data access facility

D5.7: "Report on the Evaluation of the Virtual Observatory"

A Horizon 2020 project

Grant agreement: 640276

Date: 8 March 2018

Lead Beneficiary: EUMETSAT

Nature: R

Dissemination level: PU

Work-package	**WP5 (Creation of a "Virtual Observatory" visualisation and data access facility)**
Deliverable	D5.7 Report on the Evaluation of the Virtual Observatory
Nature	R
Dissemination	PU
Lead Beneficiary	EUMETSAT
Date	08/03/18
Status	Final
Authors	Stephanie Guedj (EUMETSAT), Marie Doutriaux-Boucher (EUMETSAT), Tijl Verhoelst (BIRA), Jean-Christophe Lambert (BIRA), Neeme Loorits (TUT), Paul Green (NPL), Tom Gardiner (NPL), Jacques Descloître (ICARE), Pekka Kolmonen (FMI), Arndt Meier (EUMETSAT)
Reviewers	Jörg Schulz (EUEMTSAT), Peter Thorne (NUIM), Anna Mikalsen (NERSC)
Contacts	Joerg.schulz@eumetsat.int
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.

Version	Author(s) /Reviewers	Date	Changes
0.1	Stephanie Guedj (EUMETSAT), Marie Doutriaux-Boucher (EUMETSAT)	07/02/2018	1st draft
0.8	Stephanie Guedj (EUMETSAT)	19/02/2018	Merging reviews and important changes in section of Evaluation of the VO/BT Level 1 data
0.9	Stephanie Guedj (EUMETSAT)	21/02/2018	Complete new sub-sections and executive summary
1.0	Jörg Schulz (EUMETSAT), Peter Thorne,(NUIM), Anna Mikalsen (NERSC)	07/03/2018	Final

Executive summary

This report presents an initial quality evaluation of the final status of the Virtual Observatory (VO) developed during the GAIA-CLIM project to provide a starting point for further evolution after the project. The assessment demonstrates the achieved utility for scientific/statistical analysis of respective observations and the monitoring of instrument and product behaviour over time.

This assessment has shown that the VO provides a good demonstrator for the GAIA-CLIM outputs in terms of providing and displaying uncertainty estimates for very different types of atmospheric variables and the comparison of those derived from satellite and non-satellite data. The addressed variables represent fairly complex comparison set ups, e.g., temperatures and humidity are measured by radiosondes that are moving targets that need to be co-located with a satellite snapshot.

Overall, the VO concept has been well received as a proof of concept by the various attendees of the project roadshows as summarised in D6.9. A major achievement is the accessibility of co-located data containing all available uncertainties for download in an easy to use format. This functionality has a great potential as it may save time of many investigators that do not now need to download all data and compute the co-locations themselves. The 3D-metadata viewer is supporting the VO in such a way that a user can also look into the future using an orbit propagator and can plan the usage of specific stations with expected close co-locations.

The comparison of the GAIA-CLIM approach using ground-based to satellite data comparisons to other means, such as satellite-satellite data and NWP output-satellite data comparisons, has shown that all three are needed. The use of ground-based data ensures that the satellite data are compared to traceable standards while the other comparisons deliver robust statistics and in case of NWP outputs also global coverage. These comparisons also demonstrate that the co-location mechanism itself can still further be improved to achieve lower systematic differences and attendant uncertainties in the resulting comparisons.

The biggest shortcoming of the current VO is that no single comparison is metrologically complete containing all of non-satellite measurement uncertainty, satellite measurement uncertainty and co-location uncertainty. Thus, a big improvement potential exists for further development of the VO including the implementation as an operational facility in the future.

Contents

1 Introduction

1.1 Background

Satellite data are an essential tool to monitor and understand our changing climate system. Satellites provide global and continuous observations of many different components of the climate system that can be remotely sensed from space. They are utilised in many applications such as climate monitoring and other monitoring services, weather forecasting and supporting emergency services. To extract the full value of the satellite data, it is required that the quality of satellite measurements is continuously monitored and well understood. One key component of this is the comparison to co-located non-satellite reference observations.

The Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring (GAIA-CLIM) project aims to improve our ability to use ground-based and sub-orbital observations to better characterise satellite observations. The project focussed upon a small set of atmospheric Essential Climate Variables (ECVs). Work within the GAIA-CLIM project focussed upon geographical characterisation of non-satellite measurements, improving their metrological understanding, better quantifying co-location impacts on the uncertainty budget and the role of data assimilation systems as integrators. This underlying scientific work was integrated and presented via a Virtual Observatory (VO) facility, which serves as a demonstrator of potential utility of the non-satellite segment as a long-term calibration/validation tool for satellite measurements in future.

1.2 Purpose and scope

This document addresses the evaluation of the capabilities of the VO to provide a starting point for further evolution after the project. The assessment demonstrates the achieved utility for scientific/statistical analysis of respective observations and the monitoring of instrument and product behaviour over time. In particular an analysis for the monitored Level 1 satellite data is performed to document the value of the comparisons of non -satellite data for the characterisation of Level 1 data vis-à-vis more traditional approaches of: i) satellite-to-satellite match-ups; and ii) NWP model data comparisons. In addition, the functionalities and suitability of the VO for the quality assessment of the satellite retrievals for ozone total column contents and aerosol optical depth are analysed. This is accompanied with an evaluation of the VO in terms of the principle of the Quality Assurance (QA) methodology as developed in the EU FP7 QA4ECV project.

1.3 Structure of the document

This document is structured into sections that describe the evaluated parts of the VO and the evaluation results obtained. In addition, there is a QA assessment of the VO for the representation of the uncertainty for the addressed variables.

- Section 1 (this section) provides an introduction
- Section 2 is a brief overview of the VO including information on the accessible data and an evaluation of what additions could be possible based upon GAIA-CLIM project outcomes;
- Section 3 outlines the results of a qualitative evaluation of the VO final capabilities;
- Section 4 discusses the QA assessment results using the approach developed in the EU FP7 QA4ECV project;
- Section 5 provides a summary;
- Annex A provides a list of references;
- Annex B provides a list of acronyms.

2 Overview of the Virtual Observatory

The VO enables users to discover, select, interrogate, extract, visualise and analyse satellite to non-satellite data comparisons including all available relevant uncertainties for such a comparison. Currently, the VO considers comparisons for the following ECV products:

- Temperature and humidity profiles represented in brightness temperature space,
- Ozone total columns, and
- aerosol optical depth.

These variables can be derived from various satellite measurements operating in the visible, infrared and microwave spectral regions. The satellite estimates are compared to ground-based measurements from various networks.

The VO takes into account the systematic and random uncertainties of each of the measurements as far as available, the smoothing uncertainties (random and systematic), and the co-location mismatch uncertainty due to atmospheric variability, depending on the distance in time and space between the compared measurements that form a co-located observation pair. All these uncertainties and their basis have been described in detail in other GAIA-CLIM documentation (cf. Annex A).

The VO includes a 3D-data discovery tool as part of the Graphical User Interface (GUI) that supports the search for matching reference and satellite observations. The 3D-metadata tool has been designed to read the metadata from its own database, check for the availability of specific observations or observation locations, and, if data are available, visualise the data interactively through the GUI of the VO. Owing to time restrictions, this capability to link across to specific observations in the VO database was not realised within the project, but the potential exists to do so in future.

The VO is intended to provide the user with access to both metadata and observational data from different ground-based reference networks with co-located satellite data. Table 1 provides a list of available and accessible data in the VO at the end of the GAIA-CLIM project. The evaluation performed in this document addresses the simulated brightness temperatures for the HIRS instrument, the total column ozone estimates and the aerosol optical depth.

The actual VO database contains more data than listed in Table 1, e.g., co-locations for integrated water vapour computed from GCOS Reference Upper-Air Network (GRUAN) radiosondes and derived from the GOME-2 instrument data have been stored, but the GUI could not be updated to provide the graphical representation and data access prior to the project cessation. In addition, many other reference data have been addressed by the project that could not be integrated in the project's life time, e.g., ground-based radiometer water vapour measurements, etc. Also, on the satellite side, the number of satellite instruments that can be simulated using the GRUAN processor is much larger than showcased in the VO. For instance, data from microwave sounding and hyperspectral instruments such as IASI have not been integrated into the VO, although spatial mismatch uncertainties have been derived for IASI/GRUAN co-locations. This is entirely due to limited project resources and provides a great potential for future extensions of the VO.

