GAIA-CLIM deliverable D2.4

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring

WP2: Measurement uncertainty quantification

D2.4: "Progress report on the uncertainty estimates for the ECVs identified in Task 2.2".

A Horizon 2020 project;

Grant agreement: 640276

Date: 6th March 2017

Lead Beneficiary: BKS

Nature: Report

Dissemination level: PU

Work-package	**WP2**
Deliverable	D2.4
Nature	R
Dissemination	PU
Lead Beneficiary	BK Scientific GmbH
Date	06/03/2017
Status	Final
Authors	Karin Kreher (BKS), William Bell (MO), Domenico Cimini (CNR), Bruce Ingleby (MO), Fabio Madonna (CNR), Tim Oakley (MO), Peter Thorne (NUIM)
Editors	Peter Thorne (NUIM), Corinne Voces (NUIM)
Reviewers	Richard Davy (NERSC)
Contacts	karin.kreher@bkscientific.eu
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. Document rationale and broader context

Background. As part of WP1, the characterisation and assessment of a range of networks was undertaken using GAIA-CLIM's Measurement System Maturity Matrix (MSMM). A table providing a complete list of the networks which have been assessed so far within GAIA-CLIM is publicly available on the GAIA-CLIM web page (http://www.gaia-clim.eu/page/ maturity-matrix-assessment). In addition to the specific networks assessed, this list also includes the actual ECV(s) and measurement type(s) investigated for each network and a link to the individual assessment table. This assessment has then been used as guidance to classify these networks according to their characteristics and measurement capabilities as being of reference, baseline or comprehensive quality (for a definition, see Section 2). The results of this study are currently being summarized and will be made available in a peer-reviewed publication. Baseline networks are characterised by an average score of 3-4 in the maturity matrix and, within WP1, a maturity matrix table has been created and used to identify the capability of each of the networks of interest.

Following on from this geographical review by WP1, the aim of WP2 within this deliverable is to quantify uncertainties in baseline network measurement capabilities and outline in this progress report an introduction to a selection of baseline networks which are currently close to reference status, but still need more development in some of the assessed properties, and which may be of interest to the Virtual Observatory (VO) down the line. The aim is to quantify the uncertainties for these examples of baseline capabilities and, in the final report under deliverable D2.7, to evaluate these uncertainty estimates in the context of a corresponding reference network. One example would be to investigate how the baseline network GUAN relates to the reference network GRUAN for radiosonde temperature and water vapour profiles. The preliminary list of these examples will be extended in the final report, and best estimates for uncertainties of selected comprehensive networks will then be included as well.

2. Introduction of selected baseline networks and ECVs

As part of the systems-of-systems approach adopted and further developed within WP1 of GAIA-CLIM, we have defined the characteristics of its layers, i.e. the reference, baseline and comprehensive networks as follows:

Reference networks are defined as networks providing metrologically traceable observations, with robustly quantified uncertainty, at a limited number of locations, and/or for a limited number of observing platforms, for which traceability has been attained. This means that uncertainties arising from each step in the processing chain are fully quantified and included in the resulting data, full metadata is captured and retained, the observations program is actively managed and has a commitment to long-term operation, and any necessary changes to the measurement programme or instrumentation are carefully managed. For further details on the definition of reference observation networks see e.g. GAIA-CLIM deliverable D1.6 (Report on data capabilities by ECV and by system of systems layer for ECVs measurable from space).

Baseline networks are observing networks providing long-term records that are capable of characterising regional, hemispheric and global-scale features. Compared to reference networks, they lack the absolute traceability of their uncertainties but representative uncertainties, which are based upon understanding of instrument performance or the peer-reviewed literature, are available. Baseline measurements are also periodically assessed, either against other instruments measuring the same ECV at the same site, through comparisons to NWP/reanalyses, or through intercomparison campaigns. Ideally, such intercomparisons should include reference-quality measurements. Changes to the measurement program are minimized and managed, the observations have a long-term commitment and the metadata is retained.

Comprehensive networks are characterised as observing networks providing high spatio-temporal density data information necessary for characterising local and regional features. They provide representative uncertainties based upon, e.g., instrument manufacturer specification and knowledge of operation. If not available, uncertainty estimates based upon expert or operator judgement should be used. Long-term operation is encouraged but not required and metadata should be retained if possible.

Based on the range of networks investigated in WP1 and in WP2, Task 2.2, we have prioritised a limited initial list of baseline networks/ECVs using the MSMM assessment, specifically keeping in mind the interests of the GAIA-CLIM project and particularly their possible future use within the Virtual Observatory which will enable user access to satellite to non-satellite data comparisons. Table 1 lists the 5 networks selected thus far, and the ECVs of interest as a starting point for potential future inclusion of non-reference measurements of nevertheless well-characterised uncertainty.

In the following five sections, we will introduce the relevant selected baseline networks, measurement technique(s) and ECV(s), and discuss their MSMM and uncertainty estimates.

Section	Baseline network	Instrument/technique	ECV(s)
3	GUAN	Radiosonde	Temperature profiles Water vapour profiles
4	MWRnet	Microwave radiometer (MWR)	Temperature profiles Water vapour profiles Total column-integrated water vapour content (TWVC) Total column-integrated liquid water content (TLWC)
5	SHADOZ	Ozonesonde	Ozone profiles
6	GSN	Surface meteorology	Surface temperature
7	AERONET	Radiometer	Aerosol profiles

Table 1. Summary of selected baseline networks, specific measurement techniques and ECVs, and the section in which these will be discussed within this document.

3. GUAN radiosondes

3.1 Description of the specific network/technique/ECV

The scope of the GUAN (GCOS Upper-Air Network, see GCOS-144) is to comprise the best possible set of radiosonde stations with a spacing of between 5 to 10 degrees latitude and longitude, which is sufficient to resolve synoptic-scale waves and other hemispheric and global scale tropospheric changes. The parameters of interest are temperature, pressure (geopotential height), wind, and humidity (at least in the troposphere).

The purposes of the GUAN can be summarized as follows:

- To establish national commitments for the preservation of a minimum set of upper-air sounding stations for the foreseeable future.
- To build a collection of validated data from these stations in standardized formats.
- To provide this information to the global climate community with no formal restrictions on usage.

The most important criteria for inclusion into the GUAN are a commitment by the NMHS (National Meteorological and Hydrological Services) with regard to continuity, sufficient length and quality of historical time series, and appropriate measurement quality. The requirements should be interpreted such that every month at least one observation on each of at least 25 days should attain the Minimum Requirements (MRQs) listed below. The observing frequency (1 or 2 per day) in itself is not a criterion, although the Target Requirement (TRQ) for observation frequency is 2 per day, in accordance with World Weather Watch (WWW) regulations for radiosonde observations. Where possible, priority should be accorded to night-time ascents because these are less susceptible to radiative biases.

Minimum Requirements (MRQs) are:

- Temperature up to 30hPa.
- Humidity up to the tropopause.
- Wind direction and speed up to 30 hPa.

Target Requirement (TRQs, in addition to the MRQs) are:

- Temperature and wind up as high as possible.

Among the 171 GUAN stations (Figure 1) there are at least 11 different radiosonde designs in use and 5% (8 stations) have not reported any messages during 2016. The radiosonde type/system in current use at the GUAN stations, according to the metadata reported are as follows:

Vaisala RS92	-	71 stations (41%)
Vaisala RS41	-	18 stations (10%)

Modem M10	-	17 stations (10%)
Russian Fed. Manufacture	-	15 stations (9%)
Lockhead Martin (LMS)	-	11 stations (6%)
GRAW (DFM)	-	9 stations (5%)
Meisei	-	8 stations (5%)
China Manufacture	-	7 stations (4%)
InterMet	-	6 stations (3%)
Others (MLabor, JingYang)	-	2 stations (1%)
Silent (2016)	-	8 stations (5%)

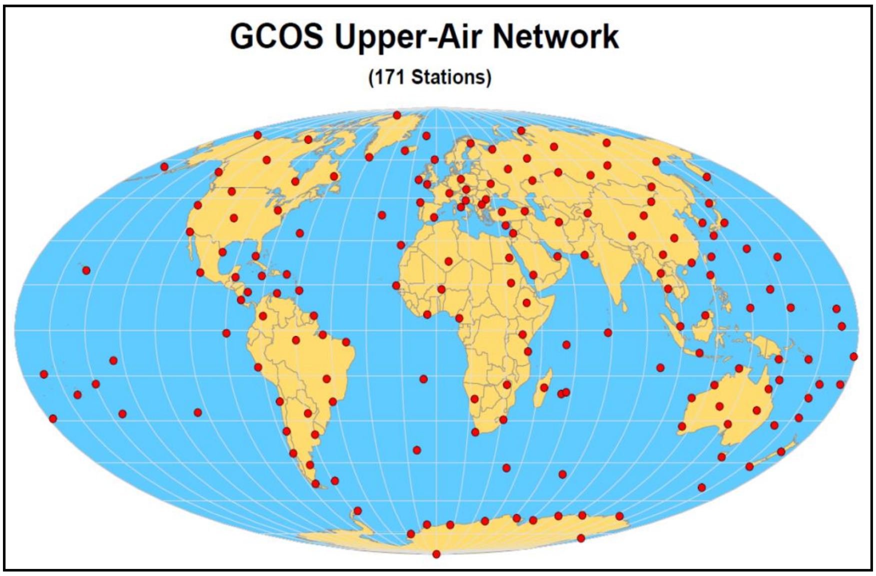

Figure 1. Map showing the 2016 status of the network, with 171 stations nominated to the GUAN.

For more details on the global distribution see the ECMWF technical memorandum (Ingleby, 2017). Within GUAN, the most frequently used type is currently the Vaisala RS92 which will cease manufacture within the next couple of years and be replaced by the RS41 type, which is currently the second most frequently used radiosonde (see list above). Anecdotally, most sites plan to move to RS-41 sondes, although competition in the marketplace is encouraged. Most radiosonde types occur predominantly in mid-latitudes, while Russian types are mainly employed at high latitudes and many of the Meisei (Japanese) radiosondes are launched at low latitudes. GUAN radiosonde types are representative of those in the broader comprehensive network of radiosonde networks. It may make more sense from the perspective of the VO to eventually consider the Global Observing System (GOS) sonde network not by GUAN / non-GUAN designation, but rather by instrument type. This is because the dominant sources of bias and uncertainty are driven by the instrument type in use at any given station.

3.2 GUAN maturity matrix

The maturity assessment for GUAN was performed within GAIA-CLIM in September 2016, http://www.gaia-clim.eu/system/files/document/GUAN%20with%20text%20.jpg (reproduced herein as Table 2). It assesses certain quantifiable aspects of typical measurement system maturity across the network for those ECVs and associated measurement systems that are relevant to GAIA-CLIM.

Typically, a reference quality measurement program would score 5s and 6s against relevant criteria, a baseline capability 3s and 4s and a comprehensive capability 1s and 2s. While some entries score higher (e.g. the metadata subcategories are all of reference quality), most subcategories within e.g. the uncertainty characterization and documentation of the GUAN maturity matrix score 3s and 4s which labels the network with baseline status. Within this report, we are particularly interested in the assessment of the uncertainty characterisation and the subcategories on measurement traceability, measurement comparability and uncertainty quantification which evaluates the extent to which uncertainties have been fully quantified and their accessibility to users. For all three of these subcategories GUAN scores 3.

GUAN

Metadata	Documentation	Uncertainty characterization	Public access, feedback and update	Usage	Sustainability	Software (optional)
Standards	Formal Description of Measurement Methodology	Traceability	Access	Research	Siting environment	Coding standards
Collection level	Formal Validation Report	Comparability	User feedback mechanism	Public and commercial exploitation	Scientific and expert support	Software documentation
File level	Formal Measurement Series User Guidance	Uncertainty Quantification	Updates to record		Programmatic support	Portability and numerical reproducibility
		Routine Quality Management	Version control			Security
			Long term data preservation			
Legend						
1	2	3	4	5	6	Not applicable

Table 2. Maturity matrix assessing the GUAN (GCOS Upper-Air Network) across seven major strands[1].

