GAIA-CLIM Report

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring:

Transition Roadmap for the Virtual Observatory

A Horizon 2020 project; Grant agreement: 640276

Date: 19/03/2018

Beneficiary: EUMETSAT

Nature: R

Dissemination level: Pu

Work Package	WP 5
Deliverable	D5.8
Title	Transition roadmap for the Virtual Observatory
Nature	R
Dissemination	PU
Beneficiary	EUMETSAT
Date	19/03/18
Status	Final
Authors	Jörg Schulz
Reviewers	Peter Thorne, Anna Mikalsen
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. 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

1. Introduction

The Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring (GAIA-CLIM) project aimed 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 included: 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 (cal/val) tool for satellite measurements in future.

The successful transition of the VO (or components thereof) into further usage, sustained operation and further evolution very much depends on the needs of currently evolving operational services and research activities, primarily in the sphere of the Copernicus services and dedicated space missions, in which the VO could play a role. Some of the described potential evolution maybe realised within existing Copernicus Climate Change Service (C3S) activities, namely 311a Lot 3 working on the characterisation of ground-based data and 311b working on the preparation of satellite data records for reanalysis. In particular, for the latter an option will be formulated that will contain the use of ground-based reference data originating from 311a Lot 3 for the validation of several satellite data records from infrared and microwave sounding instruments. To realise this, the demonstrated functionality of the VO to address the monitoring of satellite radiances needs to be transitioned into operations. However, such use for quality evaluation of climate data records inherits complexity as described in Section 5. If successfully implemented for C3S, it seems also possible to integrate this functionality into operational real-time radiance monitoring systems, but only if the reference data are available in a timely fashion.

The demonstrated capability of the VO to analyse the quality of retrieval schemes for GCOS ECVs is also a promising path for further evolution, because the comparison of new or updated satellite data products to high-quality data is a widespread standard application in operational and research activities. The concept of the VO could clearly play a role in Copernicus Services in all data quality assessment activities, e.g., the 'Evaluation and Quality Assessment' pillar of the Copernicus Climate Change Service. However, the current VO is directed toward one-to-one instantaneous co-location comparisons whereas the Copernicus Services are mostly interested in the evaluation of gridded data products. For satellite products, those data products are most often derived from aggregated single retrievals for which the quality could be assessed with the VO. Section 5 also proposes evolutions to directly address the aggregated Level-3/ Level-4 products adding new challenges to the estimation of uncertainty for such comparisons.

The retrieval assessment function of the VO could also play a role in operational cal/val activities of space agencies and selection processes of satellite retrievals (round robin) for the generation of climate data records as for instance performed in the ESA Climate Change Initiative (CCI). Further usage of the VO can also be imagined in new research projects that involve data comparisons considering all involved uncertainties.

This report provides an outline for potential future evolutions of the VO following on from the end of the GAIA-CLIM project. Although the focus is on specifically further development of the VO, it is important to stress that such developments could be achieved either via further evolution of the VO itself and / or incorporation and further development of the underlying capabilities into other relevant projects or programs. Guidance on the applications for which the VO may be used in the future is provided and what major activities would need to be undertaken to reach a functional status. The VO consists of several distinct components such as the data base of co-located measurements including all available uncertainty estimates and the Graphical User Interface (GUI). As alluded to above, it is important to understand that existing needs of applications may also be satisfied by the evolution of individual components of the VO combined with other tools.

Potential evolutions for the VO are developed starting from the status of the VO at the end of the project, ideas provided in the GAIA-CLIM User Survey and the feedback from the GAIA-CLIM 'roadshow events'. All described evolutions would need to be integrated into larger frameworks and plans of satellite agencies or European scale services to acquire the needed long-term funding commitments needed for future developments and long-term maintenance.

2. Status of the Virtual Observatory at the end of GAIA-CLIM

The VO represents the culmination of the underlying scientific activities within the GAIA-CLIM project. These activities encompassed: the characterisation of geographical capabilities; the metrological characterisation and qualification of a set of target measurement techniques; quantification of the impacts of co-location effects on measurement comparisons; and the use of data assimilation and radiative transfer modelling. These scientific underpinning work package activities occurred to a large degree in parallel to the development and deployment of the VO facility. Despite the scientific work packages delivering the materials envisaged broadly to the timescale envisaged, inevitably not all work carried out within the work packages made it into the final version of the VO.