Table 1: Data accessible from the Graphical User Interface of the VO.

ECV / Type	**Method**	**From**	**To**	**No of co-locations**	**No of stations**
Brightness temperature / Co-location	ECMWF vs HIRS on Metop-A	01.01.2013	31.12.2016	33 096	16
Brightness temperature / Co-location	UK MetOffice vs HIRS on Metop-A	01.01.2013	31.12.2016	51 816	17
Brightness temperature / Co-location	GRUAN Processor (ECMWF) vs HIRS on Metop-A	01.01.2013	31.12.2016	33 038	16
Brightness temperature / Co-location	GRUAN Processor (UK MetOffice) vs HIRS on Metop-A	01.01.2013	31.12.2016	51 674	17
Brightness temperature / Co-location	ECMWF vs HIRS on Metop-B	01.01.2013	31.12.2016	21 432	16
Brightness temperature / Co-location	UK MetOffice vs HIRS on Metop-B	01.01.2013	31.12.2016	21 720	17
Brightness temperature / Co-location	GRUAN Processor (ECMWF) vs HIRS on Metop-B	01.01.2013	31.12.2016	21 411	16
Brightness temperature / Co-location	GRUAN Processor (UK MetOffice) vs HIRS on Metop-B	01.01.2013	31.12.2016	21 689	17
Brightness temperature / Co-location	ECMWF vs HIRS on NOAA19	01.01.2013	31.12.2016	5 716	16
Brightness temperature / Co-location	UK MetOffice vs HIRS on NOAA19	01.01.2013	31.12.2016	23 540	17
Brightness temperature / Co-location	GRUAN Processor (ECMWF) vs HIRS on NOAA19	01.01.2013	31.12.2016	5 708	16
Brightness temperature / Co-location	GRUAN Processor (UK MetOffice) vs HIRS on NOAA19	01.01.2013	31.12.2016	23 497	17
Aerosol optical depth / Co-location	AERONET Sunphotometer vs AATSR on Envisat	01.07.2002	08.04.2012	30 086	527
Ozone / Co-location	NDACC DOAS vs GOME-2	23.01.2007	30.09.2017	12 710	2
Ozone / Co-location	NDACC FTIR vs GOME-2	29.01.2007	21.09.2017	17 409	10

3 Evaluation of the VO

The VO development and deployment was largely undertaken in parallel with the work on the underlying scientific work packages. Although the underlying work packages, broadly speaking, met the expectations and the timelines set out in the project plan, such parallel development and deployment was always going to be challenging. Inevitably, several potential tools, data and approaches did not make it to the final version of the VO. This document therefore begins with a qualitative evaluation of the VO functionality in the context of the broader aims and achievements of the project. In each section a list of potential recommendations for improvement based on partners evaluation of the final VO functionality are listed. These recommendations are based upon discussions at the final General Assembly and in various virtual meetings of the consortium. They are also informed by feedback on earlier versions of the VO used in the series of roadshow events facilitated by the outreach work package, of which the VO was a key component (D6.9).

The remainder of this section touches upon key salient features by going through each set of demonstrator use-cases. In each case key salient features and capabilities are highlighted and then potential limitations and future improvements outlined. Some of the suggestions apply across all use-cases considered. For brevity, these are highlighted only once, in the most appropriate sub-section. Where possible, recourse is made to comparison to current techniques for satellite characterisation to enable a comparison of potential utility.

We would note that, unfortunately, none of the developed use-cases are complete in containing all three of: an uncertainty estimate on the non-satellite measurement; an uncertainty estimate on the satellite measurement; and a comprehensive estimate of the co-location uncertainty. One or more of these estimates are, nevertheless, available in each of the demonstrator capabilities. As noted in Section 4, such a fully closed comparison should be possible shortly, and its inclusion would add substantive value.

3.1 Evaluation of the 3D-matadata tool

The 3D-metadata viewer is a data discovery tool which is accessible from the VO's GUI. It provides a 3D geographical mapping of existing non-satellite measurement capabilities (in-situ surface, in-situ sounding, columnar and profiling observations) for an extended set of ECVs compared to those accessible from the VO. In particular, networks for additional atmospheric composition ECVs, as well as temperature and salinity profiles in the ocean, are available.

The tool provides geographical mapping of the measurement sites and provides specific information for each (Figure 1). This includes access points to the different stations where the user can access more information and the full reference data sets. In addition, a maturity assessment[1] of the reference networks is presented that enables users of reference data to find the right data for a comparison exercise, in particular for long term quality monitoring.

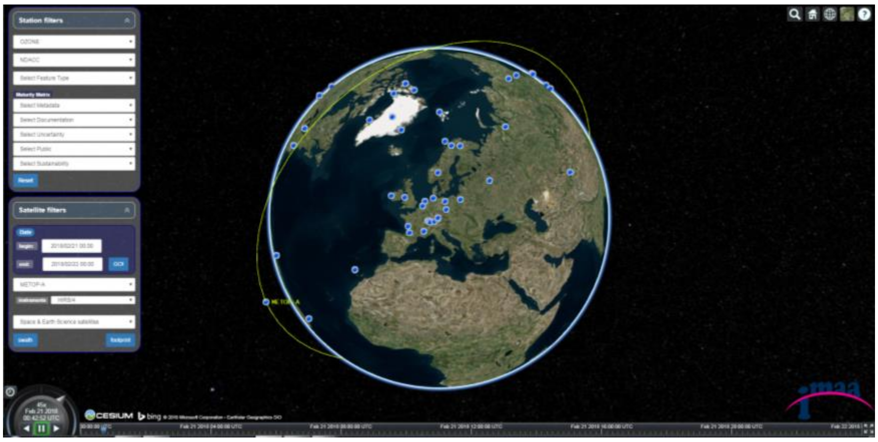

Figure 1: Screenshot of the 3D-Tool geographic capabilities illustrating locations of ozone-network stations performing columnar and profiler ozone measurements and the orbit and position of the HIRS instrument on the MetOp-A satellite.

[1] For more information on the Maturity Matrix Assesment developed within GAIA-CLIM to assess various quantifiable facets of the maturity of a measurement system, see: http://www.gaia-clim.eu/page/maturity-matrix-assessment

The tool also provides a satellite overpass mapping tool for some relevant satellites in the context of the considered ECVs. The overpass mapping tool allows a first assessment at what times of the day specific ground stations have overflights which may be used to get an impression how big the spatial and temporal mismatch will be. The tool will be sustained by CNR for the foreseeable future and a detailed description of the capabilities can be found in D1.9.

One item for improvement for climate applications would be a better treatment of the past data availability. The network configuration is static and not shown as a function of time. If a user wanted to examine available comparison data in the 1980s or 1990s, the tool does not provide the right network configuration and also more historical satellites and their orbits would need to be added.

Also, a true integration with the VO GUI could be addressed by enabling data selection, e.g., a specific site or a whole network from the 3D-metadata tool. This would work in such a way that selecting a station leads automatically to a population of the VO data selection fields and extraction from the VO database. Such capability would be possible but was not able to be realised within the timeframe of the project, although aspects of this functionality were enabled.

3.2 Level 1 HIRS radiances compared to simulated radiance from GRUAN radiosondes

3.2.1 Data selection and access

Approximately 4 years of data (2013-2016) are available through the VO for GRUAN radiosonde stations. These data have undergone an extensive traceability and quality verification within GAIA-CLIM. The GRUAN radiosonde data are converted into HIRS instrument radiances for 5 channels with weighting functions peaking between the lower stratosphere and the upper troposphere. In addition, these co-locations are enriched with forward simulations using output from two NWP models, UKMO and ECMWF, respectively. Users of the VO can select stations, time periods and sub-selections of the co-locations using more stringent criteria compared to the default. All selected co-located data can be easily downloaded in NetCDF format.

Based on the project partner's evaluation the following improvements could be undertaken in future:

- The uncertainty budget for the comparison at Level 1 is incomplete due the missing uncertainty estimate in the satellite data (supposed to come from the FIDUCEO project) and only the total uncertainty of the radiosonde measurements is propagated through the GRUAN processor, which could also be improved in the future;
- The provision of the simulated HIRS data from the NWP models has not been possible to be completed for all GRUAN stations, which should be done to allow the same analysis at all stations.

3.2.2 Measurement co-location

The GRUAN data are co-located with the satellite data by identifying all satellite overpasses that are in the vicinity of the radiosonde location at a pressure level of 300 hPa. Then a wide space (500 km) and time window (±3 hours) is used to extract co-locations for the VO.