[1] *Users of GAIA-CLIM's Maturity Matrices, should be aware that this is a first effort to systematically quantify measurement system performance. Redundant assessments suggest a minimum uncertainty arising from assessor-to-assessor variations in any category of at least 1*

3.3 Description of (indicative) measurement uncertainties

Temperature. With respect to air temperature measurements made by radiosondes, the most prominent error of upper-air temperature sensors is the radiation-induced temperature error. In other words, radiosonde temperature biases, where the sensor temperature differs from the air temperature, are mainly caused by radiative effects (typically a warm daytime bias from sunlight heating the sensor and a cold bias at night as the sensor emits longwave radiation) with smaller errors due to lags in sensor response to changing temperatures as the radiosonde rises. All factors affecting longwave and shortwave fluxes around the sensor influence the bias, including sensor physical characteristics and mounting, and environmental factors including surface temperature, solar elevation angle, temperature lapse rate, ventilation velocity, and clouds.

Many methods have been used to improve air temperature measurements. However, the most generally accepted method - reducing the size of the sensor, and hence reducing the surface emissivity/absorptivity, and increasing the speed of the gas over the sensor - are to this day always a compromise and only partially eliminate the error. Hence, it is vital to adequately quantify the remaining uncertainty.

The measurement quality of present-day operational radiosondes is best described in the report on the 8th World Meteorological Organization (WMO) Intercomparison of High Quality Radiosonde Systems, held in Yangjiang, China, in July 2010 (Nash et al., 2011). Yangjiang is located in the sub-tropics adjacent to the South China Sea. The intercomparison was carried out during the SE Asian Monsoon season in order to test the performance under extreme conditions. During this intercomparison, 11 operational radiosonde systems from 11 manufacturers were tested and compared in relation to temperature, humidity, pressure, and wind components. The 11 instruments were flown in two groups. Group working references were determined, which had to be part of most flights in a group and of good measurement quality. The LMS System was used as reference in one group, whereas the Meisei system was used as working reference in the other group. The linking between the two groups was made with the help of Vaisala, Meteolabor, Daqiao, and Multithermistor measurements. Temperature comparison statistics were computed for simultaneous samples, banded into layers 1 km thick from the surface to 33km. Statistical analysis was used to estimate systematic and random errors, day-night differences, time response, and other characteristics of each sensor.

---

score. Although the assessment may be useful for certain applications, at this time and until more broadly tested, it should not constitute a primary or unique decision making tool.

For the systematic night biases (Figure 2), at least 15 successful comparison flights of each type to 26 km, 8 flights to 29 km, and 5 flights to 33km were used. The systematic day biases (Figure 3) were calculated using 12 successful comparison flights of each type to 26 km, 12 flights to 32 km, and 10 flights to 33 km. The reference for the systematic bias plots (Figures 2 and 3) were four independent sensor types with good time constants of response at all heights (Nash et al., 2011).

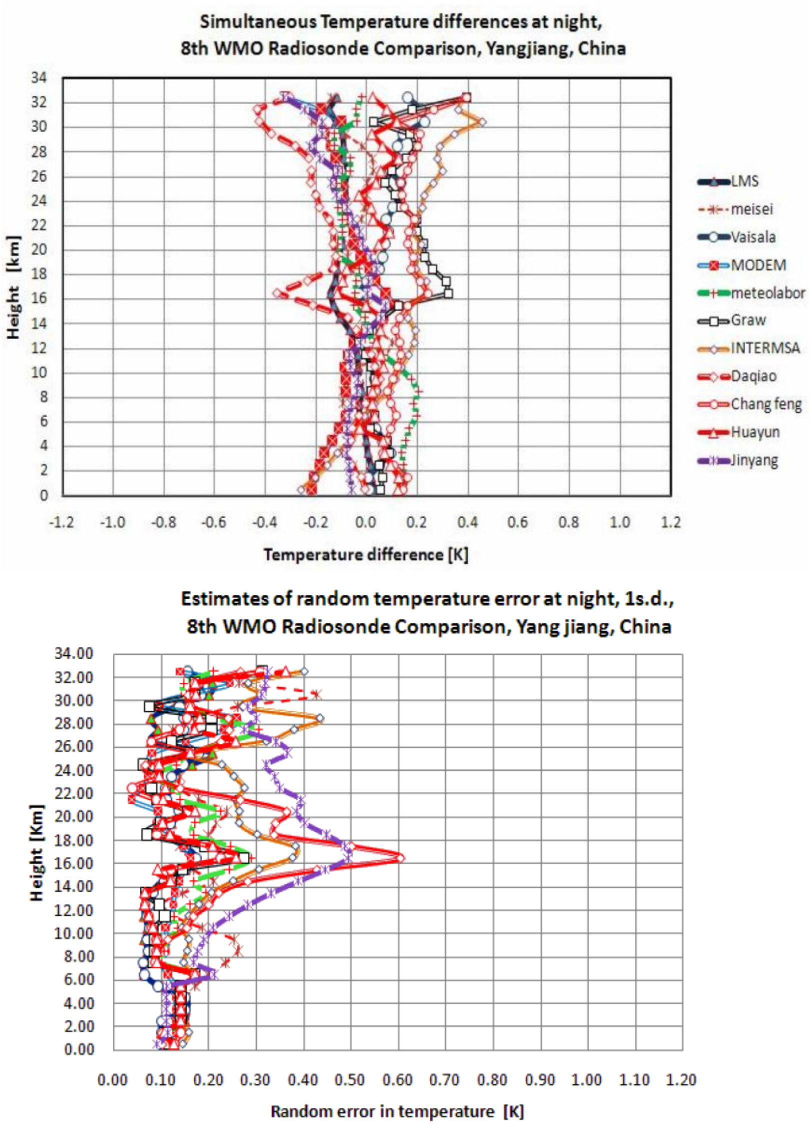

Figure 2. The systematic bias between simultaneous temperatures (K) at night is shown in the top panel, with a positive bias meaning the radiosonde reported higher values than the reference. The lower panel shows the estimated random error in temperature measurements at night (figures from Nash et al., 2011).

Figure 2 shows the systematic bias and random errors between simultaneous temperature measurements of the 11 systems at night. Below the tropopause (around 16 km) systematic differences of all systems are within more or less +/-0.2°C. The tendency for many of the radiosonde types to have larger random errors near the tropopause is probably linked to the effects of icing on the sensors when they leave dense upper-level cloud. At higher altitudes in the stratosphere, the differences are around +/-0.4°C. Overall, these results are rather encouraging, particularly if taking into account that the increase from +/-0.2 to +/-0.4ºC in the upper part is mainly due to three systems with larger biases. This analysis does, however, point to the need to apply systematic and random terms based upon the sensor type being considered as there is a clear range of both instrument systematic and random biases in evidence (Nash et al., 2011).

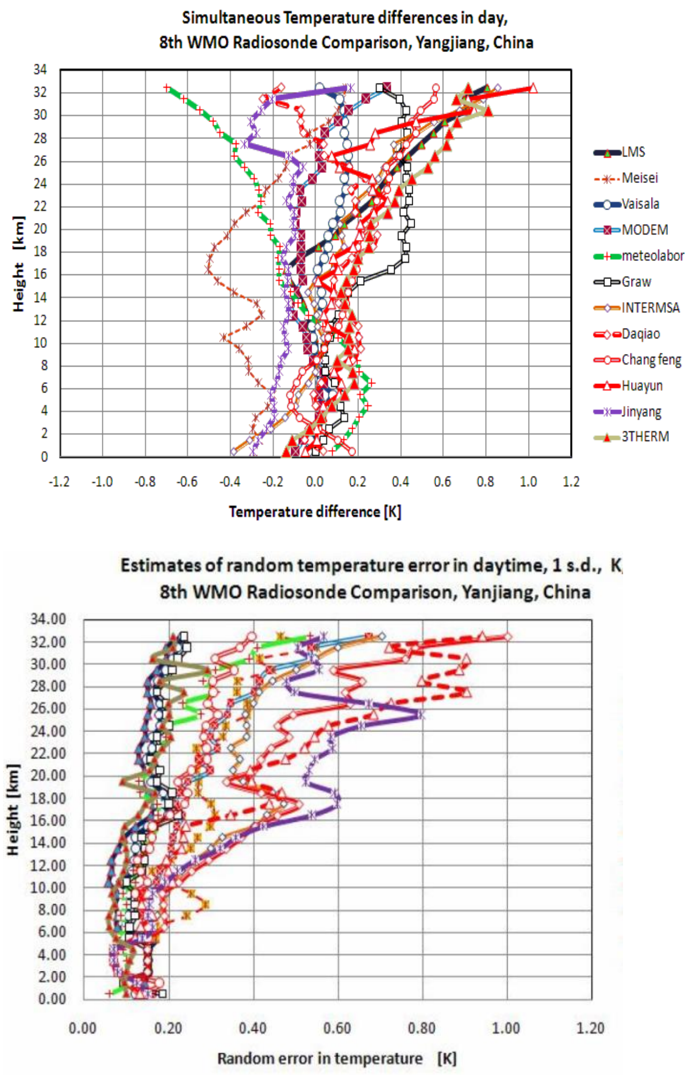

Figure 3. The systematic bias between simultaneous temperatures (K) during the day is shown in the top panel, with reference adjusted above 16 km to take into account estimate of day-night differences in geopotential height analysis. The lower panel shows the estimated random error in temperature measurements during the day (figures from Nash et al., 2011).

Figure 3 shows daytime systematic differences of simultaneous temperature measurements are on the order of +/-0.4ºC in the troposphere and around +/-0.8ºC in the stratosphere. The large divergence, particularly in the stratosphere, is primarily due to the solar radiation error, which strongly increases because of the reduced atmospheric pressure and related ventilation. Again, there is a broad range of both systematic and random behaviour which is dominated by instrument type.

In summary, with respect to temperature profile measurements at night, most radiosonde systems were found to make measurements to a high standard and provide suitable quality for both weather and climate applications. During daytime however, there is some degradation in quality and a rather large spread was found between the results of the individual sondes, particularly at high altitudes, with biases up to +/-0.8^{o}C at 32 km altitude (~10 hPa). This means that improvements and further tests are needed to reduce the systematic bias between the various radiosondes at these upper levels.

In addition to official CIMO (Commission for Instruments and Methods of Observation) intercomparisons, of which Yanjiang was the 8th such campaign, several regional activities are taken by certain national meteorological services. These are generally smaller-scale (comparing fewer sonde models) but may sample aspects such as seasonality etc. These allow an assessment of sensitivity of comparisons to synoptic conditions, solar elevation etc. They are generally less accessible to the research community. We shall aim to access and make use of such data in the final deliverable from Task 2.2 to supplement the Yanjiang based information.

From the most recent Yanjiang CIMO intercomparison results and our limited knowledge of those at mid-latitudes, it is unwise to assume that any of the current radiosonde temperature measurements can be reproduced consistently to within 0.1 K, as would be desirable for climate science applications. The results from Yangjiang show that there must be other limitations to the temperature quality than just solar heating and traceability to national metrological standards e.g. stability of the sounding system during flight. This means the origins of the uncertainty are probably not in the calibration of the sensors, but in the stability of the radiosonde sensor under different conditions during flight or in the stability of radiosonde signal electronics and processing also during flight. At night, these are the limiting factors. During daytime some improvement in the accuracy of the solar correction schemes, and references such as the Multithermistor radiosonde should be possible and are required if daytime measurements in the stratosphere are to be of most use to climatology and satellite data characterisation. Since it appears that systematic bias in the stratosphere is currently a limiting error, and this will not be stable from location to location (Nash et al., 2011).

In the final deliverable arising from Task 2.2 consideration will be given to providing look-up tables of random and systematic effects by sonde type contributing to GUAN. Inclusion of such information shall depend upon having sufficient intercomparison data, perhaps augmented by O-B statistics arising from WP4 type activities.

Relative humidity. The relative humidity sensors tested in Yangjiang had good reproducibility, but several types had large systematic errors at all heights in the troposphere, and the conclusion was that the origin of these needs to be identified and rectified as soon as possible.

The humidity sensors performance was evaluated in each of five relative humidity bands (0-20%, 20-40%, 40-60%, 60-80%, 80-100%). Apart from two sonde models, at temperatures higher than -40ºC (in the lower- to mid-troposphere), all the relative humidity sensors had good reproducibility. Many radiosonde types made measurements of poorer quality at night than in the day. Several systems showed potential for observing relative humidity in the upper troposphere in the tropics, and the new correction schemes seem to have good potential for future observations (Nash et al., 2011). It was possible to check the measurements in cloud, using cloud radar (up to 15 km) and ceilometer observations (up to 12 km) to identify some of the clouds. With more documentation and testing several systems have potential for use as reference type measurements (Nash et al., 2011).

Further details on radiosonde humidity uncertainties will be discussed in the final report. As is the case for temperature, if possible these will be disaggregated to a sonde-type based approach.