The assessment of the capability of the VO as described in deliverable D5.7 concludes that the VO constitutes 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. 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 (cf. the Report on the evaluation of the Virtual Observatory, D5.7). 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 VO in its current version is that no single comparison is metrologically "complete", containing all contributions of non-satellite measurement uncertainty, satellite measurement uncertainty and co-location uncertainty. Thus, a big potential for improvement exists for further development of the VO, including the implementation as an operational facility in the future. Importantly, although non-satellite instrument uncertainties, radiance-space comparisons, satellite uncertainties and co-location uncertainties are each available in one or more of the demonstrator case studies presented (see the VO User Guide, D5.9), unfortunately they are not all present in a single case study. Developing and showcasing such a case study would be of benefit in testing whether we have formal closure of our understanding of a comparison. Work with FIDUCEO[1] derived products and some of the data and tools developed and deployed on the VO may enable this in the future. Also, there exists the potential to consider multiple measurements arising from distinct measurement techniques of the same ECVs and this wasn't realised in any of the case studies. To that end, it is clear that while substantial pull-through occurred, by no means all that could potentially have been used was, in the end, used.

Specific work that did not make the cut is as follows:

- Work package 2 produced documented uncertainties for microwave radiometers (temperature and humidity observations) that were determined to be of reference quality, and also data for temperature, ozone and aerosols from lidars that represent high-quality estimates;
- Work package 3 produced a range of co-location effects look up tables that did not make the final version of the VO, particularly those for GRUAN temperature and humidity profile data over Europe;
- Work package 4 'GRUAN processor' work produced radiance equivalent measurements for a suite of channels of a range of microwave and infrared imagery which did not get included in the VO, which only considered a representative subset of five High-resolution Infrared Radiation Sounder (HIRS) channels;
- Work package 4 also included additional information via data assimilation that could have been used to aid the quantification of the co-location uncertainties.

These aspects are considered in Section 5 of this document for different potential evolutions of the VO.

3. Guidance from the GAIA-CLIM User Survey

The GAIA-CLIM User Survey (deliverable D6.1), performed at the beginning of the project, provided guidance to the development of the VO throughout the project. Major items for consideration were:

- Presenting long time series (because the majority of the responders were interested in climate research);
- Serving the application areas of model evaluation, climate analysis and quality assessment of ground-based data/satellite retrievals;

[1] http://www.fiduceo.eu/

- Developing very good and easily understandable explanations, especially regarding uncertainty measures and provision of a classical user guide document;
- Developing the VO as offline application (in delayed data mode) in its initial implementation but keep the desire of near-real-time availability in the view.

The first two bullets are partly met by the VO at the end of the project, but time series are still rather short, which is principally due to limited availability of non-satellite reference measurements. The functionality of climate model evaluation has not been implemented, but is still an interesting idea for further evolution and is discussed later in the document. The quality assessment of satellite retrievals has been achieved. The last bullet is pointing at an operationalisation of a VO service that is fed with data in real time which presents an option for the further evolution of the VO.

4. Guidance from the roadshow feedback

The user feedback upon the VO is summarised in the project deliverable D6.9. The need for collaboration with other tool-providers was stated very often inter-alia: as a means to avoid redundant tools; as a means to complete it with complementary building blocks; and as a means to make tools and systems harmonised and interoperable. Also often repeated was the statement that further developments and maintenance on a sustainable level were prerequisites for it being used. The roadshow feedback provided some interesting specific points and suggestions for the further evolution of the VO:

- The 3D-metadata tool could be extended to a full co-location prediction tool to create a set of best possible match-ups between reference and satellite data. It would become part of an operational system that is distributing this information to station operators to optimise scheduling of observations;
- The extension of the VO to more GCOS ECVs, covering the atmosphere, ocean and land systems, would be desirable;
- The extension to long-time series for all included ECVs to increase the usefulness of the VO for climate studies;
- Tools, such as the radiance simulator, need to be further developed to better characterise the uncertainty in the simulated radiance at the top of the atmosphere (TOA). This would include a better consideration of all scattering and absorption processes in the atmosphere, better surface emissivity modelling, etc. Most of this is covered by the major project recommendations on the conversion of non-satellite measures to TOA-radiance equivalents and their use in the Recommendations document to address gaps in observing capabilities (deliverable D6.10);
- Discovering and downloading co-located data was seen by many users as one of the most useful functionalities of the VO. Based on this finding, the VO could become an automatic co-location engine that provides its data via the internet for as many as possible combinations of non-satellite and satellite data at both Level-1 and Level-2. Such a tool would need to have an information service that informs its users about new data in the data base and could also push such data to the users’ systems. Users see a big advantage in replacing local storage and processing of large amounts of data by web-based processing;