Further possible improvements regarding the measurement co-location are:

- The co-location of the GRUAN data to the satellite data is done at a fixed reference pressure of 300 hPa, which is not optimal in particular for the HIRS water vapour channels where the vertical level of the emission depends on the water vapour loading of the atmosphere. More sophisticated methods should be developed to make the comparison as optimal as possible.

3.2.3 Visualisation

On the data selection page, a list with station names, start and end dates of the available records and number of co-locations is available to guide the selection of a specific station for analysis. After selection of the co-locations, the VO displays two panels showing time series of (a) the compared measurands including bars indicating their total uncertainty and (b) the difference between the HIRS radiance data and a selectable simulation, i.e., from GRUAN data or the NWP data. These panels can be displayed for all 5 simulated HIRS channels. Visualisation of single co-location statistics is also available, presenting bar charts of brightness temperature and relative differences for all channels and available data sources.

Identified potential improvements to visualisation are:

- The list of stations on the data selection page is very good but should be enhanced for traceability by the metadata information about the station (imported from the 3D-metadata viewer), i.e., geographical coordinates, height, etc. This could be accompanied by a geographical map;
- The co-location distance selector for time and space should also be available on the graphical display page to see the immediate effect of changing the co-location criteria;
- Additional selections should be added, e.g., latitude and continental discriminators allowing the user focussing on studies in specific areas;
- More summary information should be available for plotting, e.g., histograms for each station and for sets of stations and temporal integration, e.g., to seasonal averages would be useful;
- The plots themselves could exhibit more statistical information on averages, median, variance, etc.

3.2.4 Outline of GAIA-CLIM approach and data used in evaluation

The World Meteorological Organisation (WMO) facilitates initiatives for achieving the goal of providing well characterised data to users for climate monitoring. The Global Space-based Inter-Calibration System (GSICS) and the GRUAN are prime examples for such initiatives aiming at the provision of data and tools for the characterisation of satellite data.

Comparing satellite measurements (radiances emitted by the Earth and its atmosphere at certain wavelengths, usually expressed as brightness temperature) with reference ground-based GRUAN radiosonde measurements is not trivial. In order to achieve this, either the radiance have to be converted to geo-physical variables (e.g., temperature and humidity profiles) or the geo-physical variables measured by reference instruments have to be converted to radiances that would be measured by a satellite instrument under those surface and atmospheric conditions. The former is an ill-posed problem and the latter provides means for a more like-to-like comparison.

Therefore, the GAIA-CLIM project has developed the GRUAN processor (Carminati et al., 2016), which uses the radiative transfer model RTTOV (Matricardi, 2009) to simulate satellite radiances. RTTOV requires atmospheric as well as surface variables to simulate satellite radiances, but the GRUAN data set does not contain all such required variables and the missing information is taken from NWP model outputs.

In this section, a quantitative comparison of the two systems (space and ground-based) for the characterisation of HIRS instrument data is provided. In addition, pure simulations of the satellite radiances using the two NWP models employed in the context of GAIA-CLIM are added to this comparison.

The GRUAN measurements are SI traceable and their uncertainties are well characterised. However, the number of available reference profiles is low in space and time. GRUAN has 15 sites (mainly located in the Northern Hemisphere) and for this comparison co-located data for 6 GRUAN stations are used (Table 2). Because the radiative transfer is only accurate enough in cloud free conditions, a cloud detection scheme (Kottayil et al., 2012) has been applied to the HIRS data and only cloud free co-locations are retained. As can be seen in **Error! Reference source not found.**, only a fraction (ranging from 6% in Ny Alesund to 29% in Lamont) of the co-locations at each station is cloud free.

Table 2: Number of co-locations for radiosondes and HIRS measurements for six GRUAN stations for the period 2013 to 2016 and the fraction of clear sky cases for those co-locations.

	Barrow	Cabauw	Lamont	Lindenberg	Ny Alesund	Sodankyla
All (2013-2016)	5302	1572	3519	8545	7533	5015
% clear	15.64%	15.20%	29.01%	11.53%	6.31%	10.07%

3.2.5 Relative value of the reference data for satellite evaluation/calibration

The GSICS approach is to perform satellite to satellite comparisons. One of the satellite instruments is the monitored instrument in this case the (HIRS) and the other one is considered to be the superior quality reference instrument (in this case the Infrared Atmospheric Sounding Interferometer (IASI)). Since both considered instruments are on-board EUMETSAT's Metop-A polar-orbiting satellite, there are several thousand co-locations every day with global coverage. This comparison presents an ideal case for a satellite-satellite comparison. Co-locations between instruments on-board different polar-orbiting satellites (e.g., HIRS on-board NOAA-18 and IASI on-board Metop-A) occur rather intermittently and only over high latitudes (cold atmosphere and surface), as explained in Cao et al (2004). In such cases, there may not be enough co-locations available to produce robust estimates of bias between the considered instruments. Such comparisons have also the issue that they do not cover the dynamic range of measurements, e.g., do not allow a characterisation at warm temperatures.

Figure 2 shows the comparison of both instruments for the year 2013 for HIRS channel 3 (CO2 absorption used for temperature estimates in the upper troposphere) and channel 12 (water vapour absorption used for humidity estimates in the upper troposphere). Overall, the HIRS measurements show good agreement with the IASI measurements with differences less than ±0.5 K for both channels and little variation over time. For channel 3, distinct larger deviations are visible after July 2013, for which the cause is unknown as both instruments did not show instrument anomalies during this period.

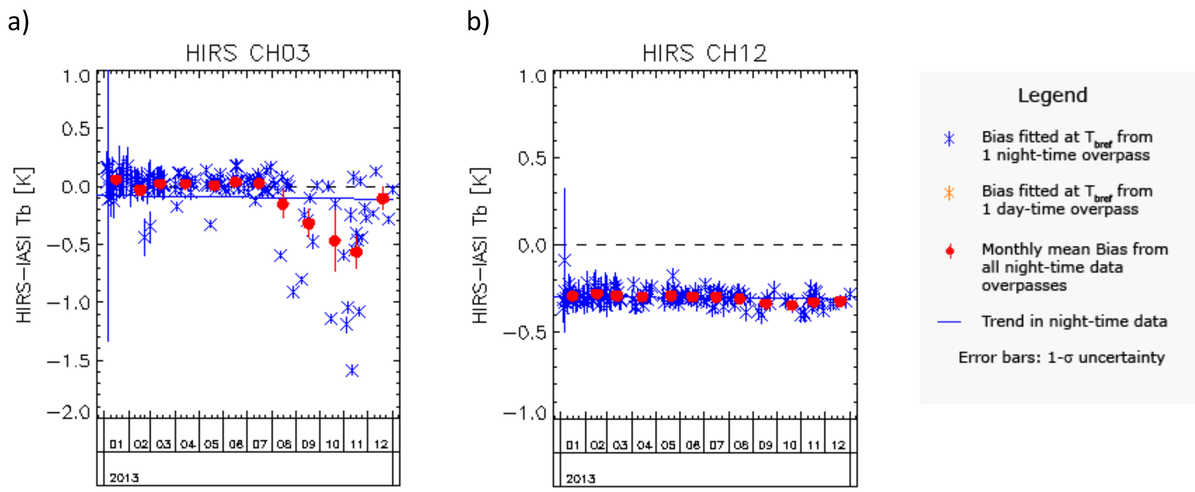

Figure 2: GSICS Inter-calibration of HIRS channel 3 and 12 using IASI hyperspectral measurements. Figure taken from http://tcweb.eumetsat.int/tcc1/proj/gsics/web/BiasMonitoring.html.

Figure 3 shows the comparison of co-located Metop-A HIRS measurements and simulated HIRS measurements from GRUAN profiles and UKMO model outputs arising from GAIA-CLIM anad available through the VO. Details of the co-location method for creating the HIRS-GRUAN match-up is described in the GAIA-CLIM deliverable D5.9. The two selected channels (3 and 12) are the best to compare between HIRS measurements and GRUAN simulations because most of the signal originates from emission in the upper troposphere or in the lower stratosphere, and thus comparisons are not unduly affected by unknown surface parameters and the missing mid-stratospheric information in the GRUAN profiles. So, it is important to note that these kinds of comparisons cannot be used for monitoring and characterising all HIRS channels, for example the surface or lower tropospheric sensitive channels, on infrared sounding instruments.