3.4 Literature and reports

List of literature/reports describing the instrumentation, and measurement and analysis technique including uncertainty analysis:

GCOS-144 - Guide to the GCOS Surface Network (GSN) and GCOS Upper-Air Network (GUAN) (2010 Update of GCOS-73), http://www.wmo.int/pages/prog/gcos/Publications/GCOS-144_en.pdf

GCOS-182 - Workshop on the review of the GCOS Surface Network (GSN), GCOS Upper-Air Network (GUAN), and related atmospheric networks, Ispra, Italy, April 2014, http://www.wmo.int/pages/prog/gcos/Publications/gcos-182.pdf

Ingleby B., An assessment of different radiosonde types 2015/2016, ECMWF Technical Memoranda, to be submitted 2017.

Nash, J., T. Oakley, H. Vömel, and LI Wei, WMO intercomparison of high quality radiosonde systems: Yangjiang, China, 12 July-3 August 2010. IOM Rep. 107, WMO/TD-1580, 2011. See also ref for WMO IOM 107 in section below

Philipona, R., A. Kräuchi, G. Romanens, G. Levrat, P. Ruppert, E, Brocard, P. Jeannet, D. Ruffieux and B. Calpini, Solar and thermal radiation errors on upper-air radiosonde temperature measurements, J. of Atmos. and Oceanic Technol., 30, 2382-2393, doi:10.1175/JTECH-D13-00047.1, 2013.

Philipona, R., A. Kräuchi, G. Romanens, G. Levrat, P. Ruppert, D. Ruffieux and B. Calpini, Upper-air radiosonde intercomparisons and uncertainty estimates, 2014. https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-116_TECO-2014/Session%203/K3_Philipona_etal_Radiosonde_Uncertainty.pdf

3.5 Intercomparisons and corresponding literature references

WMO radio intercomparisons go back as far as 1950 when an international comparison of six different radiosonde types was organised in Payerne, Switzerland. The results of this intercomparison were rather confusing because large differences between various radiosondes were found and the participants were unable to agree on suitable reference values or instruments. A second comparison was held in Payerne in 1956. This time, fourteen different radiosondes types were compared. It turned out that significant systematic differences still existed between different types and the participants were convinced that routine radiosondes should be compared with internationally accepted reference radiosondes (Huovila, 1987).

The development of the reference radiosondes proceeded slowly but in 1967 the CIMO working on radiosonde and radiowind measurements noted the existence of five reference thermos-metric sonde systems and recommended to organise a series of comparisons between them. Following this recommendation, three bilateral comparisons were held in Germany in 1968, in Japan in 1968 and in the (then) USSR in 1969. It turned out that the systematic temperature differences between the participating sondes were very small and it was concluded that any of these types may be used as a temperature reference (Huovila, 1987).

Furthermore, intensive research and development was initiated by several countries in the 1970's to modernize and improve standard radiosondes. The rapid evolution of electronics provided an opportunity to introduce superior radiosonde components, digital technology and partial or total automation of data treatment. The ninth WMO Congress in 1983 recommended therefore the organization of an international WMO radiosonde intercomparison which lead to a series of WMO radiosonde intercomparisons conducted under the auspices of CIMO (Commission for Instruments and Methods of Observation) which took place in the UK in 1984, in the USA in 1985, in the former USSR in 1989, in Japan in 1993, in the USA/Russian Federation in 1995-97 in Brazil in 2001, and in Mauritius Island in 2005 (Jeannet et al., 2008). The most recent intercomparison, the WMO Intercomparison of High Quality Radiosonde Systems, was held in Yangjiang, China in 2010 (also discussed in the previous section). While in earlier comparisons the different parameters were measured up to 100 hPa, measurements are made up to 10 hPa since the 1980s. Figure 4 shows the temperature differences observed during the night at the different intercomparisons and the improvements made since 1984 (Philipona et al., 2014).

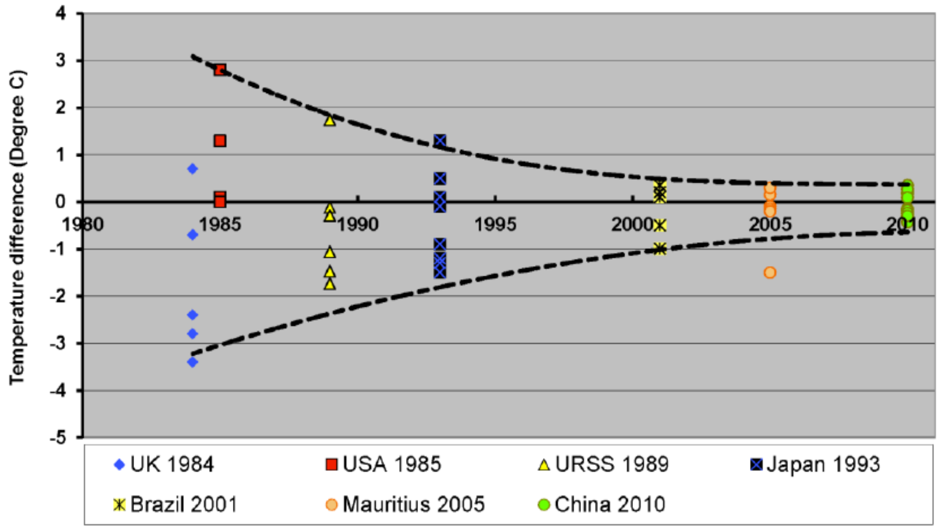

Figure 4. Mean temperature differences observed during the night at 10 hPa (figure from Philipona et al., 2014).

The first four CIMO Radiosonde Comparisons in the 1980s compared most of the radiosonde types in operational use between 1984 and 1993. However, further advances in computing and electronics suitable for use on radiosondes, plus an understanding of sensor errors that could be avoided without excessive expense have greatly improved the performance of the better-quality radiosondes.

The 5th Radiosonde Comparison (1995-97) was mainly devoted to humidity measurements and accelerated the change from the carbon hygristor and goldbeaters skin to thin film capacitative humidity sensors, which now prevail in current radiosonde systems.

The 6th Radiosonde Comparison in Brazil (2001) saw the introduction of new GPS radiosonde designs, that had much better radiofrequency electronics than was generally used until that time. This solved the problems of measuring winds with GPS in tropical conditions, which was the major issue with radiosonde systems at the time.

By the time of the 7th Comparison in Mauritius (2005), the new designs had matured and it was possible for the first time to see the high reproducibility of GPS geopotential height measurements. The basic GPS geometric heights were readily converted into geopotential heights using the variation of gravity with height at the given latitude. The results suggested that it might no longer be necessary to use pressure sensors on the radiosondes, since the pressure could be computed from the geopotential height plus the temperature and relative humidity profile to that height. The test also showed that some operational radiosonde systems were starting to measure relative humidity in the upper troposphere without excessive errors. Day-night differences in relative humidity were quantified for many types of operational radiosonde.

The corresponding references are listed below:

Huovila, S., Summary of WMO radiosonde intercomparisons, ECMWF/WMO Workshop on Radiosonde Data Quality and Monitoring, 14-16 December 1987, Conference Paper, ECMWF, http://www.ecmwf.int/sites/default/files/elibrary/1987/10062-summary-wmo-radiosonde-intercomparisons.pdf

Jeannet, P., C. Bower and B. Calpini, Global criteria for tracing the improvements of radiosondes over the last decades, IMO Rep. 95, WMO/TD-1433, 32pp., 2008.

WMO IOM 28 - WMO International Radiosonde Comparison - Phase I, Beaufort Park (UK), 1984 (A. H. Hooper (United Kingdom), 1986).

WMO IOM 29 - WMO International Radiosonde Comparison - Phase II, Wallops Island (USA), 1985 (F. Schmidlin (USA)).

WMO IOM 40 - WMO International Radiosonde Comparison - Phase III, Dzhambul (USSR), 1989 (A.Ivanov, A. Kats, S. Kurnosenko, N. Zaiseva (all USSR) and J. Nash (UK)).

WMO IOM 59 - WMO International Radiosonde Comparison - Phase IV, Tsukuba (Japan), 1993 (S. Yagi, A. Mita and N. Inoue (all Japan), 1996).

WMO IOM 90 - WMO Intercomparison of GPS Radiosondes – (Phase V), Alcantara (Brazil), 2001 (R.B. da Silveira, G. Fisch, L.A.T. Machado, A.M. Dall´Antonia Jr., L.F. Sapucci, D. Fernandes (all Brazil), and J. Nash (UK), 2006).

WMO IOM 83 – The WMO Intercomparison of Radiosondes Systems – (Phase VI), Vacoas (Mauritius), 2005 (J. Nash, R. Smout, T. Oakley (all UK), B. Pathack (Mauritius), S. Kurnosenko (USA), 2006).

WMO IOM 107 – WMO Intercomparison of High Quality Radiosondes Systems – (Phase VII), Yangjiang (China), 2010 (J. Nash, T. Oakley (all United Kingdom), H. Vömel (Germany), LI Wei (China), 2011).

4. MWRnet

4.1 Brief description of the ECV/technique/network

Ground-based microwave radiometers (MWR) provide, depending on channels and characteristics, atmospheric temperature and humidity profiles as well as integrated values of water vapour and liquid water. MWR can make continuous observations (time scales of seconds to minutes) in a long-term unattended mode in nearly all weather conditions.

MWR measure natural atmospheric thermal emission at one to several frequency channels. The measured voltages are calibrated into brightness temperatures (Tb) at each channel and then processed to derive the geophysical parameters through the application of inversion techniques.

The atmospheric parameters that can be retrieved from MWR observations depend upon the channel specifications. Channels in the 22-35 GHz band provide information on vapour and cloud liquid water. Two channels (usually 23.8 and 30-31 GHz) are enough to retrieve the column-integrated total water vapour content (TWVC) and total liquid water content (TLWC) simultaneously. More channels provide information on the vertical distribution of water vapour content, i.e. as a function of altitude, though at low vertical resolution. Observations at 50-60 GHz band provide information on atmospheric temperature, either by single-channel observations at several elevation angles or by multi-channel observations at one or more elevation angles. Units with channels in both the 22-30 and the 50-60 GHz bands are often called temperature and humidity profilers.

The International Network of Ground-based Microwave Radiometers (MWRnet) currently links about 61 members, operating more than 100 MWR worldwide (including dual-channel units, water vapour profilers, single-channel temperature profilers, multi-channel temperature profilers, and temperature and water vapour profilers), of which about 30 are located in Europe (mainly temperature and water vapour profilers). More information on MWRnet and the network map are available at the MWRnet website (http://cetemps.aquila.infn.it/mwrnet/).

4.2 MWRnet maturity matrix

The maturity assessment for MWRnet was performed within GAIA-CLIM in September 2016 http://www.gaia-clim.eu/system/files/document/MWRnet%20with%20text%20.jpg. It assesses certain quantifiable aspects of typical measurement system maturity across the network for those ECVs and associated measurement systems that are relevant to GAIA-CLIM.

MWRnet						
Metadata	**Documentation**	**Uncertainty characterization**	**Public access, feedback and update**	**Usage**	**Sustainability**	**Software (optional)**
Standards	Formal Description of Measurement Methodology	Traceability	Access	Research	Siting environment	Coding standards
Collection level	Formal Validation Report	Comparability	User feedback mechanism	Public and commercial exploitation	Scientific and expert support	Software documentation
File level	Formal Measurement Series User Guidance	Uncertainty Quantification	Updates to record		Programmatic support	Portability and numerical reproducability
		Routine Quality Management	Version control			Security
			Long term data preservation			
Legend						
1	2	3	4	5	6	Not applicable

Table 3. Maturity matrix assessing the MWRnet (International Network of Ground-based Microwave Radiometers)[2].

Typically, a reference quality measurement program would score 5s and 6s against relevant criteria, a baseline capability 3s and 4s and a comprehensive capability 1s and 2s. While some entries score higher, all subcategories in the uncertainty characterization and public access, feedback and update categories of the MWRnet maturity matrix score 3s and 4s which labels the network clearly as of baseline status.

Due to the bottom-up nature of MWRnet, there are currently different levels of maturity within the network itself. For example, the MWR belonging to the U.S. Atmospheric Radiation Measurement (ARM) program has indeed higher scores than average concerning metadata, documentation, public access, and sustainability. In Europe, MWR contributing to HD(CP)2 and Towards Operational ground-based PROFiling with ceilometers, doppler lidars and microwave radiometers (TOPROF) also have higher scores than average. In addition, the GRUAN MWR guidelines document has the ambition to set the guidelines for producing reference level MWR observations. The MWR observations contributing to the GAIA-CLIM Virtual Observatory are

[2] *Users of GAIA-CLIM's Maturity Matrices, should be aware that this is a first effort to systematically quantify measurement system performance. Redundant assessments suggest a minimum uncertainty arising from assessor-to-assessor variations in any category of at least 1 score. Although the assessment may be useful for certain applications, at this time and until more broadly tested, it should not constitute a primary or unique decision making tool.*

from a selected subset of the above virtuous members. This guidance and assessment pertains to the remainder of the network only.