- Integration of tools into larger tool sets, which could be done in the context of Copernicus Data and Information Access Services (DIAS), e.g., the specific DIAS developed by EUMETSAT, ECMWF and MERCATOR OCÉAN[2]. It was encouraged that this should work as a plug-and-play framework creating the ability for users to upload their own data sets and tools. Such an enhancement has the potential to boost the number of users of the VO.

5. Evolution of the Virtual Observatory

5.1. Operational monitoring of satellite Level-1 data quality

Classical monitoring of Level-1 instrument data quality is mostly based on the assessment of the evolution of time series of the data, calibration target measurements and other ancillary parameters of the instrument such as instrument housing and satellite temperatures. In addition, automatic satellite/satellite comparisons, as well as monitoring of instruments based on NWP forecast and analysis, is used to detect anomalies in the satellite measurements (see deliverable D4.7 for the latter). As the gap analysis undertaken within GAIA-CLIM has shown (see the Gap Analysis and Impact Document, D6.11), none of these comparisons constitute a traceability to a reference measurement. In space, traceability would be achievable for some regions of the electromagnetic spectrum with planned missions such as the Climate Absolute Radiance and Refractivity Observatory (CLARREO) (Wieliki et al., 2013). However, this mission is very expensive and has been postponed by NASA till 2023. Instead, the operational Global Space-based Inter-Calibration System (GSICS) is using non-traceable satellite instruments, such as the Infrared Atmospheric Sounding Interferometer (IASI), for detecting issues in other instruments measuring in the infrared (IR). The basis for this is that IASI measurements are almost an order of magnitude more stable compared to other measurements in the IR, e.g., from older radiometers in geostationary orbits. Although the newest microwave imager (Global Microwave Imager on the Global Precipitation Measurement mission satellite) is also very stable (see the individual reports on validation of satellites, D4.6), there is no reference measurement in the microwave spectrum such as IASI represents in the IR.

GAIA CLIM has developed and used the 'GRUAN processor' in a demonstration mode to improve the evaluation of NWP model temperature and humidity fields with reference measurements and to enable a better use of NWP data in satellite Level-1 data characterisation. This was done to address a gap (G4.01 in the GAID) on the lack of traceable uncertainty estimates for NWP and reanalysis fields and equivalent TOA radiances related to temperature and humidity. In addition, GAIA-CLIM has also applied the GRUAN processor directly to the GRUAN radiosonde data (topped and tailed with NWP data) and compared such estimates to co-located satellite measurements.

An analysis of these three complementary approaches (see details in D5.7) has shown that both, the satellite/satellite and NWP forward model approach provide robust statistics, while the direct co-location is hampered by very low numbers of co-locations. To improve upon this the relevant GAIA-CLIM recommendations (see deliverable D6.10) on time scheduling, coherency of satellite and non-satellite measurements and on the provision of reference-quality measurements on a continuous basis, need to be implemented.

[2] https://www.eumetsat.int/website/home/DIAS/DIAS/index.html

The most promising way of using the reference measurements, based upon the results presented in GAIA-CLIM, would be to exploit them to increase the credibility of NWP TOA radiance estimates. As the used radiative transfer model RTTOV has been demonstrated to be usable for many satellite instruments, this has the potential to become a powerful tool for operational monitoring of Level-1 satellite data. However, following the GAIA_CLIM recommendations on conversion of non-satellite measures to TOA radiance equivalents and their use (see deliverable D6.10), further efforts would be required to operationalise the availability of the GRUAN processor and generalise the processor to include other reference-quality measurements from further non-satellite measurement techniques. This would also take up 'leftovers' in terms of additional ground-based reference measurements from GAIA-CLIM work package 2 as described in Section 2.