Monthly mean and standard deviation of the differences are shown in Figure 3. Also shown are the monthly mean difference estimated for HIRS against the reference sensor IASI. HIRS observations show significant bias against both GRUAN and UKMO simulated radiance. For channel 3, the bias is larger than 1 K for certain time periods. The simulated radiance from the GRUAN data and the UMKO NWP output agree with each other, which is expected as over land NWP model output is tuned towards radiosonde measurements (although note that it is the manufacturer processed data and not GRUAN data ingested into the NWP systems). This is significantly different from the GSICS monitoring results where the bias is much smaller. Results for channel 12 show even larger differences, in some months up to -5 K which is an order of magnitude larger than for the GSICS monitoring results.

One of the major issues in the comparison between HIRS and GRUAN is the lack of sufficient co-locations at a reasonable time scale, e.g., monthly. Depending on the GRUAN station locations, there are only 10-40 collocations available per month as shown in the bottom panels of Figure 3, which are not always sufficient to compute robust estimates of biases and may result in spurious variability in the bias time series as evident from the comparison against GSICS results.

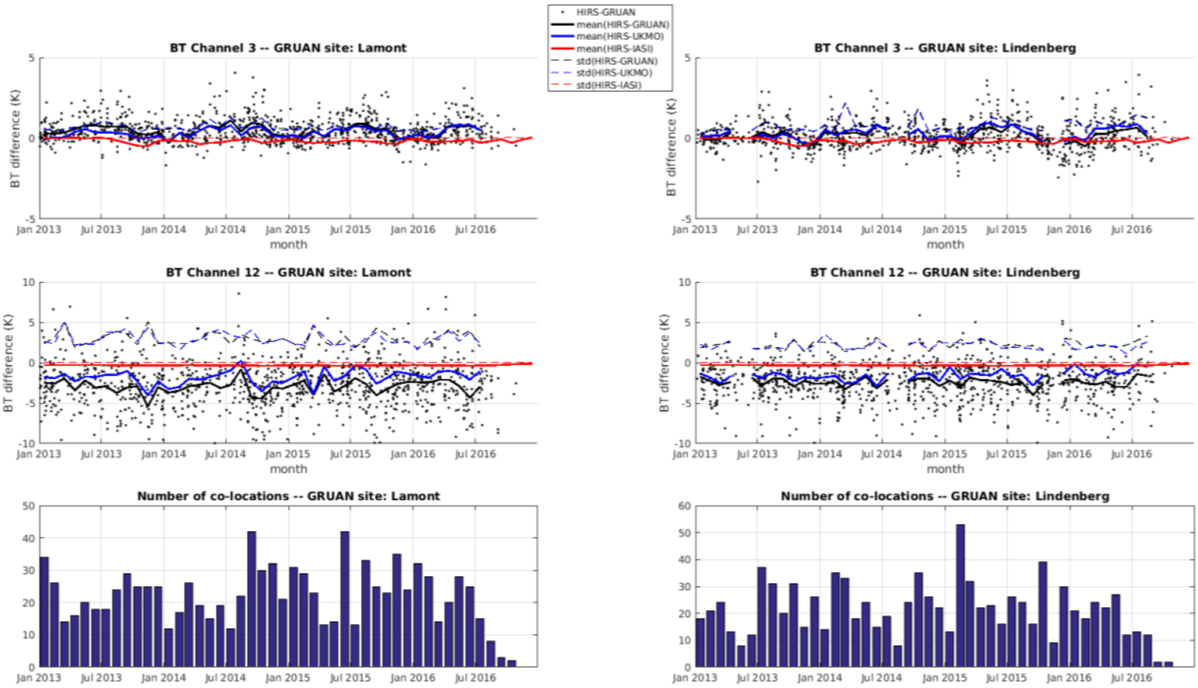

Figure 3: Time-series of monthly bias (solid lines) and standard deviation (dashed lines) at Lamont and Lindenberg between co-located HIRS Brightness Temperature (BT) channel 3 and 12 against GRUAN radiosondes propagated into brightness temperature space (black lines) as well as NWP BT simulations using UKMO NWP model output (blue lines). GSICS determined biases and standard deviations between HIRS and IASI are also plotted (red lines). Black dots represent the difference for each pair of HIRS BT minus GRUAN BT. Bar plots display the amount of co-locations available to compute the monthly mean. If the number of co-locations drops below 10 no monthly mean is computed.

Another approach to evaluate satellite measurements is by the use of NWP model outputs as already discussed. In contrast to pure satellite-satellite or ground-based-satellite method as described above, co-locations are globally distributed (Figure 4a). However, the differences (also called first-guess departures or Observation minus Background) include biases arising from mixed information: 1) instrumental anomalies, 2) mis-representation of the atmosphere/surface in the model, and 3) errors in the radiative transfer model used to simulate the radiances.

Figure 4b shows observations minus simulations for a set of infrared instruments. One can note that, except for HIRS ozone channel 9 that could not be efficiently simulated by NWP models, biases are generally within ±1.5K. With the exception of HIRS channels 1, 13, and 16, biases are quite similar for each instrument. Channels showing a bias of a similar amplitude for all instruments indicate that this bias is more likely to come from the NWP system. On the contrary, channels that carry inconsistent bias for each instrument underline instrumental anomalies that may require improvement in the calibration. However, changes in NWP systems can significantly affect the Observation minus Background (O-B) statistics. The operational NWP systems are regularly updated and therefore long-term radiance changes may not be easily detected unless a reanalysis output is used.

A summary assessment of the three methods to characterise satellite data is provided in Table 3 using HIRS/4 on-board Metop-A as the monitored instrument. The bias estimates for the comparisons to IASI and the 6 GRUAN radiosonde stations are for the time period 2013-2016. The bias estimates for the NWP data are taken from Saunders et al. (2013) using data from 2010 and only over ocean, and thus represent a favourable case for radiance simulation. Satellite-satellite comparison has an upper hand in terms of bias and its very low variability because both instruments are on the same satellite and thus resulting in thousands of co-locations every day with minimal co-location mismatch effects and very robust statistics. Biases from the NWP output are similar for channel 12 but significantly higher for channel 3. Despite a very large number of co-locations, the standard deviations of the O-B statistics are higher compared to the satellite-satellite approach. This is due to the fact that the NWP system is not fully representing the atmosphere and surface which the satellite is measuring and the comparison procedure is rather complex. The standard deviations are much larger for surface sensitive channels which are not shown here. As discussed above the direct comparison with GRUAN data has largest biases with significant temporal variability.

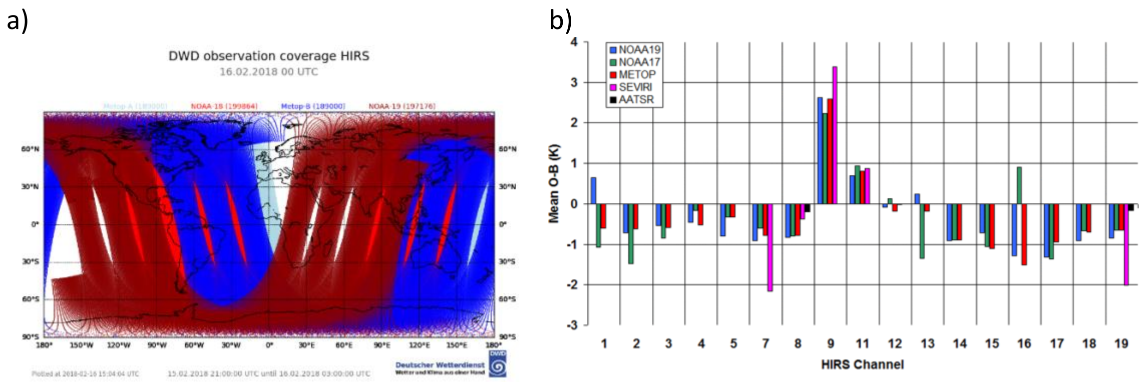

Figure 4: a) Observation coverage from HIRS on board Metop-A/B, NAOO18/19 on the 16/02/2018 at 0UTC. b) Global mean O–B biases of HIRS channels over the sea for 2010. The shortwave channels are only for night-time data. The corresponding SEVIRI- and AATSR-channel biases are also plotted for those channels which are common in wavelength (Saunders et al., 2013).

Table 3: Summary table on the methodologies used to validate/characterise satellite radiances. NWP results are taken from Saunders et al., 2013.