4.3 Description of (indicative) measurement uncertainties

Nowadays, off-the-shelf commercial MWR are robust instruments providing continuous unattended operations and real time accurate atmospheric observations at ~1 min temporal resolution under nearly all-weather conditions. The MWR instruments are typically commercial grade instruments. Random and systematic uncertainties are typically instrument dependent. Systematic uncertainties will also depend upon periodic calibrations against cryogenic targets. Commercial MWR units are usually offered with azimuth- and elevation-angle scanning capability. For individual MWR instruments, manufacturer claimed estimates are available. These have been validated by independent investigators (Maschwitz et al., 2013). When properly calibrated, a MWR provides Tb with an absolute accuracy of ~0.3-1.0 K.

The uncertainty of derived products is typically estimated from simulated datasets and, when available, from rms difference with respect to independent reference measurements (usually radiosondes). **Typical uncertainties for derived products are** (Cimini et al., 2006; Löhnert and Maier, 2012):

TWVC ~ 1.0 mm (or kg/m^2)

TLWC ~ 0.02 mm (or kg/m^2)

T(z) ~ 0.5 - 2.0 K (decreasing from surface up)

WV(z) ~ 0.2 - 1.5 g/m^3

4.4 Literature and reports

List of literature and reports describing the instrumentation, and measurement and analysis technique including uncertainty analysis. Manufacturers provide specification documents for their MWR instruments. These documents typically include manufacturer claimed uncertainties.

Cadeddu, M. P., Liljegren, J. C., and Turner, D. D.: The Atmospheric radiation measurement (ARM) program network of microwave radiometers: instrumentation, data, and retrievals, Atmos. Meas. Tech., 6, 2359-2372, doi:10.5194/amt-6-2359-2013, 2013.

Cimini D., O. Caumont, U. Löhnert, L. Alados-Arboledas, R. Bleisch, J. Fernández-Gálvez, T. Huet, M. E. Ferrario, F. Madonna, O. Maier, F. Nasir, G. Pace, and R. Posada, An International Network of Ground-Based Microwave Radiometers for the Assimilation of Temperature and Humidity Profiles into NWP Models, Proceedings of 9th International Symposium on Tropospheric Profiling, ISBN 978-90-815839-4-7, L'Aquila, ITALY, 3-7 September 2012.

Cimini, D., T. J. Hewison, L. Martin, J. Güldner, C. Gaffard and F. S. Marzano, Temperature and humidity profile retrievals from ground-based microwave radiometers during TUC, Met. Zeit., Vol. 15, No. 1, 45-56, 2006.

Löhnert U. and O. Maier, Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges, Atmos. Meas. Tech., 5, 1121-1134, doi:10.5194/amt-5-1121-2012, 2012.

Maschwitz, G., U. Löhnert, S. Crewell, T. Rose, and D.D. Turner, 2013: Investigation of Ground-Based Microwave Radiometer Calibration Techniques at 530 hPa, Atmos. Meas. Tech., 6, 2641–2658, doi:10.5194/amt-6-2641-2013.

EU COST EG-CLIMET Final report: http://cfa.aquila.infn.it/wiki.eg-climet.org/index.php5/Final_Report

Chapter on MWR: http://cfa.aquila.infn.it/wiki.eg-climet.org/index.php5/ Microwave_radio-meter

Chapter on MWRnet: http://cfa.aquila.infn.it/wiki.eg-climet.org/index.php5/Final_Report#MWRnet

4.5 Intercomparisons and corresponding literature references

Caumont O., D. Cimini, U. Löhnert, L. Alados-Arboledas, R. Bleisch, F. Buffa, M. E. Ferrario, A. Haefele, T. Huet, F. Madonna, G. Pace: Assimilation of humidity and temperature observations retrieved from ground-based microwave radiometers into a convective-scale model, Quart. Jour. Roy. Met. Soc., doi:10.1002/qj.2860, 2016.

Cimini D., O. Caumont, U. Löhnert, L. Alados-Arboledas, R. Bleisch, T. Huet, M. E. Ferrario, F. Madonna, A. Haefele, F. Nasir, G. Pace, and R. Posada, A data assimilation experiment of temperature and humidity profiles from an international network of ground-based microwave radiometers, Proc. Microrad 2014, Pasadena, USA, 24-27 March, ISBN: 978-1-4799-4645-7, 978-1-4799-4644-0/14/$31.00, 2014.

Cimini, D., T. J. Hewison, L. Martin, Comparison of brightness temperatures observed from ground-based microwave radiometers during TUC, Meteorologische Zeitschrift, Vol.15, No.1, 2006, pp.19-25, 2006.

Cimini, D., E. R. Westwater, Y. Han, and S. J. Keihm: Accuracy of Ground-based Microwave Radiometer and Balloon-Borne Measurements During WVIOP2000 Field Experiment, IEEE Transactions on Geoscience and Remote Sensing, Vol. 41, n. 11, pp. 2605-2615, 2003.

Illingworth, A., D. Cimini, C. Gaffard, M. Haeffelin, V. Lehmann, U. Loehnert, E. O'Connor, D. Ruffieux, Exploiting Existing Ground-Based Remote Sensing Networks To Improve High Resolution Weather Forecasts, Bull. Amer. Meteor. Soc. doi: 10.1175/BAMS-D-13-00283.1, February, 2015.

Reports of EU COST TOPROF Joint microwave calibration experiment (J-CAL) 1 and 2: http://www.toprof.imaa.cnr.it/images/toprof/sub_working_group/Scientific_Report_20140825_Lindenberg.pdf,

http://www.toprof.imaa.cnr.it/images/toprof/sub_working_group/Scientific_Report_JCAL2_final.pdf

5. SHADOZ – ozonesonde network

5.1 Brief description of the ECV/technique/network

The electrochemical concentration cell (ECC) ozone instrument, developed by Komhyr (1969), combines the basic principle of the reaction of ozone and iodide within a redox cell sensor coupled to a nonreactive air-sampling Teflon pump. The ozone instrument, when interfaced with a balloon-borne radiosonde, provides a reliable and relatively inexpensive method to measure ozone concentrations from the surface to altitudes of about 35 km.

The earliest ECC ozonesonde sites began measuring ozone profiles in the late 1960s. Since then, ozonesondes have become increasingly important in monitoring tropospheric and lower stratospheric ozone. For example, trends derived from ozonesonde data at several sites are a significant part of ozone assessments. Ozonesondes have been used extensively in projects such as SHADOZ (Thompson et al., 2003a; 2003b, 2004, 2007, 2012) where ozone profiles are gathered and compared from many different locations. Therefore, it is important that ozonesonde data sets are consistent and that the measurement uncertainties are well understood.

The SHADOZ (Southern Hemisphere Additional OZonesondes) network brings together ozonesonde data from 13 tropical and sub-tropical sites (Figure 5) and has become the central repository for vertical profiles of ozone in the tropics/sub-tropics which makes SHADOZ an important network for equatorial tropospheric ozone research.

The rationale for SHADOZ is to: (1) validate and improve model and remote sensing techniques for estimating tropical ozone, (2) contribute to climatology and trend analyses of tropical ozone and (3) provide research topics to scientists and help educate students, especially in participating countries. SHADOZ is envisioned as a data service to the global scientific community by providing a central public archive location via the internet: http://croc.gsfc.nasa.gov/shadoz. The SHADOZ website maintains a standard data format for the archive and it informs users of the differing sites' preparation techniques and data treatment. Data from launches from various SHADOZ supported field campaigns, such as, the Indian Ocean Experiment (INDOEX) and Sounding of Ozone and Water in the Equatorial Region (SOWER) are also available.

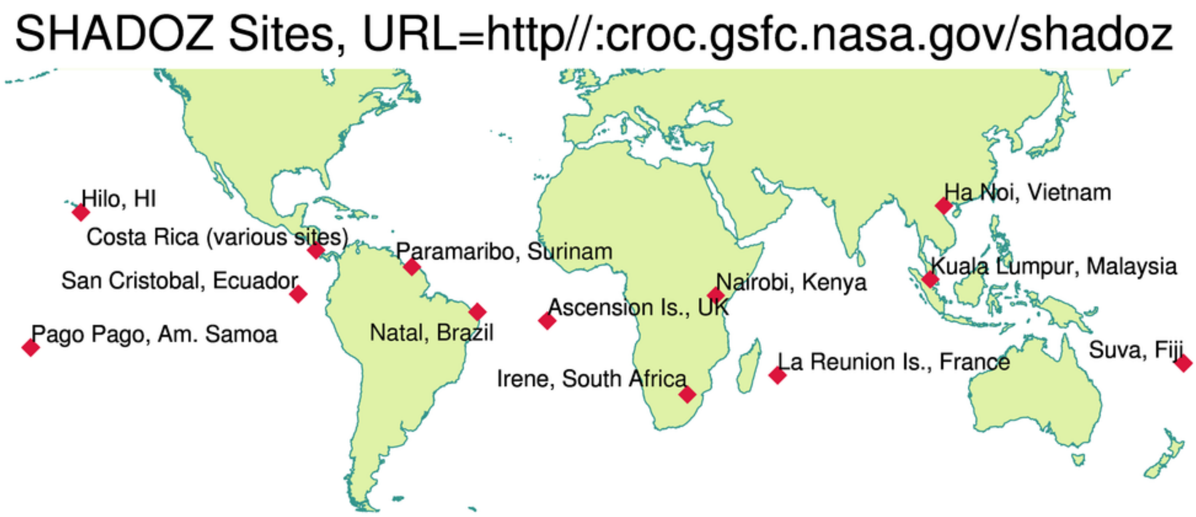

Figure 5. Map showing the SHADOZ ozonesonde stations.

5.2 SHADOZ maturity matrix

The maturity assessment for SHADOZ was performed within GAIA-CLIM in September 2016, http://www.gaia-clim.eu/system/files/document/SHADOZ%20with%20text%20.jpg. It assesses certain quantifiable aspects of typical measurement system maturity across the network for those ECVs and associated measurement systems that are relevant to GAIA-CLIM.

Typically, a reference quality measurement program would score 5s and 6s against relevant criteria, a baseline capability 3s and 4s and a comprehensive capability 1s and 2s. And although many of the entries for SHADOZ, such as all entries for the documentation category, have a score of 5 or 6, there are several other subcategories, especially for the uncertainty characterization category that are of score 3 and 4 which means that overall SHADOZ currently has a baseline status.

SHADOZ						
Metadata	Documentation	Uncertainty characterization	Public access, feedback and update	Usage	Sustainability	Software (optional)
Standards	Formal Description of Measurement Methodology	Traceability	Access	Research	Siting environment	Coding standards
Collection level	Formal Validation Report	Comparability	User feedback mechanism	Public and commercial exploitation	Scientific and expert support	Software documentation
File level	Formal Measurement Series User Guidance	Uncertainty Quantification	Updates to record		Programmatic support	Portability and numerical reproducability
		Routine Quality Management	Version control			Security
			Long term data preservation			
Legend						
1	2	3	4	5	6	Not applicable

Table 4. Maturity matrix assessing SHADOZ (Southern Hemisphere Additional OZonesondes). [3]

5.3 Description of (indicative) measurement uncertainties

The assessment of the uncertainty budget of ozone measured by ozonesondes is a complex task. Measurement uncertainties in ozonesonde data should, in the first instance, be characterized in the laboratory, and preferably, these laboratory results should be supported with field campaigns which are likely to test ozonesondes in an environment closer to their standard operating environment than in the laboratory environment e.g. the BESOS (Balloon Experiment on Standards for Ozonesondes) campaign (Deshler et al., 2008).