To systematically integrate the non-satellite reference data into Level-1 monitoring activities, work needs to be undertaken to integrate a further updated GRUAN processor into an operational environment in which the NWP, GRUAN and other data would be available in real time. The application of the GRUAN processor and the co-location procedure, potentially based on predicted orbits from an enhanced 3D-metadata tool, would need to be run in an operational environment under complete configuration control. Such a development can be proposed for operational activities at space agencies including EUMETSAT, e.g., the Satellite Application Facility on NWP for enhancing and maintaining the GRUAN processor.

5.2. Offline satellite Level-2 retrieval validation

The VO has also been used to demonstrate comparison of satellite data at Level-2 for ozone total columns and aerosol optical depth showing parts of the uncertainty budget for the comparison. Section 2 has indicated a couple of additional characterised measurement systems arising from WP2 (D2.8) that could immediately be added to the VO to achieve multiple comparisons for a single ECV and to achieve comparisons for additional ECVs.

One essential function to be added to the VO in order to become a successful retrieval validation tool is a data upload function for the user. This could be realised by:

- use of the data download function for co-located satellite instrument data for a reference measurement;
- application of the retrieval scheme at the user's computer;
- upload of the retrieval results to the VO by the user;
- automatic integration of the new retrieval results into the VO data base.

The GUI would need to be enhanced to enable display of several satellite products versus the same reference at the same time. This would be a useful enhancement of the VO in the context of satellite retrieval round-robin activities for selecting retrieval schemes. It could also support individual researchers during the improvement of satellite retrievals.

This enhancement of the VO could be realised in a research and development project with involvement of an institution that sustains the VO.

5.3. Operational monitoring of satellite Level-2 data quality

The VO could become a tool for operational monitoring of Level-2 satellite products as for instance provided by EUMETSAT, e.g., through its Satellite Application Facilities. For such an operationalisation to be successful, basically the same issues as described in Section 5 need to be addressed. Also for Level-2 monitoring, the inclusion of actual NWP-analysis data from at least two models would be useful. In addition, the VO would need to be expanded to be able to deal with many more variables compared to the status at the end of the project. This would, naturally, require additional work on measurement characterisation and co-location effects quantification for the expanded suite of variables.

Such an enhancement should be realised in an operational environment with support from research units ensuring correct implementation of uncertainty computations and supporting design of graphical displays. A successful implementation, permitting the operational use of non-satellite reference measurements in satellite product validation, would likely have a beneficial effect on the sustainability of funding of the ground-based measurements. With such long-term reference measurements, the VO could also address climate questions in the future.

5.4. Offline satellite climate data record evaluation

The validation of long climate data records from satellite data (>30 years) is a scientific challenge because measurements of sufficient quality that can be used for comparison are scarce. Also the often used GSICS satellite-satellite comparison approach is not possible as the number of missions in the 1970s and 1980s was very limited and the data of generally lower quality (less stable). Comparisons to reanalyses are also complicated as with few additional observations assimilated, the assessment becomes potentially substantively circular. However, reanalyses data can be used to support the detection of issues in satellite Level-1 data due to a relative high temporal consistency.

Using ground-based data would require the use of non-reference quality measurements because reference measurements, such as those arising from the GRUAN radiosondes, have typically been established only 10-15 years ago. To make non-reference ground-based measurements useful for validation exercises, the quality of such data needs to be understood as best as possible. Work within the C3S 311a Lot 3 activity, led by GAIA-CLIM participants, is attempting to implement improved characterisation of such data back to 1979. The use of data from this activity together with the application of the GRUAN forward simulation tool will be proposed by the C3S 311b to enable an additional tool to validate Level-1 data records. It will lead to an enhanced co-location data base containing more and better characterised measurements that could also be made available via the internet. The usage of the data base in this context would require the implementation of more diagnostic tools such as trend models (Weatherhead et al., 1998, Mieruch et al., 2014) and breakpoint analysis (Gallagher et al., 2013). The tools could also be implemented into an updated GUI.

5.5. Inversion of approach to use satellite data and NWP to characterise non-reference-quality non-satellite data

The GAIA-CLIM project was specifically limited by the nature of the call to the use of the non-satellite segment to characterise the satellite data. The restriction is scientifically incorrect in that the value in any multi-platform comparison can and should work in all directions. Indeed, use of satellite data to characterise non-satellite data is appealing as the satellite instrument is the same across all comparisons (under the assumption that careful satellite design considerations have substantively limited and / or eliminated all orbital effects).