Criteria / Reference	Co-locations (time/space)	Traceability of uncertainties	Complexity of comparison	Robustness of statistics	Bias (clear sky) (HIRS – REFERENCE)
Satellite (IASI)	1000s/day, global[1]	Medium (in case of hyper-spectral measurements)	Low	High	CH03 = -0.13±0.08 K CH12 = -0.18±0.05 K
NWP (UKMO)	1000s/day, global	Low	Medium	Medium	CH03 = -0.59±0.25 K CH12 = -0.17±1.65 K (Saunders et al., 2013)
Radiosondes (GRUAN)	A few/day, local (NH)	High (only in reference measurements)	High	Low	CH03 = -0.47±0.87 K CH12 = -2.30±2.41 K

[1]this applies only for the HIRS-4 to IASI comparison on Metop-A/B, polar SNOs would have hundreds of co-locations per day, when such events happen, with slightly different bias estimates resulting in Medium robustness of statistics.

However, both the use of the GSICS and the NWP model outputs do not provide a full traceability for the uncertainty estimates of the reference datasets. Reference measurements such as from the GRUAN network provide fully traceable uncertainty characteristics, but the full potential of these uncertainty characterisation cannot be easily utilised for the evaluation of the satellite data because of their mismatch in representativeness with satellite data. That is to say that the satellite measurement is a snap shot, while radiosonde measurements take several tens of minutes to measure the same air mass vertically. Similarly, radiosonde measurements are drifting point measurements horizontally but satellite measurements are area averages over footprints of several kilometres. This combined with low number of co-locations are leading to apparently low robustness in the derived statistics for GRUAN-satellite comparisons.

Nevertheless, GRUAN data can be used for characterising NWP systems and then the well characterised NWP systems may be used to characterise satellite data because NWP outputs and satellite data can be co-located globally resulting in thousands of co-locations every day. This has also the advantage that a fully traceable uncertainty chain can be established and can be traced back to the reference measurements, once uncertainties of each individual step in the comparison are estimated. This work was showcased in Work Package 4 of GAIA-CLIM and is documented in the suite of deliverables arising therefrom, specifically D4.4 and D4.7.

3.3 Ozone Total Column Amount

3.3.1 Data selection and access

More than 10 years of data are available through the VO for ground-based stations with a DOAS instrument at 2 sites and FTIR at 10 sites. These data sets have undergone an extensive traceability and quality verification within GAIA-CLIM, although the final assessment is that they have not quite attained reference quality (D2.8; and caveats on the data accessible through the VO are described in the VO User Guide D5.9). The reference measurements are co-located to data from the 2nd release of the reprocessed satellite ozone total column data record from the EUMETSAT AC SAF based on EUMETSAT Metop-A and B GOME-2 measurements[2]. Users can select stations, time periods and sub-selections based upon co-location distance in time and space. By their nature the measurements at the ground require a clear-sky view. Users can download the selected co-location data sets in NetCDF format.

Based on project partner’s evaluation further improvments could be done to the VO GUI:

- Need for improved provenance tracking on the data sources;
- The possibility to export co-located ozone data is implemented. Ideally, this would offer the user both the original data files and the harmonised format used within the VO, including the co-location specific information (distance, time difference, estimated uncertainty terms due to co-location mismatch).

3.3.2 Measurement co-location

The data accessible through the VO are co-located if matched to one another within 500km and 6h. For each pair, the actual separation in linear space and time are reported in the co-location overview panel. Further filtering of the co-locations can be done by the user of the GUI both in linear distance and in time difference using user-defined limits.

Further possible improvements regarding co-location filtering are:

- It is currently unclear whether measurements are ensured to be used only once in the resulting data set. Tick boxes ensuring “unique measurements” (separately for both ground instrument and

[2] https://acsaf.org/datarecords/o3_vcd.html

satellite) would be a desirable option to avoid the introduction of unwanted correlations in the comparison results;

- Further filtering options allowing the user control over quantities of influence such as solar zenith angle, satellite viewing angle etc. would be useful and may help the user identify issues in the FTIR and DOAS data.

3.3.3 Visualisation

Two graphics panels (Figure 5) are currently produced by default following data selection: (a) the time series of co-locations with uncertainties and (b) the time series of satellite-FTIR or satellite-DOAS differences. In the latter panel, the user can underlay different shaded regions representing different uncertainty components, including those related to co-location mismatch, and separating random and systematic measurement uncertainties. More detailed descriptions of the uncertainty components are provided at the bottom of the VO page. This 2nd panel in particular represents a major advancement with respect to the state-of-the art in online ground-satellite comparison platforms and the visualisation is clear and easy to interpret.

Identified potential improvements to visualisation are:

- For full traceability, it would be good to have more details on the ground-station (latitude, longitude, altitude above sea level) at the top of the graphs;
- More quantities could be added to the current graphs, such as mean and median values in the “differences” graph, running means or medians on both graphs (e.g. 3-monthly, to highlight seasonal features), linear drift estimates on the difference graph, etc.;
- An additional panel could be introduced to visualise the distribution (histogram) of the differences, with indications of the expected range based on the uncertainty information. This would, for instance, allow a user to interpret the significance of outliers. Also, correlation graphs could be introduced;
- Beyond these per station graphs, it would be interesting to present to the user some information on the comparison statistics at the network level. For instance: mean satellite minus non-satellite differences as a function of station latitude;
- The quantitative analysis that can be done with the VO is currently limited to a computation of mean values of the total ozone column time series, for both satellite and ground-based data sets, shown as horizontal lines within the left-hand graph. This could be extended to include more diagnostics such as the median, the standard deviation or variance, etc. Similar diagnostics should also be computed for the differences.

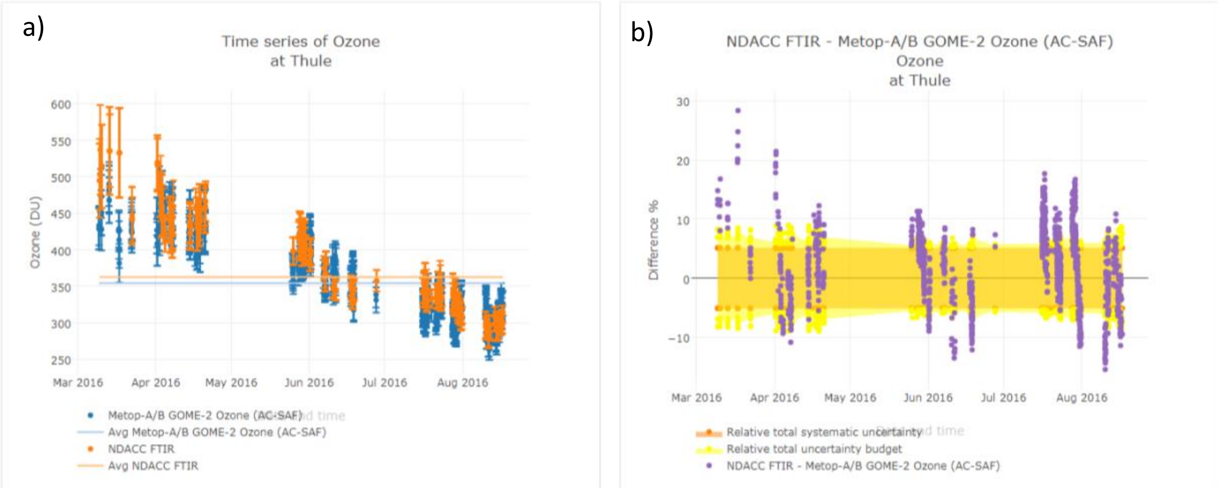

Figure 5:5: Screenshot of time-series produced into the VO at Thule for Ozone a) Ozone measurements from both satellite and ground-based observations and b) Relative differences (in percent) between satellite and ground-based with total uncertainties.

3.4 Aerosol Optical Depth

3.4.1 Data selection and access

The VO allows the comparison of satellite based AOD (aerosol optical depth) product to individual sites and visualisation of the results with uncertainty information included. The GAIA-CLIM VO provides co-located AOD at 550 nm retrieved from the ESA Envisat/AATSR satellite measurements and from NASA's ground-based reference measurements AERONET. The user can access high quality and traceable AOD observations from up to 527 AERONET sites. These observations were co-located with the Envisat AATSR (up to 15 km and 30 min providing closer match-ups compared to other cases owing to the smaller spatiotemporal correlation scales of aerosols). The data were produced at ICARE[3], where the co-location with satellite data from the Envisat was computed and validated. The sample size of the available data for each site varies from 1 to more than 300 measurements over a time period of 10 years (2002-2012) for most stations. Time series of AOD estimates and relative differences of satellite to non-satellite measurements are accessible via the VO. The user can vary the co-location maximum distance and time to create even closer match-ups. The export tool permits to download the co-located data that includes the uncertainty information and run further analysis offline. This is a clear advancement compared to similar existing platforms. The capability to visualise different uncertainty types for two co-located measurements, separately or altogether, is very unique and usually not found in online validation tools.