Factors influencing sonde performance: Each ozone sounding is made with a new instrument which must be characterized prior to flight. Consistency of instruments with regard to their quality and characteristics, but also standardization of operating procedures, is a pre-requisite to assure consistent sonde measurements. Several instrumental and procedural parameters and their uncertainties can have a substantial influence on the quality of the ozonesonde

[3] *Users of GAIA-CLIM's Maturity Matrices, should be aware that this is a first effort to systematically quantify measurement system performance. Redundant assessments suggest a minimum uncertainty arising from assessor-to-assessor variations in any category of at least 1 score. Although the assessment may be useful for certain applications, at this time and until more broadly tested, it should not constitute a primary or unique decision making tool.*

measurements. Changes of these parameters through changes in instrument, operating procedures, or environmental conditions can have a significant impact on the long term ozone trends derived from ozonesonde measurements. From intercomparisons between different sounding stations using the same sonde type it has been shown that observed differences are, in large part, due to differences in the preparation and correction procedures applied at the different launch sites (e.g. GAW Report No. 201, 2013 and Smit et al., 2007).

The Ozone Uncertainty Equation

The partial pressure of ozone measured by the electrochemical sensor is a function of the measured sensor current (I_M), the background current (I_B), the conversion efficiency (η_C), the temperature of the gas sampling pump (T_P) and the volumetric flow rate (Φ_P).

$$P_{O_3} = P_{O_3}\left(\mathrm{I}_{M,}, \mathrm{I}_B, \eta_C, \Phi_P, \mathrm{T}_P \right)$$

The instrumental uncertainty of the electrochemical ozone sensor for the measurement of ozone is a composite of the contributions of the individual uncertainties of the different instrumental parameters listed above. Some of the contributions depend on air pressure in such a way that the overall uncertainty of the ozone measurement will be a function of pressure i.e. altitude.

Instrumental uncertainties are assumed to be random and gaussian, and therefore follow the gaussian law of error propagation. The overall relative uncertainty of ozone partial pressure (P_{O3}) can be expressed as the sum of the squares of the uncertainty in each term of the P_{O3} equation as can be seen below in the WMO/GAW ozone uncertainty equation as taken from Smit et al, (2014):

$$\frac{\Delta \mathbf{P_{O3}}}{\mathbf{P_{O3}}} = \sqrt{\frac{(\Delta \mathbf{I_M})^2 + (\Delta \mathbf{I_B})^2}{(\mathbf{I_M} - \mathbf{I_B})^2} + \left(\frac{\Delta \eta_C}{\eta_C}\right)^2 + \left(\frac{\Delta \mathbf{\Phi_P}}{\mathbf{\Phi_P}}\right)^2 + \left(\frac{\Delta \mathbf{T_P}}{\mathbf{T_P}}\right)^2}$$

The overall relative uncertainty of P_{O3} determined by the contributions from the individual uncertainties of the different instrumental parameters are demonstrated for an ECC type ozonesonde in Figure 6 for a typical mid-latitude (A) and tropical (B) vertical ozone profile. In the troposphere the overall uncertainty of P_{O3} (red solid line) is dominated by the contribution of the uncertainty of the background current I_B (blue solid line), while in the stratosphere the uncertainties of the conversion efficiency η_C (purple solid line) and pump flowrate Φ_P (cyan solid line) are the major contributions. **Figure 6 shows the relative uncertainty in ozone in % as well as the absolute uncertainty in the ozone partial pressure (P_{O3}) in mPa for altitudes from 0 to 35 km.** The different instrumental uncertainties and their influence on the overall uncertainty of P_{O3} are discussed in more detail in Sections 3.2.2 - 3.2.5 of the GAW Report No. 201 (2013).

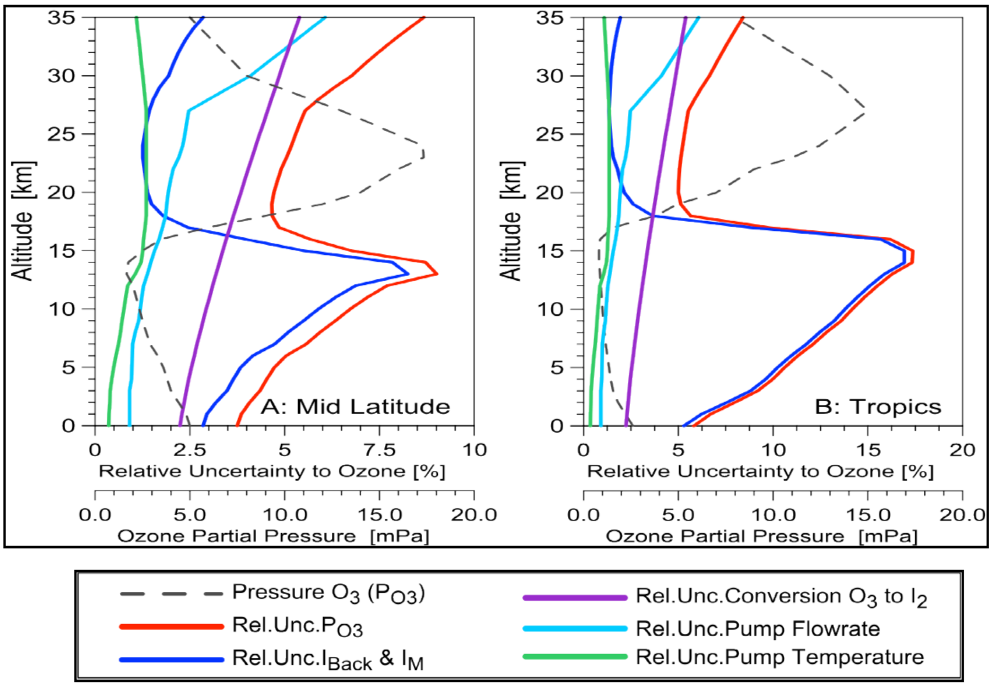

Figure 6. For the ECC-ozonesonde type, the overall relative uncertainty of P_{O3} and the contributions from the individual uncertainties of the different instrumental parameters are shown. Individual uncertainties are: the measured cell current (I_M=0.1-5μA, ΔI_M=±0.01μA) and background current (I_B≈0.07μA, ΔI_B≈±0.04μA), conversion efficiency (η_C≈1, $\Delta\ \eta_C$≈±(0.05-0.07)), pump flowrate Φ_P, and pump temperature T_P as a function of altitude for a typical mid-latitude (A) and tropical (B) ozone profile.

5.4 Literature and reports

Komhyr, W.D., Electrochemical concentration cells for gas analysis, *Ann.Geoph.*, 25, 203-210, 1069.

Smit, H. G. J. and the Panel for the Assessment of Standard Operating Procedures for Ozonesondes (ASOPOS), Quality assurance and quality control for ozonesonde measurements in GAW, World Meteorological Organization, 2014.

Thompson, A.M., et al., Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements, J. Geophys. Res., 108, 8238, doi:10.1029/2001JD000967, 2003a.

Thompson, A.M., et al., Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2000 tropical ozone climatology 2. Tropospheric variability and the zonal wave-one, J. Geophys. Res., 108, 8241, doi:10.1029/2002JD002241, 2003b.

Thompson, A., J. Witte, S. Oltmans, and F. Schmidlin, Shadoz - A tropical ozonesonde radiosonde network for the atmospheric community, Bulletin of the American Meteorological Society, 85(10), doi: 10.1175/BAMS-85-10-1549, 2004.

Thompson, A.M., J.C. Witte, H.G.J. Smit, S.J. Oltmans, B.J. Johnson, V.W.J H. Kirchhoff, and F.J. Schmidlin, Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles, J. Geophys. Res., 112, D03304, doi:10.1029/2005JD007042, 2007.

Thompson, A. M., et al., Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone climatology (2005-2009): Tropospheric and tropical tropopause layer (TTL) profiles with comparisons to OMI-based ozone products, Journal of Geophysical Research-Atmospheres, 117, doi: 10.1029/2011JD016911, 2012.

5.5 Intercomparisons and corresponding literature references

To assess the performance of the ozonesondes and to quantify any systematic differences among the various sonde types, several intercomparisons have been carried out since 1970. A comprehensive review of the performance of ozonesondes in terms of precision and accuracy has been given in the *SPARC-IOC-GAW Assessment of Trends in the Vertical Distribution of Ozone* (1998).

Some earlier intercomparisons included only ozonesondes, without comparisons to a reference profile measured by a different technique while other intercomparisons did use a reference profile measured by other techniques. However, to further quantify the precision and accuracy of the three different types of ozonesondes, several comparison studies of sondes with other ozone profiling techniques have been made since 1970 (Smit et al., 2014 and references within). Most intercomparison studies, particularly before 1990, focussed exclusively on sonde performance in the stratosphere. However, due to the much lower concentrations of ozone in the troposphere compared to the stratosphere the performance of the sondes and their typical instrumental and operational factors determining precision and accuracy are rather different in the two regions of the atmosphere. Since the 1990's, investigations have been completed to also address the performance of ozonesondes also in the troposphere (GAW report and references therein).

These short-term intercomparisons are more or less "snap shots" and may not necessarily reflect the performance of ozonesondes under operational field conditions. Long-term comparison studies of ozonesonde data with other simultaneously operating ozone monitoring devices like lidar or microwave are more suitable to assess the data quality of the ozonesonde measurements in regular operation.

In addition, intercomparisons like the series of JOSIE (Jülich Ozonesonde Intercomparsion Experiment) experiments (Smit et al., 2007) were conducted in a controlled environmental chamber capable of simulating real sounding conditions. The JOSIE experiments were conducted to specifically address accuracy and response of the ozonesondes as a function of sonde type, altitude, and ozone level. The different ozonesonde types were tested under a variety of conditions and compared with an accurate ozone UV-photometer as reference. Special attention was paid to the influence of pre-launch procedures on in-flight performance. In addition, the influence of certain procedures, such as background signal correction and total ozone column normalization, was investigated. The JOSIE 1996-2000 experiments, their design and results are presented in detail in three GAW reports (Smit and Kley, 1998 and Smit and Straeter, 2004a & 2004b) and evaluated in peer reviewed literature by Smit et al. (2007).

In 2004 during BESOS (Balloon Experiment on Standards for Ozonesondes), the results from JOSIE were tested in the field by a balloon flight carrying a core of 12 ECC sondes and an in-situ photometer (Deshler et al., 2008). The JOSIE and BESOS experiments and the ASOPOS (Assessment of Standard Operating Procedures for OzoneSondes) initiative have improved the homogeneity and quality of ozone sounding data and ensure more confidence in the observed trends.

Table 5 shows a summary of the precision and accuracy of the different types of ozonesondes over 5 km altitude bins from the surface up to 35 km. The presented data are a synthesis obtained from SPARC-IOC-GAW (1998), JOSIE (Smit et al., 2007) and BESOS (Deshler et al., 2008). Precision here is confined to reproducibility or repeatability and can be characterized in terms of the standard deviation of the sonde measurements when exposed to the same ozone input. Precision is related to random errors of the sonde. This is in contrast to bias, which is the difference of the sonde from an ozone reference instrument and is associated with systematic errors of the sonde. Accuracy is here defined as the sum of bias and precision. The precision and accuracy listed here were obtained from experimental intercomparison fits within the theoretical estimate of the overall relative uncertainty of an ozonesonde described in the previous section (see Figure 6).

Sonde Type	ECC SPC-6A SST1.0			ECC ENSCI-Z SST0.5			BM MOHp			KC96 JMA		
Altitude Range [km]	Bias [%]	Precision [%]	Accuracy [%]	Bias[%]	Precision [%]	Accuracy [%]	Bias [%]	Precision [%]	Accuracy [%]	Bias [%]	Precision [%]	Accuracy [%]
30-35	2	7	9	6	6	12	-2	13	15	11	7	18
25-30	4	5	9	7	5	12	0	5	5	5	6	11
20-25	1	5	6	5	5	10	0	5	5	-2	6	8
15-20	1	5	6	2	3	5	-1	3	4	-7	5	12
10-15	-2	4	6	-1	3	4	-3	5	8	-6	5	11
5-10	0	4	4	-1	4	5	-3	8	11	-2	4	6
0-5	-3	4	7	-1	4	5	-3	10	13	-7	6	13

Table 5. Survey of the average relative bias to UV-Photometer and relative precision of ECC-sonde types SPC-6A (operated with SST1.0: 1.0%KI&full buffer) and ENSCI-Z (operated with SST0.5: 0.5%KI and half

buffer), BM-sonde operated by Meteorological Observatory Hohenpeissenberg (MOHp) and KC96-sonde operated by Japan Meteorological agency (JMA). The listed data are averaged over altitude bins of 5 km. The accuracy is determined as the sum of bias and precision. Results are representative for mid-latitude atmospheric conditions and are derived from SPARC-IOC-GAW (1998), JOSIE (*Smit et al.*, 2007) and BESOS (*Deshler et al.*, 2008) experiments.