Assuming that the satellite data are characterised using metrological approaches as outlined in the H2020 FIDUCEO project and that all uncertainties involved in the comparison to characterise reference measurements can be quantified, it may be possible to use satellite data as a travelling reference for the broader global observing system components that are not reference quality. Taking the example of radiosondes, the travelling reference of an infrared sounding instrument such as the Advanced Infrared Sounder (AIRS) and IASI could be used to assess the quality of all Global Upper Air Network (GUAN) radiosondes that are not reference standard back to the year 2000. An assessment of GUAN radiosondes can at least reveal systematic effects of specific radiosonde types, e.g., related to radiative effects on the radiosonde measurement. Satellite-satellite comparisons, e.g., of IASI to its predecessor the HIRS can help to extend the understanding of radiosonde-to-satellite differences further into the past. Such an analysis can be supported by using reference standard satellite measurements such as radio occultation for differentiating radiosonde temperature errors resulting from instrument characteristics as demonstrated by He et al. (2009) and Ho et al. (2017) for upper tropospheric and stratospheric temperature measurements. New methods to derive radio occultation bending angle profiles should offer more potential to extend such analysis into the mid and lower troposphere.

The data base of the VO would need to be enhanced with access to additional ground-based data beyond the reference data and more satellite instrument data, in particular from radio occultation instruments. These needed to be co-located to all available ground-based data to enable the described analysis. This evolution may be proposed for implementation by the C3S 311b activity to support the use of ground-based data for the quality evaluation of Level-1 climate data records.

5.6. Extension to comparison at Level-3/ Level-4

Most ECV climate data records are served as Level-3 or Level-4 products. These are attractive as they are most often designed to not contain temporal and spatial coverage gaps. Level-3 products are created from geophysical variables retrieved from satellite data and kept in satellite projection by mapping of the data to regular grids. This process may comprise of resampling, compositing, averaging and interpolation of the original retrieved values. Level-4 products originate if the resulting gridded product includes several sources of data and is supported by a dynamical model, as for instance in data assimilation. A challenge for all products at Level-3/ Level-4 is how to propagate the uncertainty from the Level-2 products to the gridbox averages through the different mapping methods. The resulting uncertainty also heavily depends on the chosen spatial scale of a product and the temporal aggregation of the data to produce the gridded series.

Another challenge is if gridded values, e.g., monthly means shall be compared to point measurements as they exist for many reference measurements that are not representative at the spatiotemporal scale of such a gridded Level-3/ Level-4 product.

Despite all these issues, an extension to Level-3/ Level-4 data would be a natural extension of the VO application that may attract many more users. It would also increase its applicability in the context of Copernicus Service’s evaluation and quality control activities where the quality of many gridded data products need to be evaluated also coming from climate models. However, many tools are currently built in this area, e.g., Copernicus DIAS. The VO can be distinctive if a traceable chain of uncertainty estimates from Level-1 to Level-3/ Level-4 satellite data including uncertainty budgets for comparison to the ground-based data could be established.

Enabling the comparison to climate model data can have a positive effect of supporting research activities in this area, e.g., the Observations for Model Intercomparison Projects (Obs4MIPs)[3] activity that collates and makes available satellite observations via the Earth System Grid (ESG) to support the quality monitoring of coupled climate model runs in the framework of Phase 6 of the Coupled Model Intercomparison Project (CMIP) (see Eyring et al., 2016a for further details on CMIP6). The satellite data collected by obs4MIPs are currently Level-3 gridded products, which are compared to model outputs mapped to the same grids. An evolution of the VO could be to develop a service that enables comparison of the observations containing uncertainty estimates for Level-3 products with climate model output online using existing tools such the Earth System Model eValuation (ESMVal) Tool (Eyring et al., 2016b). Such a service was already proposed by potential users of the VO and is captured in the GAIA-CLIM User Survey results. However, it would need a dedicated project to be realised.