Potential recommendations regarding data selection and access, in addition to those articulated in prior sub-sections are:

- Currently, the user cannot select several sites simultaneously. A typical (satellite validation) user might want to select all available AERONET sites (or a subset of them), find all co-locations fulfilling the selected co-location parameters, and select the validation metric to be applied, e.g. the correlation coefficient. The user could then play with the co-location parameters and see how they affect, e.g. the number of co-locations for the whole (sub)set of AERONET sites.

---

[3] http://www.icare.univ-lille1.fr/

3.4.2 Measurement co-location

The principle of the co-location is to identify all AATSR overpasses that fall in the vicinity (defined by a given radius around the site) of the location of an AERONET instrument that performed at least one measurement within a given time range around the overpass time. For each co-location, one unique value (average) of satellite and reference AOD is provided.

Within the chosen window, there are usually multiple measurements for each co-location that are averaged, even if the presence of clouds can lower the number of coincident measurements (Breon et al., 2011). The time-space co-location criteria are based on a study with various thresholds for space and time windows performed by Virtanen et al. (2018) as part of GAIA-CLIM.

Further possible improvements regarding the measurement co-location approach are:

- The standard deviation is a good indicator for precision if several samples are available. If too few comparison samples are available the precision cannot be computed. The data integrated into the VO database contain every single sample, but the user of the VO has no information about the sample size of the individual co-location. More information should be displayed in the VO and data filtering as function of sample size could be implemented as well;
- The spatial and temporal standard deviations partly take into account the mismatch error, respectively the distance between the AERONET location and the AATSR pixel centers, and the difference between the AATSR overpass time and the time of the AERONET measurements. However, the standard deviations do not account for the mismatch error due to the instrumental observation geometries. The satellite AOD is integrated along the line of sight to the ground while AERONET AOD is integrated along a direct path to the Sun. The air masses sounded by the respective lines of sight are different. The mismatch due to differences in observed air mass should be studied and potentially be modelled to provide an estimate of this mismatch uncertainty.

3.4.3 Visualisation

As shown in Figure 6, the time-series of co-located AOD observations from the AERONET network is compared to the satellite product of Envisat AATSR. Measurements of AOD as well as averaged values are presented in Figure 6a. Also shown for the satellite retrievals are relative uncertainties. This information is not available for the AERONET network observations (see D2.7 for present summary of understanding of these measurements). The relative difference (standard deviation and uncertainties) is shown in Figure 6b.

Identified potential improvements to visualisation are*:*

- In addition to time series, it would be useful to have 2D density plots (AATSR vs AERONET AOD) and histograms of the absolute differences plotted;
- In the AOD plot, single uncertainty components can be plotted separately. The option to plot the total uncertainty should be added;
- In the difference plot, the overall uncertainty is plotted but individual uncertainty components would be useful;
- In the difference plot, the relative differences and relative uncertainties are plotted. It is advisable to plot the absolute AOD difference and the AOD uncertainties (i.e. not normalised). Indeed, low AOD are very common and normalisation by AOD is not suitable;
- Displaying the uncertainty vs AOD difference would be a useful option.

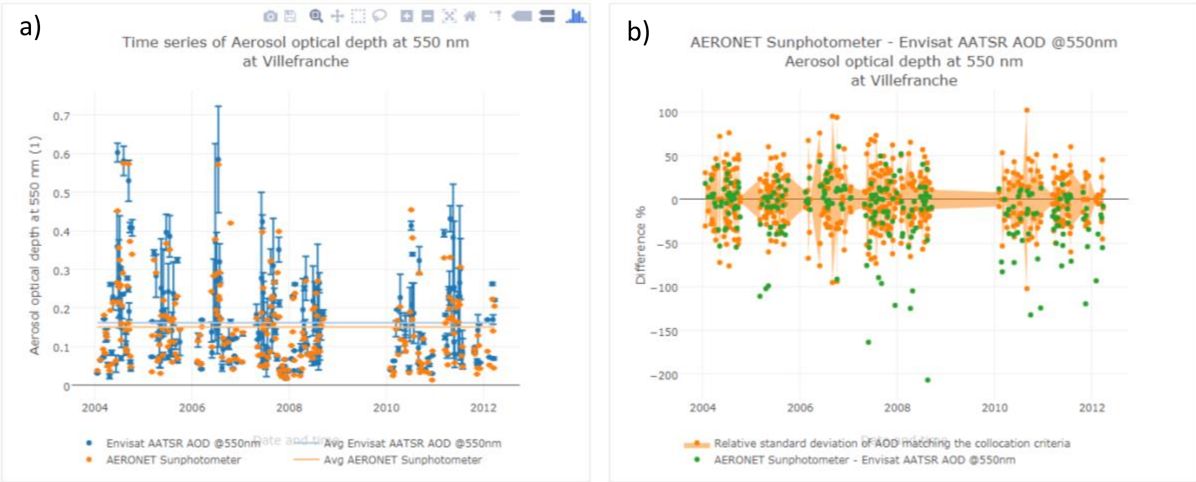

Figure 6:6: Screenshot of time-series produced into the VO at Villefranche for Aerosol Optical Depth (AOD) a) AOD measurements from both the Envisat satellite product and ground-based observations from the AERONET network and b) Relative standard deviation (in percent) with the uncertainties.

4 Quality Assurance assessment

An evaluation of the VO non-satellite products against the QA4ECV (Quality Assurance for Essential Climate Variables) principles has been undertaken to assess the information content of the VO, and its applicability to the user community. The QA4ECV principles were developed in the FP7 QA4ECV project[4] which followed on from the earlier QA4EO project[5]. The QA4ECV principles are being applied to a number of projects[6] across the EO data community, to ensure that data is of sufficient quality for the users' applications.

The QA4ECV Summary Product QA reports collate all of the information input by the data provider into a summary template that consists of 8 sections including:

1. Product Details (basic metadata about the product)
2. Product Availability (where the product and code can be obtained)
3. Product Documentation (core documents and relevant publications)
4. Product Generation (input data and algorithm traceability chain)
5. Product Quality (completeness, quality flags, uncertainties)
6. Product Evaluation (validation, inter-comparisons)
7. Product Applications (recommendations for usage and associated references)
8. Assessment Against Standards (GCOS and Maturity Matrix assessment)

These are further summarised into a six criteria Quality Indicator (QI) evaluation checklist (Figure 7) which indicates three levels of increasing compliance with recognised best practices for each QI. These levels are:

---

[4] http://www.qa4ecv.eu/qa-system
[5] http://qa4eo.org
[6] CHARMe, CLIPC, C3S, Core Climax, ESA CCI, EUPORIAS, FIDUCEO, GAIA-CLIM, GeoViQua, Obs4MIPS & UERRA – See QA4ECV D2.3

Basic – Some information is provided on the quality of the product to allow the users to make a simple distinction between the product and alternatives.

Intermediate – Detailed information is provided on the product, allowing the user to understand how it was made, and the quality and uncertainty information available to them.

Advanced – Significant detailed information is provided on the product, providing the user with enough information to make an informed decision about how the product should be used.

In addition to the above, there is a baseline amount of information required for each product providing rudimentary information about what the product is. Figure 7 provides a summary of the type of information that should be provided for each QI to achieve each of the three audit levels. This table provides product producers with the information they need to understand how their quality records are assessed: http://www.qa4ecv.eu/qa-system.

QA4ECV Quality System:
Quality Checklist

Aim: Checking the level of detail provided about an ECV data product helps to improve the available information. The checklist provided here states what information should be provided and details levels against which the information provided can be assessed.