The corresponding references are listed below:

Deshler, T., J. Mercer, H.G.J. Smit, R. Stubi, G. Levrat, B.J. Johnson, S.J. Oltmans, R. Kivi, J. Davies, A.M. Thompson, J. Witte, F.J. Schmidlin, G. Brothers, T. Sasaki, Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes, J. Geophys. Res., 113, D04307, doi:10.1029/2007JD008975, 2008.

Smit H.G.J and D. Kley (1998), JOSIE: The 1996 WMO International intercomparison of ozonesondes under quasi flight conditions in the environmental simulation chamber at Jülich, WMO Global Atmosphere Watch Report No. 130, WMO TD No. 926, World Meteorological Organization, Geneva.

Smit, H.G.J., and W. Straeter (2004a), JOSIE-1998, Performance of ECC Ozone Sondes of SPC-6A and ENSCI-Z Type, WMO Global Atmosphere Watch Report No. 157, WMO TD No. 1218), World Meteorological Organization, Geneva.

Smit, H.G.J., and W. Straeter (2004b), JOSIE-2000, Jülich Ozone Sonde Intercomparison Experiment 2000, The 2000 WMO international intercomparison of operating procedures for ECCozonesondes at the environmental simulation facility at Jülich, WMO Global Atmosphere Watch Report No. 158, WMO TD No. 1225, World Meteorological Organization, Geneva.

Smit, H.G.J., W. Straeter, B. Johnson, S. Oltmans, J. Davies, D.W. Tarasick, B. Hoegger, R. Stubi, F. Schmidlin, T. Northam, A. Thompson, J. Witte, I. Boyd, F. Posny, Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Jülich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res., 112, D19306, doi:10.1029/2006JD007308, 2007.

Smit, H. G. J. and the Panel for the Assessment of Standard Operating Procedures for Ozonesondes (ASOPOS), Quality assurance and quality control for ozonesonde measurements in GAW, World Meteorological Organization, 2014.

SPARC-IOC-GAW Assessment of Trends in the Vertical Distribution of Ozone (1998) (http://www.sparc-climate.org/fileadmin/customer/6_Publications/SPARC_reports_PDF/1_Ozone_SPARCreportNo1_May1998_redFile.pdf).

6. GSN

6.1 Brief description of the ECV/technique/network

The GCOS Surface Network (GSN) is a designated subset of the WMO standard meteorological networks. GSN stations are required to produce monthly CLIMAT reports. The GSN operations are covered by GCOS-144 (http://www.wmo.int/pages/prog/gcos/Publications/GCOS-144_en.pdf), which documents the scope of operations that is intended.

The GSN is intended to comprise the best possible set of land stations with a spacing of 2.5 to 5 degrees of latitude (Figure 7), thereby allowing coarse-mesh horizontal analyses for some basic parameters (primarily temperature and precipitation). The criteria for selection include:

- Commitments by NMHSs with regard to continuity
- Geographical representativeness of observations
- Length and quality of historical time series
- Available parameters.

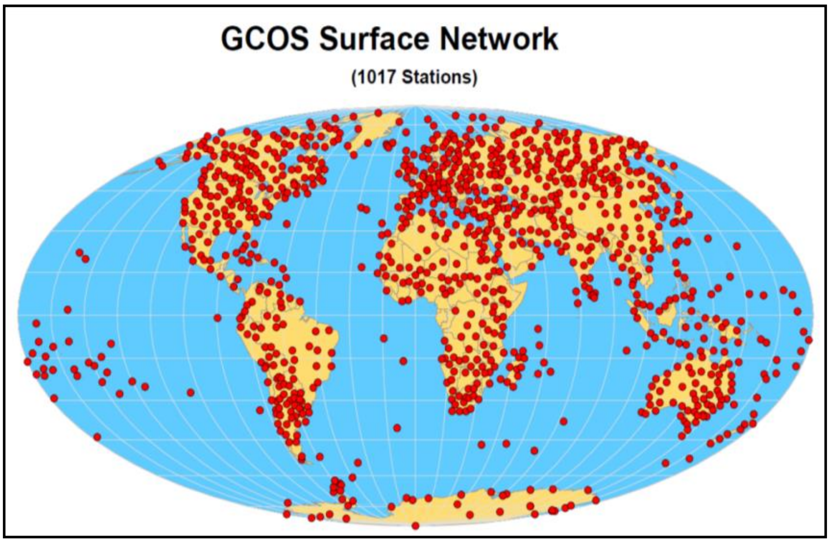

Figure 7. Map showing the 1017 GSN (GCOS Surface Network) stations in 2014.

It is recognized that the coarse network density limits the applicability for some applications. For parameters having smaller-scale horizontal variability (e.g., precipitation), it is accepted that the network data generally should be supplemented by those from networks with a finer mesh. The purposes of the GSN are the following:

- To establish national commitments for the preservation of a set of valuable climate stations for the foreseeable future.
- To build a collection of validated data from these stations in standardized formats.
- To provide this information to the global climate community with no formal restrictions.

- To create a baseline and benchmark data set for more enhanced regional and sub-regional climate networks and for newly-developed observing systems, including remote-sensing systems.

The GSN network is supported by two monitoring centres and a set of CBS lead centres whose mandate is to aid members in ensuring operation of and access to data from their GSN stations. The DWD hosted monitoring centre is at http://www.dwd.de/EN/specialusers/water_management/gsnmc/_node_gsnmc_dataset.html and provides quality controlled monthly summaries.

Per GCOS-144 ECVs and reporting requirements for GSN are:

Minimum Requirements:

- Monthly means of daily maximum, minimum and mean temperature
- Monthly precipitation amounts.

Target Requirements in addition:

- Pressure: monthly mean values, station level and mean sea level
- Daily precipitation amounts
- Precipitation: number of days with precipitation if daily precipitation amounts are not provided
- Temperature: daily mean, minimum and maximum
- Pressure: daily mean, station level and mean sea level
- Subdaily data: historical and real-time synoptic or hourly reports, with all the data normally reported in synoptic transmissions, for the full period of record for the station.

If only monthly values are available, the number of days used in the calculation should be provided as a Minimum Requirement. In reality, GSN performance is mixed with only a subset of the designated network reporting at any given time (Figures 8 and 9).

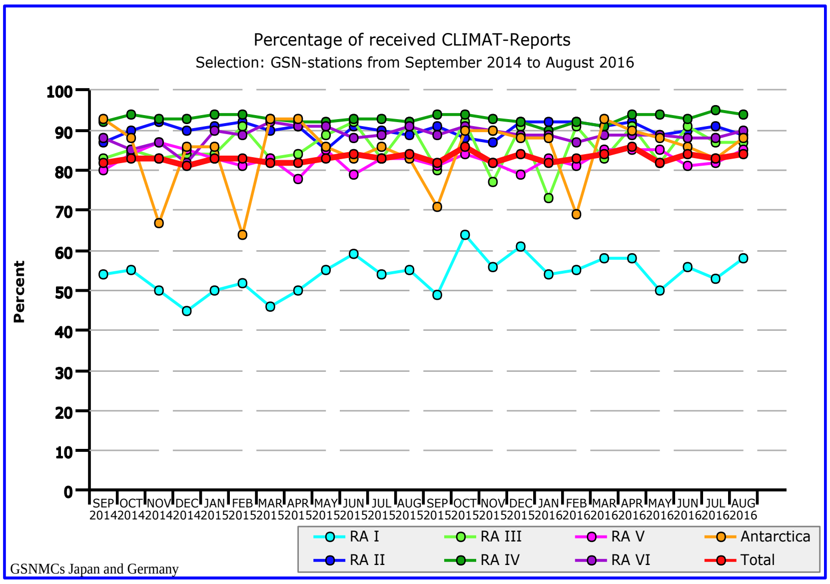

Figure 8. GSN reports received and passing QC by WMO Region. Source: DWD GSN Monitoring Centre.

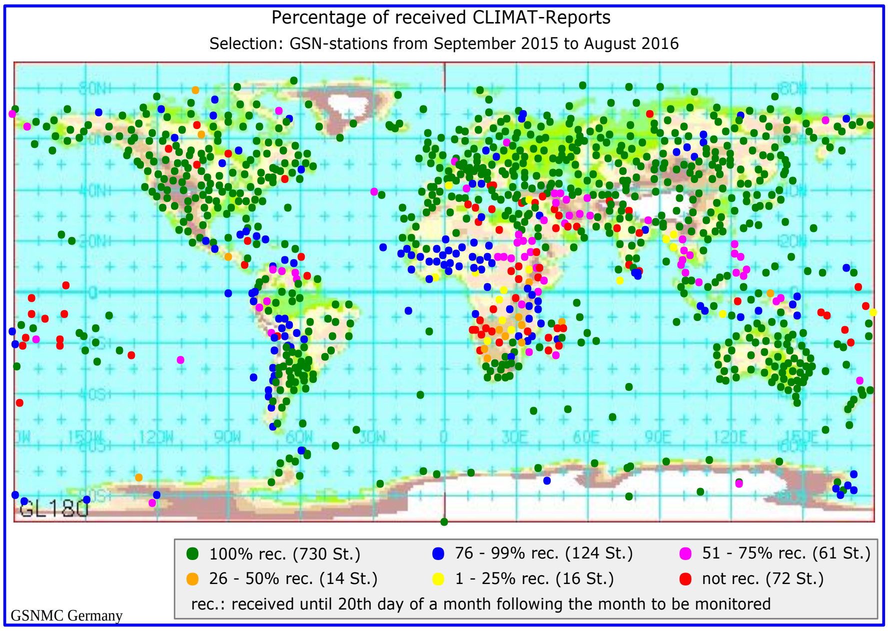

Figure 9. GSN station report reliability over past 12 months. Source: DWD GSN monitoring centre.

In addition to GSN there are the RBCN and RBSN networks that consist, again, of in-situ surface observations (http://www.wmo.int/pages/prog/www/ois/rbsn-rbcn/rbsn-rbcn-home.htm).

Regulatory material around these networks can be found at http://library.wmo.int/pmb_ged/wmo_544-v1-2015_en.pdf and their implementation in practice at http://library.wmo.int/pmb_ged/wmo_488-2013_en.pdf. Moves are afoot to change RBCN and RBSN to a new RBON network concept. For practical purposes this shall not change the network characteristics, but may greatly change the network geographical composition.

6.2 GSN maturity matrix

The maturity assessment for GSN was performed within GAIA-CLIM in September 2016, http://www.gaia-clim.eu/system/files/document/GSN%20with%20text.jpg. It assesses certain quantifiable aspects of typical measurement system maturity across the network for those ECVs and associated measurement systems that are relevant to GAIA-CLIM.

GSN						
Metadata	**Documentation**	**Uncertainty characterization**	**Public access, feedback and update**	**Usage**	**Sustainability**	**Software (optional)**
Standards	Formal Description of Measurement Methodology	Traceability	Access	Research	Siting environment	Coding standards
Collection level	Formal Validation Report	Comparability	User feedback mechanism	Public and commercial exploitation	Scientific and expert support	Software documentation
File level	Formal Measurement Series User Guidance	Uncertainty Quantification	Updates to record		Programmatic support	Portability and numerical reproducibility
		Routine Quality Management	Version control			Security
			Long term data preservation			
Legend						
1	2	3	4	5	6	Not applicable

Table 6. Maturity matrix assessing the GSN (GCOS Surface Network).[4]

Typically, a reference quality measurement program would score 5s and 6s against relevant criteria, a baseline capability 3s and 4s and a comprehensive capability 1s and 2s. And although all subcategories within the metadata category score 5s and 6s, most other entries in the GSN maturity matrix score 3s and 4s and hence GSN is clearly a baseline network.

[4] *Users of GAIA-CLIM's Maturity Matrices, should be aware that this is a first effort to systematically quantify measurement system performance. Redundant assessments suggest a minimum uncertainty arising from assessor-to-assessor variations in any category of at least 1 score. Although the assessment may be useful for certain applications, at this time and until more broadly tested, it should not constitute a primary or unique decision making tool.*

6.3 Description of (indicative) measurement uncertainties

There are a range of measurement uncertainties that shall relate to:

1. Instrumentation deployed
2. Site representativity
3. Observational practices
4. Transmission errors.

Taking these in turn:

The GSN network has a heterogeneous mix of instrumentation including manual and automatic instrumentation. The instruments are generally commercial grade instruments. For manual measurements representative random uncertainties are, at best:

- +/-0.1 or +/-0.5K for temperature and dewpoint temperature
- nearest mm for rainfall
- Nearest hPa for pressure.