5.7. Extension to more ECVs

The GAIA-CLIM roadshow events have clearly shown that users are interested in the VO to cover more GCOS ECVs. This means, a general extension to oceanic and terrestrial variables, but also the inclusion of more atmospheric ECVs. Adding ECVs would clearly increase the number of user communities and may make development of the VO a larger community activity. However, it is evident that such an extension is very challenging as for many ECVs traceable reference measurements do not exist. Thus substantive work would be required on measurement characterisation, as well as issues such as co-location effects quantification to underpin such an extension.

A realistic approach to enable this would be to contribute to the Quality Assessment of ECV products of the C3S. This could start with building up co-location data bases containing all available uncertainties for ground-based measurements compared to satellite products in the Climate Data Store (CDS) and could be extended to such ECVs covered by reanalyses data that will become available from the CDS. For the latter, the uncertainty propagation to Level-3/Level-4 products needs to be addressed first.

The co-location data base could be made accessible through the CDS or could be available in a Copernicus DIAS hosted processing environment that provides cloud-based processing capabilities that can be run on virtual machines with pre-installed tools, e.g., Jupyter Notebooks. This could be used by many users to perform quality analysis using the co-located data, upload own data and therefore has a huge potential for the development of additional tools for data analysis and graphical displays, which would foster the development of applications within a DIAS. The shaping of the DIAS environments for Copernicus is ongoing and it is currently not predictable if an inclusion of the VO concept into it is possible or not.

The described activity would implement activities related to several GAIA-CLIM recommendations on the provision of operational user tools that enable exploitation of reference measurements, but would

[3] https://www.earthsystemcog.org/projects/obs4mips/

also address recommendations on non-satellite data availability and quality because the need in the described service requires uninterrupted availability and high quality.

5.8. Training and capacity building tool

The VO contains a more complete representation of uncertainty estimates for satellite to ground-based comparison compared than other existing tools. Although the VO's GUI is only in a demonstrational mode at the end of the project, it has demonstrated the capability of displaying these uncertainties in a meaningful way. Making the existing comparisons metrologically complete could make the VO a very attractive tool for capacity building and would partly address the GAIA-CLIM final recommendation on education and training. The VO can be used to increase awareness on what is important in a comparison of satellite data with ground-based measurements and can be used to explain how to build up an uncertainty budget.

The current graphical display is a good starting point to assess the usefulness of the VO GUI for training purposes on uncertainties. In particular, the comparison of the ozone estimates from the EUMETSAT AC SAF utilising the Metop GOME-2 instrument to the ground-based FTIR and DOAS measurements at some stations has the potential to be used in training activities. This is because within the GUI many different uncertainty measures are explained and shown on actual data comparisons. The assessment of the VO capabilities (D5.7) contains additional proposals as to how to enhance the graphical capability in the VO for the existing variables. An updated display could be further enhanced by selecting and adding examples where different types of uncertainties dominate the comparison. Experienced capacity building personnel may develop specific "user journeys" through the VO to explain different aspects of uncertainty. The EUMETSAT training team has been informed about the VO and may use it in the future for training on climate data records.

6. Conclusions

Based on findings from the GAIA-CLIM user survey, user workshops and roadshow events, as well as from the initial assessment of the VO functionality in D5.7, this report outlines several options for the future evolution of the VO after the GAIA-CLIM project.

The report identifies eight potential areas for further evolution, which are:

- Operational monitoring of satellite Level-1 data quality;
- Offline satellite Level-2 retrieval validation;
- Operational monitoring of satellite Level-2 data quality;
- Offline satellite climate data record evaluation;
- Inversion of approach to use satellite data and NWP to characterise non-reference quality non-satellite data;
- Extension to comparison at Level-3/ Level-4;
- Extension to more ECVs;
- Training and capacity building tool.

Each of the potential evolution areas are described including the challenges in adoption and what component of the VO would need to be enhanced. The most important component of the VO to be further developed is the data base of co-locations containing the uncertainty estimates. The data base could easily be enhanced with more measurements and could serve co-located data for many applications. It could also be easily connected to other graphical display systems independently of the VO.

Most of the potential evolutions also require further development of uncertainty estimates for more ground-based networks and satellite data, additional quantification of co-location effects, as well as an extension of uncertainty propagation methods to Level-3/ Level-4 products. In addition, the radiative-transfer capability to enable Level-1 comparisons used in GAIA-CLIM would need to be used for more instruments and enhanced to enable a better uncertainty propagation into the simulated Level-1 data. The extension to more ECVs would require substantive activities in characterising ground-based measurements and may also require the extension of metrological approaches to active satellite measurements, which could be addressed in follow up EU research projects before being implemented to any service.