	QA4ECV-Basic	QA4ECV-Intermediate	QA4ECV-Advanced
Details	Documentation available with source code and details on product completeness + consistency	Parameter details provided; ATBD / PUG provides significant detail; forum provided for users	Parameter details split spatially / temporally; ATBD / PUG in line with guidance; forum is monitored
Traceability	High level diagrams with basic information on algorithm provided	Detailed diagrams with relevant sub-chains; detailed information for steps	Detailed information on most steps as well as uncertainty information provided
Flags	Simple flags available in the product with basic information available on derivation and usage	Several flags provided allowing easy distinction of data quality; details provided for each flag	Comprehensive set of flags / ancillary data provided to allow detailed understanding of quality
Validation	Assessed against LPV hierarchy; validation report available and some campaign details	Justification for LPV hierarchy provided; good level of detail provided for validation and intercomparison.	Validation guidance closely followed and comprehensive information on campaigns provided
Uncertainty	Details of uncertainty calculations provided including how they are made available in product and how they should be used	Contributors to uncertainty analysis and calculation details provided with enough information to allow immediate use	Uncertainty significance estimates for all contributors
Assess	Maturity matrix filled in to some extent; GCOS – basic details provided.	All maturity matrix filled in; comments provided for GCOS.	All boxes filled with consensus between producer and auditor.

QA4ECV Help: QA4ECV_helpme@npl.co.uk   Version 1.0 – July 2017

Figure 7.7. Quality indicator evaluation checklist, from QA4ECV D2.9 Framework for a prototype QA Service in support of C3S

In the remainder of this section the GAIA CLIM VO co-locations highlighted in Section 2 and Table 1 are assessed against the QA4ECV principles. An assessment against the six quality indicator evaluation checklist criteria is made with justification for the assessment given in each case.

4.1 GRUAN processed Brightness Temperature (BT)

Assessment category	Level	Comment
Details	Intermediate	Peer-reviewed instrument/uncertainty paper available. Comprehensive PTU available for the profile, but not the GRUAN processor BT product.
Traceability	Intermediate	Detailed traceability diagram in PTU for profile product, but the GRUAN processor BT product traceability is not covered, particularly the spectroscopy aspects.
Flags	Intermediate	See netcdf
Validation	Intermediate	Validation activities underway by UKMO/ECMWF but not formally reported. See WP4 results.
Uncertainty	Intermediate	Intermediate level uncertainty information in netcdf, with detailed uncertainties for all contributions to profile in the PTU documentation. Additional steps the BT product not assessed.
Assess	Basic	WP1 maturity matrix completed for profile product only.

4.2 NDACC FTIR ozone

Assessment category	Level	Comment
Details	Intermediate	PTU available for the FTIR ozone product No single peer-reviewed instrument/uncertainty paper available.
Traceability	Intermediate	Traceability diagram in PTU for the FTIR ozone product. Traceability of spectroscopic parameters remains an issue.
Flags	Intermediate	See netcdf
Validation	Intermediate	Laboratory and field validation undertaken for some PTU elements.
Uncertainty	Intermediate	Intermediate level uncertainty information in netcdf, with uncertainties for major contributions in the PTU documentation.
Assess	Intermediate	WP1 maturity matrix completed for NDACC network as a whole.

4.3 NDAAC UV-Vis DOAS ozone

Assessment category	Level	Comment
Details	Intermediate	Peer-reviewed instrument/uncertainty paper available. PTU available for the DOAS UV-vis product.
Traceability	Advanced	Detailed traceability diagram in PTU for DOAS product.
Flags	Intermediate	See netcdf
Validation	Intermediate	Field inter-comparisons undertaken.
Uncertainty	Intermediate	Intermediate level uncertainty information in netcdf, with uncertainties for some contributions in the PTU documentation.
Assess	Intermediate	WP1 maturity matrix completed for NDACC network as a whole.

4.4 AERONET Aerosol Optical Depth

Assessment category	Level	Comment
Details	Intermediate	No PTU is available for the AERONET aerosol product. No current single peer-reviewed instrument/uncertainty paper is available although it is well described in the wider literature.
Traceability	Intermediate	No detailed traceability diagram is available. AERONET aerosol was not assessed in GAIA CLIM WP2
Flags	Intermediate	See netcdf
Validation	Intermediate	Field inter-comparisons undertaken.
Uncertainty	Intermediate	Intermediate level uncertainty information in netcdf.
Assess	Intermediate	WP1 maturity matrix completed for AERONET network.

5 Summary

This report presents an initial quality evaluation of the final status of the VO developed during the GAIA-CLIM project, with the understanding that the VO constitutes solely a set of demonstrator tools. This evaluation addressed the data accessible from the VO, the specific value of the Level 1 HIRS data for instrument data characterisation, an assessment of the utility to plot and analyse data in the VO, and a QA assessment using methodology from the EU QA4ECV project.

This assessment has shown that the VO provides a good demonstrator for the GAIA-CLIM outputs in terms of provision and display of uncertainty estimates for very different types of atmospheric variables and the comparison of those derived from satellite and non-satellite data. The addressed variables represent fairly complex comparison set ups. Temperatures and humidity are measured by radiosondes that are moving targets that need to be co-located with a satellite snapshot. The ozone and aerosol ECVs are retrieved using data in the visible part of the electromagnetic spectrum where different observation geometries from ground and satellite easily dominate the uncertainty budget of a comparison.

Overall, the VO concept has been well received as a proof of concept by the various attendees of the project roadshows as summarised in D6.9. A major achievement is the accessibility of co-located data containing all available uncertainties for download in an easy to use format. This functionality has a great potential as it may safe time of many people that do not need to download all data and compute the co-locations themselves.

The 3D-metadata viewer is supporting the VO in a way that a user can also look into the future using an orbit propagator and can plan the usage of specific stations with expected close co-locations.

The comparison of the GAIA-CLIM approach using ground-based to satellite data comparisons to other means such as satellite-satellite data match-ups and NWP output-satellite data comparisons has shown that all three are needed. The use of ground-based data ensures that the satellite data are compared to traceable standards, while the other comparisons deliver robust statistics and in case of NWP outputs also global coverage. These comparisons also demonstrate that the co-location mechanism itself can still further be improved aiming at lower systematic differences. This is in particular true for water vapour absorption channels where the height from which the signal originates that the satellite instrument measures, depends on the actual water vapour loading of the atmosphere, which has not been taken into account in the current version.

The biggest shortcoming of the current VO is that no single comparison is metrologically complete containing all of non-satellite measurement uncertainty, satellite measurement uncertainty and co-location uncertainty.

The QA assessment is clearly documenting that most assessment categories have been assigned an intermediate status leaving potential for improvement. As a consequence, further work on both the VO and the underlying preparation of the datasets and associated uncertainties is still required. In particular, the outputs of the H2020 FIDUCEO project would greatly help to provide the needed uncertainty estimates for the Level 1 satellite data that are currently missing in the VO.

The assessment of the graphical displays has resulted in various proposals for additional graphical representation of the co-located data, which are straightforward for implementation. The online analysis requires rather strong computing power at the server side to deal quick enough with simultaneous requests to the database being actioned. This was not available in the development environment for the VO with the consequence that online data interrogation is a distinct weakness. In addition, it requires a very good internet connection at both sides to become rather independent of the number of users using the VO at the same time. For this point, a big improvement potential exists if the VO would be made an operational facility.

Annex A References

Cao, C., M. Weinreb, and H. Xu, 2004: Predicting Simultaneous Nadir Overpasses among Polar-Orbiting Meteorological Satellites for the Intersatellite Calibration of Radiometers. J. Atmos. Oceanic Technol., 21, 537–542, https://doi.org/10.1175/1520-0426(2004)021<0537:PSNOAP>2.0.CO;2.

Carminati F., F., W. Bell, S. Migliorini, S. Newman and A. Smith, 2016: An introduction to the GRUAN processor. Satellite Applications Tech Memo 46n, January 15.

Dirksen, R., M. Sommer, F. Immler, D. Hurst, R. Kivi, and H. Vömel, 2014: Reference quality upper-air measurements: Gruan data processing for the vaisala RS92 radiosonde. Atmospheric Measurement Techniques 7, 12, 4463–4490.

Eck, T. F.; B. N. Holben, J.S. Reid, O. Dubovik, A. Smirnov, N.T. O'Neill, I. Slutsker, S. Kinne, 1999: Wavelength dependence of the optical depth of biomass burning, urban and desert dust aerosols. J. Geophys. Res., 104, 31333 – 31350.

EUM-PG, EUMETSAT ATOVS Level 1b Product Guide, EUM/OPS-EPS/MAN/04/0030 v3, January 2010, http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=pdf_v2a_atovs_level_1b&RevisionSelectionMethod=LatestReleased&Rendition=Web.