For automated measurements, random uncertainties are generally at the precision of reporting, which varies by measurement system. Systematic uncertainties are instrument dependent and will depend upon whether, and if so how, periodic recalibrations against secondary standards are performed. For individual instruments, manufacturer claimed estimates may be available. Unfortunately, instrumentation metadata is not regularly exchanged.

Site representativity uncertainties will matter for many satellite calibration activities. There are two substantive issues. One is immediate site representativity, that is a function of micro-climate exposure. The second is the uncertainty that arises if the instrument is based in a heterogeneous satellite pixel footprint, which shall typically be the case. Satellite footprints are typically of the order several kilometres, and over most of the global land surface a combination of the heterogeneity of land cover types and orography, mean that any single point within such a radius is not exactly indicative of the footprint characteristics taken as a whole.

Siting guidance exists for the stations (see http://www.wmo.int/pages/prog/www/IMOP/SitingClassif/ SitingClassif.html) but few of the stations audited under this guidance attain the highest siting quality. Siting includes both location and vertical datum of measurements. Different uncertainties introduced by degraded siting are given in http://www.wmo.int/pages/prog/www/IMOP/SitingClassif/Siting-Classif_2008Ed_Up2010.pdf.

The majority of audited sites fall into categories 3 or 4 by the siting criteria. For a category 4 siting the additional uncertainty quantified in the above document amounts to:

- 2K in temperature
- 25% in precipitation.

Representativity of the observation site across a broader satellite pixel shall depend upon the satellite pixel size and the surface characteristics around the site. In principle, such an uncertainty could be inferred from high resolution digital elevation and surface type maps but would need be calculated around each site in turn which is a substantive undertaking and would depend upon the radiance spectrum space being sensed by the satellite, particularly so for (near-)window channels.

Observational practices can add additional uncertainties. Synoptic measurements reported for a given hour are typically taken at 50 minutes past the preceding hour. Some countries report the observations as instantaneous values, others as 1-minute, 5-minute or 10-minute averages. The matter becomes more complex for daily and monthly reports where a range of averaging and calculation methods can be in use both nationally and internationally which varies through time. Knowledge of the uncertainties in these is somewhat limited within the climate community, although literature may exist in the meteorological literature or grey literature.

Transmission errors occur for a variety of reasons. Poorly formatted messages can be transmitted. There are several types of error, many of which will be easily flagged (see e.g. www.metoffice.gov.uk/hadobs/hadisd). Common errors include repeated values transmission (observation over observation, day over day and month over month) and implicit decimal place errors (order of magnitude, implicit because the values are multiplied by 10 or 100 and transmitted as integers in alpha-numeric code).

Overall uncertainties are therefore hard to quantify, arising from a number of sources and requiring in-depth metadata on a per measurement site basis that is typically unavailable. The move to Binary Universal Form for the Representation of Meteorological data (BUFR) may help with metadata but only if WMO members fill the BUFR metadata headers correctly and transmit up-to-date metadata information, neither of which, sadly, is guaranteed.

Ways of getting at the uncertainty on a per site basis may include:

- Neighbour checks in data rich regions
- O-B fields from reanalyses. However, O-B systematic ‘biases’ in many instances are likely to be dominated by differences in apparent station altitude and reanalysis gridpoint versus station representativity issues
- Intercomparisons, but these may not be representative viz. the performance of the network as a whole
- Parallel measurements studies undertaken by ISTI POST (http://www.surfacetemperatures.org/databank/parallel_measurements)

Filing availability of a per site basis estimate via expert solicitation may give a reasonable estimate of the uncertainty which would be broadly applicable. Based upon the considerations outlined above an initial estimate as to the different sources of uncertainty and their potential magnitude is given in Table 7 below. This table would require solicitation from a broader range of experts before it shall be considered fit for purpose.

Uncertainty source	Temperature	Humidity (wet bulb T)	Precipitation	Pressure	Trace
Instrumental	0.5K random 1-5K systematic	0.5K random 1-5 K systematic	1mm random 1-20mm systematic	1hPa random 1-10 hPa systematic	Primarily available intercomparisons and manufacturer specifications
Representa-tivity	2K (siting, systematic) Unknown (footprint representa-tivity)	2K (siting, systematic) Unknown (footprint representa-tivity)	25% (siting, systematic) Unknown (footprint representa-tivity)	Unknown	WMO guidance cited in text
Obs practices	0.5K random	0.5K random	1/6th random for hourly measurements	Unknown	Limited trace. Needs improvement
Trans-mission	Error dependent	Error dependent	Error dependent	Error dependent	QC to HadISD and GHCND products documented in literature

Table 7. Summary of principal sources of observational uncertainty.

Based upon Table 7 and assuming independence across sources of uncertainty an indicative set of absolute uncertainties would be:

- **Temperature: 6K**
- **Dewpoint: 6K**
- **Rainfall: >30% (higher for low rainfall amounts)**
- **Pressure: 10hPa**

These may not be exhaustive and further refinement would appear necessary prior to final production of the Task 2.2 deliverable. Some of these may be systematic rather than random, so it is important to consider the desired application. For stations that are systematically biased, yet are broadly representative and have low-random uncertainties, they may be applicable to a broad range of applications despite their large absolute error.

6.4 Literature and reports

List of literature/reports describing the instrumentation, and measurement and analysis technique including uncertainty analysis. Most manufacturers provide specification documents. These typically shall include manufacturer claimed random and systematic uncertainties. Often the provenance of these claims is unknown.

Boroneant, C., M. Baciu, A. Orzan. On the statistical parameters calculated for the essential climatological variables during 2-years of parallel observations with automatic and classical stations in Romania. 5th seminar on homogenization and data quality in the climatological databases, Budapest, May 29 - June 2, 2006.

Baciu, M., V. Copaciu, T. Breza, S. Cheval, I.V. Pescaru. Preliminary results obtained following the intercomparison of the meteorological parameters provided by automatic and classical stations in Romania. WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2005), Bucharest, Romania, 4-7 May, 2005.

Brandsma, T. and J.P. van der Meulen. Thermometer Screen Intercomparison in De Bilt (the Netherlands),Part II: Description and modeling of mean temperature differences and extremes. Int. J. Climatology, 28, pp. 389-400, 2008.

Dobesch, H. and H. Mohnl. Comparison of time series of sunshine duration measuered by the Campbell-Stokes Recorder and the Haenni Solar System: Instruments and Observing methods. Report No. 49, WMO/TD-No. 462. WMO Technical Conference on Instruments and Methods of Observation (TECO 92), 1992.

Doesken, N.J. The National Weather Service MMTS (Maximum-Minimum Temperature System) - 20 years after. 15th Conference on Applied Climatology, 20—24 June, Savannah, Georgia, no. JP1.26, 2005.

Gerbush, M.R., J. Carlin, D.A. Robinson, C. Speciale, P.J. Croft, Long-term comparison of temperatures observed from multiple sensors at the New Brunswick, NJ NWS Cooperative Weather Station, 20th Conference on Applied Climatology, 5-10 January, 2013, Austin, Texas.

McPherson, R. A., C. A. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, S. J. Stadler, A. D. Melvin, A.J. Sutherland, and H. Shrivastava, 2007: Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. J. Atmos. Oceanic Technol., 24, 301-321.

Sotelino, L.G., N. De Coster, P. Beirinckx, P. Peeters. Intercomparison of cup anemometer and sonic anemometers on site at Uccle/Belgium. WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2012), Brussels, Belgium, 16-18 October 2012.

Suter, S., T. Konzelmann, C. Mühlhäuser, M. Begert, A. Heimo. SWISSMETNET – The new automatic meteorological network of Switzerland: transition from old to new network, data management and first results. Report MeteoSwiss, 2006.

Van der Meulen, J.P. and T. Brandsma. Thermometer screen intercomparison in De Bilt (The Netherlands), Part I: Understanding the weather-dependent temperature differences). Int. J. Climatol., 28, pp. 371-387, doi: 10.1002/joc.1531, 2008.

Warne, J. A preliminary investigation of temperature screen design and their impacts on temperature measurements. Instrument test report number 649. Physics laboratory OEB, 9 June 1998.

6.5 Intercomparisons and corresponding literature references

WMO Field Intercomparison of Thermometer Screens/Shields and Humidity Measuring Instruments: Ghardaïa, Algeria, November 2008 – October 2009
http://library.wmo.int/pmb_ged/wmo-td_1579.pdf

WMO Field Intercomparison of Rainfall Intensity Gauges
http://library.wmo.int/pmb_ged/wmo-td_1504.pdf

Further potential WMO literature at
http://www.wmo.int/pages/prog/www/IMOP/publications-IOM-series.html

ISTI POST has a list of existing parallel measurements and comprehensive literature at
http://ourproject.org/wiki/Projects/parallel

7. AERONET

7.1 Brief description of the ECV/technique/network

The AERONET (AErosol RObotic NETwork) program is a federation of ground-based remote sensing aerosol networks established by NASA and LOA-PHOTONS (CNRS) and has been greatly expanded by collaborators from national agencies, institutes, universities, individual scientists, and partners (aeronet.nasa.gsfc.gov). The program provides a long-term, continuous and readily accessible public domain database of aerosol optical, microphysical and radiative properties for aerosol research, and characterization & validation of satellite retrievals, and synergism with other databases. AERONET collaboration provides globally distributed observations of spectral Aerosol Optical Depth (AOD), inversion products, and precipitable water in diverse aerosol regimes (Figure 10). AOD data are computed for three data quality levels: Levels 1.0 and 1.5 are provided in near real-time, 12-month or longer delay (due to final calibration and manual inspection) ensures that the highest quality data can be found in Version 2, Level 2.0 data products. Inversions, precipitable water, and other AOD-dependent products are derived from these levels and may implement additional quality checks.

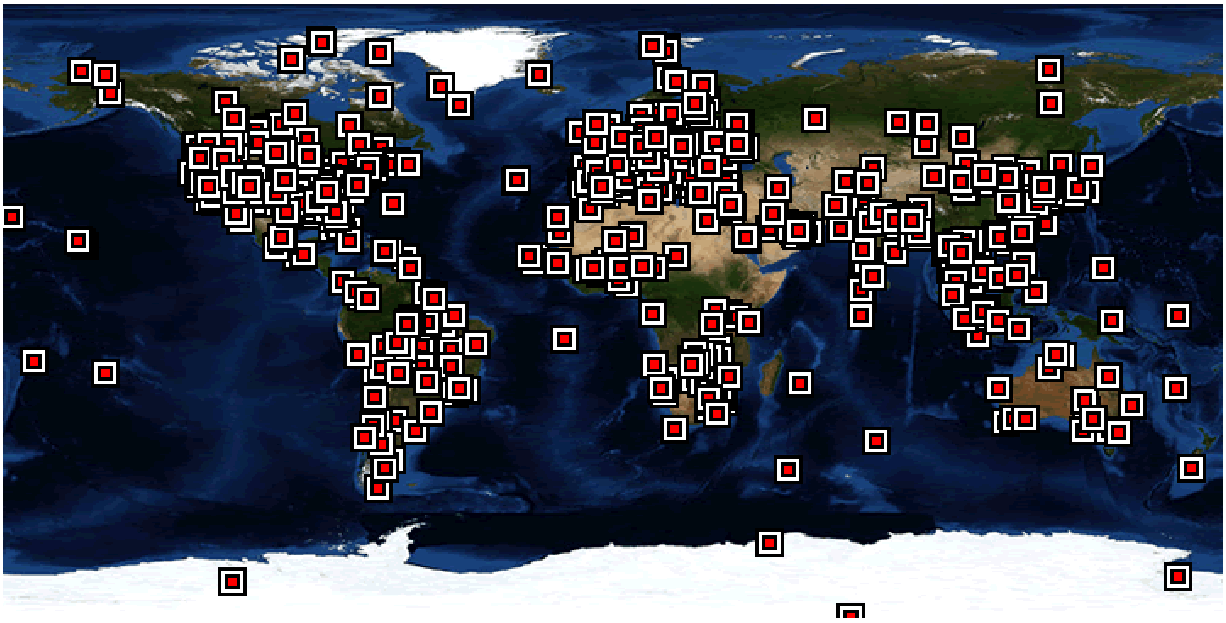

Figure 10. Map showing the AERONET data station available worldwide.

In July 2016, the release of Version 3 near real-time AOD data has been announced. The differences compare to Version 2 are listed below:

- Level 1.0: Minor corrections, selective high AOD restoration applied to all levels
- Level 1.5: Improved cloud clearing, high airmass data included, and automatic data quality assurance applied

- Level 2.0: Manual QA replaced by automatic QA.