Although no cost estimates are provided for the evolution areas, it is evident that all potential evolutions described would need to be integrated into larger frameworks and plans of European space and research agencies, as well as Copernicus and national services to acquire the needed level of secure funding for future developments and long-term maintenance. Offline satellite climate data record evaluation and, as part of it, the inversion of approach to use satellite data and NWP to characterise non-reference quality non-satellite data, may be realised in C3S as described in Section 5. For other evolutions, dedicated projects need to be developed in which more work will need to be invested into establishing a user community that can drive further evolution or can directly contribute by providing elements, e.g., in hosted processing environments such as being developed for Copernicus DIAS.

Annex A References

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D., 2016a: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, *Geosci. Model Dev.*, **9**, 1747-1802, doi:10.5194/gmd-9-1747-2016.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E., 2016b: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev*., **9**, 1937-1958, doi:10.5194/gmd-9-1937-2016.

Gallagher, C., R. Lund, and M. Robbins, 2013: Changepoint Detection in Climate Time Series with Long-Term Trends. *J. Climate*, **26**, 4994–5006, https://doi.org/10.1175/JCLI-D-12-00704.1 .

He, W., S.-P. Ho, H. Chen, X. Zhou, D. Hunt, and Y.-H. Kuo, 2009: Assessment of radiosonde temperature measurements in the upper troposphere and lower stratosphere using COSMIC radio occultation data. *Geophys. Res. Let.*, **36**, L17807, doi:10.1029/2009GL038712.

Ho, S.-P., Peng, L., and Vömel, H., 2017: Characterization of the long-term radiosonde temperature biases in the upper troposphere and lower stratosphere using COSMIC and Metop-A/GRAS data from 2006 to 2014, *Atmos. Chem. Phys.*, **17**, 4493-4511, https://doi.org/10.5194/acp-17-4493-2017.

Mieruch, S., M. Schröder, S. Noël, and J. Schulz, 2014: Comparison of decadal global water vapor changes derived from independent satellite time series, *J. Geophys. Res. Atmos.*, **119**, 12,489–12,499, doi:10.1002/2014JD021588.

Weatherhead, E. C., G. C. Reinsel, G. C. Tiao, X.-Li Meng, D. Choi, W.-K. Cheang, T. Keller, J. DeLuisi, D. J. Wuebbles, J. B. Kerr, A. J. Miller, S. J. Oltmans, and J. E. Frederick, 1998: Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J. Geophys. Res.*, **103**, D14, 17,149-17,161.

Wielicki, B.A., D.F. Young, M.G. Mlynczak, K.J. Thome, S. Leroy, J. Corliss, J.G. Anderson, C.O. Ao, R. Bantges, F. Best, K. Bowman, H. Brindley, J.J. Butler, W. Collins, J.A. Dykema, D.R. Doelling, D.R. Feldman, N. Fox, X. Huang, R. Holz, Y. Huang, Z. Jin, D. Jennings, D.G. Johnson, K. Jucks, S. Kato, D.B. Kirk-Davidoff, R. Knuteson, G. Kopp, D.P. Kratz, X. Liu, C. Lukashin, A.J. Mannucci, N. Phojanamongkolkij, P. Pilewskie, V. Ramaswamy, H. Revercomb, J. Rice, Y. Roberts, C.M. Roithmayr, F. Rose, S. Sandford, E.L. Shirley, W.L. Smith, B. Soden, P.W. Speth, W. Sun, P.C. Taylor, D. Tobin, and X. Xiong, 2013: Achieving Climate Change Absolute Accuracy in Orbit. *Bull. Amer. Meteor. Soc.*, **94**, 1519–1539, https://doi.org/10.1175/BAMS-D-12-00149.1

GAIA-CLIM deliverables:

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

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.

D4.6: Individual reports on validation of satellites v3, December 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, February 2018.

D5.7: Report on the evaluation of the Virtual Observatory

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

D6.1: Report on results of user survey, July 2015.

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

D6.10: Recommendations document to address gaps in observing capabilities, February 2018.

D6.11: Gap Assessment and Impacts Document (GAID), final version, March 2018.