Hendrick, F., J.-P. Pommereau, F. Goutail, R.D. Evans, D. Ionov, A. Pazmino, E. Kyrö, G. Held, P. Eriksen, V. Dorokhov, M. Gil, M. and Van Roozendael, 2011: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations. Atmos. Chem. Phys., 11, 5975-5995, https://doi.org/10.5194/acp-11-5975-2011.

Holben B. N. et al., 1998: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. Remote Sensing of Environment, 66, 1 – 16.

Immler, F., and M. Sommer, 2011: Brief description of the RS92 gruan data product (rs92-gdp). GRUAN. Tech. Rep., Document GRUAN-TD-4, 17 pp. Available online at http://www.gruan.org.

Kottayil, A., S.A. Buehler, V.O. John, L.M. Miloshevich, M. Milz, and G. Holl, 2012: On the Importance of Vaisala RS92 Radiosonde Humidity Corrections for a Better Agreement between Measured and Modeled Satellite Radiances. J. Atmos. Oceanic Technol., 29, 248–259, https://doi.org/10.1175/JTECH-D-11-00080.1

Kreher, K., P.V. Johnston, S.W. Wood, and U. Platt, 1997: Groundbased measurements of tropospheric and stratospheric BrO at Arrival Heights (78 S). Antarctica, Geophys. Res. Lett., 24, 3021–3024.

Matricardi, M., 2009: Technical note: An assessment of the accuracy of the RTTOV fast radiative transfer model using IASI data. Atmos. Chem. Phys 9, 2.

Ohring G., B. Wielicki, R. Spencer, B. Emery, R. Datla, 2005: Satellite instrument calibration for measuring global climate change. Bull. Amer. Meteorol. Soc., Vol. 86, No. 9, pp. 1303-1313.

Platt, U. and J. Stuz, 2008: Differential Optical Absorption Spectroscopy (DOAS), Principles and Applications. ISBN 978-3-540-21193-8, Springer, Berlin-Heidelberg.

Pommereau, J.-P. and F. Goutail, 1988: O_3 and NO_2 ground-based measurements by visible spectrometry during arctic winter and spring 1988. Geophys. Res. Lett., 15, 891–894.

Richter, A., M. Eisinger, A. Ladstätter-Weißenmayer, F. Wittrock, and J.P. Burrows, 1999: DOAS zenith-sky observations: 2.Seasonal variations of BrO over Bremen (53 N) 1994–1995. J. Atmos. Chem., 32, 83–99.

Rodgers, C. D., 2000: Inverse methods for atmospheric sounding: Theory and Practice. Series on Atmospheric, Oceanic and Planetary Physics – Vol. 2, World Scientific Publishing Co., Singapore.

Saunders R. W., T. A. Blackmore, B. Candy, P. N. Francis and T. J. Hewison, 2013: Monitoring Satellite Radiance Biases Using NWP Models. IEEE Transactions on Geoscience and Remote Sensing, Vol. 51, No. 3, pp. 1124-1138, doi: 10.1109/TGRS.2012.2229283.

Schmid B. and C. Wehrli, 1995: Comparison of Sun photometer calibration by use of the Langley technique and the standard lamp. Applied Optics, 34, 4500 – 4512.

Shi, L., C. J. Schreck III, and V. O. John, 2013: HIRS channel 12 brightness temperature dataset and its correlations with major climate indices. Atmos. Chem. Phys., 13, 6907-6920, https://doi.org/10.5194/acp-13-6907-2013.

Smith, A., 2014a: Radiance simulator v1.0 user guide. Tech. Rep. NWPSAF-MO-DS-033, EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWPSAF).

Smith, A., 2014b: Radiance simulator v1 top-level design. Tech. Rep. NWPSAF-MO-DS-028, EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWPSAF).

Van Roozendael, M., P. Peters, H.K. Roscoe, H. De Backer, A.E. Jones, L. Bartlett, G. Vaughan, F. Goutail, J.-P. Pommereau, E. Kyrö, C. Wahlstrom, G. Braathen, and P.C. Simon, 1998: Validation of ground-based visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers. J. Atmos. Chem., 29, 55–83.

Viatte, C., M. Schneider, A. Redondas, F. Hase, M. Eremenko, P. Chelin, J.-M. Flaud, T. Blumenstock, and J. Orphal, 2011: Comparison of ground-based FTIR and Brewer O3 total column with data from two different IASI algorithms and from OMI and GOME-2 satellite instruments. Atmos. Meas. Tech., 4, 535-546, https://doi.org/10.5194/amt-4-535-2011.

Vigouroux, C., T. Blumenstock, M. Coffey, Q. Errera, O. García, N.B. Jones, J.W. Hannigan, F. Hase, B. Liley, E. Mahieu, J. Mellqvist, J. Notholt, M. Palm, G. Persson, M. Schneider, C. Servais, D. Smale, L. Thölix, and M. De Mazière, M., 2015: Trends of ozone total columns and vertical distribution from FTIR observations at eight NDACC stations around the globe. Atmos. Chem. Phys., 15, 2915-2933, https://doi.org/10.5194/acp-15-2915-2015.

Vigouroux, C., M. De Mazière, P. Demoulin, C. Servais, F. Hase, T. Blumenstock, I. Kramer, M. Schneider, J. Mellqvist, A. Strandberg, V. Velazco, J. Notholt, R. Sussmann, W. Stremme, A. Rockmann, T. Gardiner, M. Coleman, and P. Woods, 2008: Evaluation of tropospheric and stratospheric ozone trends over Western Europe from ground-based FTIR network observations. Atmos. Chem. Phys., 8, 6865–6886, http://www.atmos-chem-phys.net/8/6865/2008/.

GAIA-CLIM deliverables:

All GAIA-CLIM deliverables are available at: http://www.gaia-clim.eu/page/deliverables.

D1.9: Final version of a 3D tool for the online visualisation of existing measurements, August 2017.

D2.7: Report summarising the uncertainty estimates for the ECVs identified in Task 2.2, February 2018.

D2.8: Final report on the measurement uncertainty gap analysis from each subtask under Task 2.1 of WP2, February 2018.

D3.7: Final version of tools for quantification of co-location mismatch and smoothing uncertainties and associated documentation for integration in the virtual observatory that reflects any subsequent updates arising as a result of a. feedback from WP5 and b. any subsequent finessing in tasks T3.1 and 3.2, November 2017.

D4.4: Publicly available web based monitoring pages showing a comparison of GRUAN observations with Met Office and ECMWF data assimilation systems as an input to Virtual Observatory, February 2017.

D4.7: Report detailing approach to the calibration and validation of (atmospheric state variable) EO data, and detailing proposed approach to other ECVs and associated EO data.

D5.5: Virtual Observatory Product User Guide and Implementation Description, December 2017.

D5.9: Final version of the Virtual Observatory visualisation and data access facility, March 2018.

D6.9: Report on external stakeholder consultation exercise, January 2018.

Annex B Acronyms

AATSR	Advanced Along-Track Scanning Radiometer
ADV	AATSR Dual-View
AE	Angstrom Exponent
AERONET	AErosol Robotic Network
AMF	Air Mass Factor
AOD	Aerosol Optical Depth
BIRA-IASB	The Royal Belgium Institute for Space Aeronomy
BT	Brightness Temperature
DOAS	Differential Optical Absorption Spectrometer
ECV	Essential Variable climate
ECMWF	European Center for medium-range Weather Forecast
ESA	European Space Agency
FMI	Finnish Meteorological Institute
FTIR	Fourier Transform Infrared Spectroscopy
GAIA-CLIM	Gap Analysis for Integrated Atmospheric ECV - CLImate Monitoring
GSFC	Goddard Space Flight Center
GUI	Graphical User Interface
GSICS	Global Spaced-based Inter-Calibration System
GRUAN	Global Climate Observing System (GCOS) Reference Upper-Air Network
HIRS	High-resolution Infrared Radiation Sounder
NDACC	Network for the Detection of Atmospheric Composition Change
NASA	National Aeronautics and Space Administration
NWP	Numerical Weather Prediction
PHOTONS	PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire
PTU	Product Traceability and Uncertainty
RMSE	Root Mean Square Error
SAOZ	Système d'Analyse par Observation Zénithale
SCD	Slant Column Density
STD	Standard deviation
SZA	Solar Zenith Angle
TCA	Total Column Amount
TUT	Tallinn University of Technology
UKMO	United Kingdom Meteorological Office
UV	Ultra Violet
VCD	Vertical Column Density
VO	Virtual Observatory
WP	Work Packages