For AERONET users the implications of the Version 3 are the following:

- Level 1.0: Minor difference, captures fine mode aerosol plumes at very high AOD. Latency of ~30 minutes or less from the time data are received for processing.
- Level 1.5: Latency of 30 minutes or less from the time data are received for processing. Less cloud contamination, less instrumental anomalies, and more accurate data. Because the cloud clearing is improved and a series of automatic quality assurance algorithms are further removing compromised data, the Level 1.5 may be close to Level 2.0.

AERONET's recommendation is that Level 1.5 should provide excellent data for near real time operational comparisons, such as for satellite and model validation and for model assimilation. Note that all Level 1.5 products are reprocessed multiple times within the first several weeks to utilize the most recent and best ancillary inputs that originate from satellites, radiative transfer models and reanalysis models for both AOD and inversion products. The Level 1.5 products may or may not change during the first six weeks after data collection and/or after a post field calibration is applied prior to Level 2. Thus, we do not recommend using Level 1.5 data for publication.

For Level 2.0, manual QA has been replaced by automatic QA with no latency in Level 2.0 designation. The database is too large for manual QA. By implementing automatic QA algorithms, Level 2 is generated immediately after the post-field calibration is applied. Version 2.0 should be fully supported and processed by December 2017 after which time all data should be locked and archived. Version 3 data are already available in the AERONET website though this part of the website is still under implementation.

More information on the release of AERONET V3 Level 1.0 and Level 1.5 near-real time (NRT) database, announced on June 22nd, 2017 are available at:
http://aeronet.gsfc.nasa.gov/new_web/Documents/AERONET_V3_AOD.pdf

Measurement technique

AERONET stations are equipped with a CIMEL sun photometer which is a multi-channel, automatic sun-and-sky scanning radiometer that measures the direct solar irradiance and sky radiance at the Earth's surface (Figure 11). It provides the aerosol optical depth (AOD) at 340, 380, 440, 500, 675, 870, 1020 and 1640 nm, along with the water vapour column content and the estimation of several optical and microphysical aerosol properties, such as the refractive index and the size distribution.

Figure 11. Picture of the sun photometer measuring since 2004 at Potenza AERONET station (40.6°N, 15.72°E, 760 m a.s.l.).

The direct (collimated) solar radiation is used to calculate the columnar aerosol optical depth (AOD). AOD can be also used to compute columnar water vapor (precipitable water) and estimate the aerosol size using the Angstrom parameter relationship.

The system is fully automatic and powered using solar panels. The measured radiances are automatically sent to the NASA-GSFC (NASA – Goddard Space Flight Center) where they are processed according to the AERONET data analysis. Cloud products are also available for a part of the network. At several stations, sun photometer measurements are collocated with Raman lidar measurements.

There are two other versions of the sun photometer available in the network: the first is equipped with a polarization channel able to provide more information about the particle shape and to improve the inversion products; the second, more relevant for the AOD measurement, is the Sun Sky Lunar Multiband Photometer which is able to perform daytime and night-time photometric measurements using the sun and the moon as light source (Barreto et al., 2016). First night-time data are routinely released thought the AEROENT website.

7.2 AERONET maturity matrix

The maturity assessment for AERONET was performed within GAIA-CLIM in September 2016 http://www.gaia-clim.eu/system/files/document/AERONET_PHOTONS%20with%20text.jpg. It assesses certain quantifiable aspects of typical measurement system maturity across the network for those ECVs and associated measurement systems that are relevant to GAIA-CLIM.

AERONET_PHOTONS						
Metadata	Documentation	Uncertainty characterization	Public access, feedback and update	Usage	Sustainability	Software (optional)
Standards	Formal Description of Measurement Methodology	Traceability	Access	Research	Siting environment	Coding standards
Collection level	Formal Validation Report	Comparability	User feedback mechanism	Public and commercial exploitation	Scientific and expert support	Software documentation
File level	Formal Measurement Series User Guidance	Uncertainty Quantification	Updates to record		Programmatic support	Portability and numerical reproducability
		Routine Quality Management	Version control			Security
			Long term data preservation			
Legend						
1	2	3	4	5	6	Not applicable

Table 8. Maturity matrix assessing the AERONET (AErosol RObotic NETwork). [5]

Typically, a reference quality measurement program would score 5s and 6s against relevant criteria, a baseline capability 3s and 4s and a comprehensive capability 1s and 2s. While most entries score higher (e.g. all subcategories of the sustainability, usage and public access categories are of reference quality), some subcategories within the uncertainty characterization, documentation and metadata of the AERONET maturity matrix score 3-4 which currently still labels the network with baseline status for the purposes of GAIA-CLIM.

7.3 Description of (indicative) measurement uncertainties

AERONET imposes standardization of instruments, calibration, processing and data distribution (e.g. http://aeronet.gsfc.nasa.gov/new_web/PDF/AERONETcriteria_final1.pdf). The influence of various instrumental, calibrational, atmospheric and methodological factors affecting the precision and the accuracy of optical depth determination requires their minimization (see for example Shaw, 1976, Reagan et al., 1986 and Russel et al., 1993). AERONET documentation and publications on uncertainties are reported on the website, such as

[5] *Users of GAIA-CLIM's Maturity Matrices, should be aware that this is a first effort to systematically quantify measurement system performance. Redundant assessments suggest a minimum uncertainty arising from assessor-to-assessor variations in any category of at least 1 score. Although the assessment may be useful for certain applications, at this time and until more broadly tested, it should not constitute a primary or unique decision making tool.*

https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140008327.pdf.

Instrument uncertainty due to electro-optical precision is considered for all practical purposes insignificant for a properly operating instrument. The variability of the atmosphere is characterized by the variability of the triplet optical thicknesses which may at times be cloud contaminated. This uncertainty is computed, can be used as a screening tool, and may be retrieved from the AERONET data base (https://aeronet.gsfc.nasa.gov/valdesaire/val.html). Additionally, the uncertainty due to calibration is tracked with time dependent data and may also be retrieved from the data base.

Typically, the total uncertainty in AOD (aerosol optical depth) from a newly calibrated field instrument under cloud free conditions is <±0.01 for wavelengths>440 nm and <±0.02 for shorter wavelengths (Eck et al., 1999).

As far as we know, the AERONET team is working hard to provide a newer data version which includes a fully traceable uncertainty estimation for the AOD measurement and for the other aerosol and water products released at NASA-GSFC.

7.4 Literature and reports

AERONET NASA-GSFC website official documents can be found at aeronet.gsfc.nasa.gov.

Barreto, Á., Cuevas, E., Granados-Muñoz, M.-J., Alados-Arboledas, L., Romero, P. M., Gröbner, J., Kouremeti, N., Almansa, A. F., Stone, T., Toledano, C., Román, R., Sorokin, M., Holben, B., Canini, M., and Yela, M.: The new sun-sky-lunar Cimel CE318-T multiband photometer – a comprehensive performance evaluation, Atmos. Meas. Tech., 9, 631-654, doi:10.5194/amt-9-631-2016, 2016.

Eck, T.F., Holben, B.N., Reid, J.S., Dubovik, O., Smirnov, A., O'Neill, N.T., Slutsker, I. and Kinne, S., Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. Journal of Geophysical Research 104: doi: 10.1029/1999JD900923. issn: 0148-0227, 1999.

Holben, B. N., Eck, T. F., Slutsker, I., Tanr´e, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Reagan, J.A., L.W. Thomason, B.M. Herman, and J.M. Palmer, Assessment of atmospheric limitations on the determination of the solar spectral constant from ground-based spectroradiometer measurements, IEEE Trans. Geosci. Rem. Sens.,GE-24, 258-265, 1986.

Russel, P.B., J.M. Livingston, E.G. Dutton, R.F. Pueschel, J.A. Reagan, T.E. Defoor, M.A. Box, D. Allen, P. Pilewski, B.M. Herman, S.A. Kinne and D.J. Hofmann, Pinatubo and pre-pihatubo optical depth spectra: Mauna Loa measurements, compariosns, inferred particle size distributions, radiative effects and relationships to lidar data, J. Geophys. Res., vol. 98, no. D12, pp. 22,969-22,985, 1993.

Shaw, G.E., Error analysis of multi-wavelength sun photometry, Pageoph., 114, 1-14, 1976.

7.5 Intercomparisons and corresponding literature references

Cheymol, A., L. Gonzalez Sotelino, K. S. Lam, J. Kim, V. Fioletov, A. M. Siani, and H. De Backer, Intercomparison of Aerosol Optical Depth from Brewer Ozone spectrophotometers and CIMEL sunphotometers measurements , Atmos. Chem. Phys., 9, 733–741, 2009.

Friess, U., H. Klein Baltink, S. Beirle, K. Clémer, F. Hendrick, B. Henzing, H. Irie, G. de Leeuw, A. Li, M.M. Moerman, M. van Roozendael, R. Shaiganfar, T. Wagner, Y. Wang, P. Xie, S. Yilmaz, and P. Zieger, Intercomparison of aerosol extinction profiles retrieved from MAX-DOAS measurements, Atmos. Meas. Tech., 9, 3205–3222, doi:10.5194/amt-9-3205-2016, 2016.

Mazzola, M., R.S. Stone, A. Herber, C. Tomasi, A. Lupi, V. Vitale, C. Lanconelli, C. Toledano, V.E. Cachorro, N.T. O'Neill, M. Shiobara, V. Aaltonen, K. Stebel, T. Zielinski, T. Petelski, J.P. Ortiz de Galisteo, B. Torres, A. Berjon, P. Goloub, Z. Li, L. Blarel, I. Abboud, E. Cuevas, M. Stock, K.-H. Schulz, A. Virkkula, Evaluation of sun photometer capabilities for retrievals of aerosol optical depth at high latitudes: The POLAR-AOD intercomparison campaigns, Atmospheric Environment 52, 4-17, doi:10.1016/j.atmosenv.2011.07.042, 2011.

Nyeki S., Halios C.H., Eleftheriadis K., Amiridis V., Gröbner J., Wehrli C., Ground-Based Aerosol Optical Depth Inter-Comparison Campaigns at EUSAAR Sites in Athens, Greece. In: Helmis C., Nastos P. (eds) Advances in Meteorology, Climatology and Atmospheric Physics. Springer Atmospheric Sciences. Springer, Berlin, Heidelberg, 2013.

Piters, A. J. M., Boersma, K. F., et al., The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results, Atmos. Meas. Tech., 5, 457–485, doi:10.5194/amt-5-457-2012, 2012.

Schafer, J. S., Eck, T. F., Holben, B. N., Thornhill, K. L., Anderson, B. E., Sinyuk, A., Giles, D. M., Winstead, E. L., Ziemba, L. D., Beyersdorf, A. J., Kenny, P. R., Smirnov, A., Slutsker, I., Intercomparison of aerosol single-scattering albedo derived from AERONET surface radiometers and LARGE in situ aircraft profiles during the 2011 DRAGON-MD and DISCOVER-AQ experiments, J. Geophys. Res. Atmos. 119, 12, http://dx.doi.org/10.1002/2013JD021166, 2014.

Shinozuka Y., J. Redemann, J. M. Livingston, P. B. Russell, A. D. Clarke, S. G. Howell, S. Freitag, N. T. O'Neill, E. A. Reid, R. Johnson, S. Ramachandran, C. S. McNaughton, V. N. Kapustin, V. Brekhovskikh, B. N. Holben, and L. J. B. McArthur, Airborne observation of aerosol optical depth during ARCTAS: vertical profiles, inter-comparison and fine-mode fraction, Atmos. Chem. Phys., 11, 3673–3688, doi:10.5194/acp-11-3673-2011, 2011

8. Summary

In this deliverable, we have introduced and discussed five baseline network/ECV combinations which were selected because they are, firstly, close to reference quality and, secondly, of interest for potential future inclusion into the GAIA-CLIM Virtual Observatory. For the five selected network/ECV combinations, we have provided some general information regarding the network, measurement technique(s) and ECV(s), a maturity matrix assessment, a description and estimate of the measurement uncertainties, information on measurement intercomparisons and any relevant background literature.

In the final report (deliverable D2.7), we will extend the assessment of the uncertainties for the selected examples of baseline capability and evaluate, where possible, their uncertainty estimates in the context of a corresponding reference network. One apparent example is to investigate how the baseline network GUAN relates to the reference network GRUAN for radiosonde temperature and water vapour profiles. In the final report, we will also extend the preliminary list of the five examples of baseline networks further, and provide best estimates for uncertainties of selected comprehensive networks. Further examples of based on their MSMM assessments performed within WP1.