GAIA-CLIM deliverable D2.5

Gap Analysis for Integrated Atmospheric ECV CLImate Monitoring

WP2: Measurement uncertainty quantification. Review of the methodologies and tools used for uncertainty quantification

D2.5: "Progress report on development of best practices under Task 2.3".

A Horizon 2020 project;

Grant agreement: 640276

Date: 28th February 2017

Lead Beneficiary: NPL

Nature: Other

Dissemination level: PU

Work-package	**WP2**
Deliverable	D2.5
Nature	R
Dissemination	PU
Lead Beneficiary	NPL
Date	28/02/2017
Status	Final
Authors	Paul Green
Editors	Tom Gardiner, Karin Kreher
Reviewers	Richard Davy, Corinne Voces, Peter Thorne
Contacts	
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

Contents

Scope of Document

This document describes the progress made within WP2 task 2.3 in creating and disseminating metrological best practise guidance to the wider WP2 activities in the development of reference quality measurement capabilities and uncertainty quantification for the target geophysical quantities and instrumentation.

Introduction

Task 2.3 stipulates the continuous review of the methodology and tools devised in tasks 2.1 and 2.2 to ensure that the uncertainty traceability and quantification techniques rigorously follow best practice and that the procedures used for the different instruments and measurement techniques are demonstrably comparable.

To facilitate these aims, NPL has been developing a best practise guide, alongside constructing a template for the reporting of instrument uncertainties, with worked examples; to allow a broad range of different instrumental techniques to be directly compared within the same structure.

The best practise guide & reporting template have been packaged as a short and accessible 'Guide to Uncertainty in Measurement & its Nomenclature' document, now on its second substantive version. The dissemination of the document within the WP2 partners has been supplemented by a number of rounds of teleconferences at the task 2.1.x level. In these calls, NPL has endeavoured to describe the document content; in terms of both the technical detail together with the overarching philosophy of the best practise.

As well as the best practise guide a Go/NoGo template has been developed in a simple checklist format as a mechanism to demonstrably assess a geophysical product against the GAIA CLIM-developed metrics for metrological robustness and it's published supporting evidence. The checklist acts as a simple summary indicating whether a product has the evidential basis to be deemed reference quality and acceptable to be stated as such in the GAIA-CLIM Virtual Observatory.

At the recent General Assembly, the Guide and Go/NoGo template were presented to the wider GAIA-CLIM consortium membership within the main meeting, but also more expansively discussed in a full day supplementary workshop, attended by at least one participant of each of the task 2.1.x groups.

The presentation from the supplementary session is appended to this report.

In the remainder of this document, the guidance document and Go/NoGo checklist are briefly described, with a final future activities section describing the activities foreseen before M30 when the guidance document is submitted as deliverable D2.6.

The Guide to Uncertainty in Measurement & its Nomenclature document

The Guide to Uncertainty in Measurement & its Nomenclature has been developed as a short and accessible guide to the metrological robust treatment of uncertainties, written to be both uncomplicated and pragmatic, for use by the EO user community.

The first six sections describe the 'Guide to the Expression of Uncertainty in Measurement' (GUM), the concepts of traceability and some of the basic methods & terminology used in metrology.

Section 7 describes traceability chains, the philosophy of thinking about the function of the measurement instrument & processing of the resultant data in terms of the physical, processing and metrological models. The three models are tools to encourage the user to consider all the contributing elements to the measurement process, so fully populate all the instrumentation & process elements. These elements should include all the active processes, but also all the potentially invisible sources of uncertainty that arise from assumptions in the model. (For example, a process may not include a detector non-linearity correction, so assumes linearity; but there is an uncertainty associated with this assumption that should be included in the assessment.)

Section 8 is the practical guide to producing traceability chains within GAIA CLIM, including the template element table for each of the steps in the traceability chains. The element table identifies the information needed for each element such that the uncertainty contribution can be treated robustly, including correlations with other effects and the timescales over which correlations exist. This section ends with the traceability confidence assessment, where the impact of elements with lower levels of traceability on the overall uncertainty is considered.

The final sections of the guidance document include suggested further reading and a glossary of vocabulary.

The Guide is aligned with similar advice issued by key sister projects. The traceability chain physical, processing and metrological models take work developed in the QA4ECV project (which is in turn based on the previous QA4EO project) tailored for GAIA CLIM application. The element tables in section 8 have significant synergies with our sister project FIDUCEO, again tailored for the GAIA CLIM instrumentation set.

The Guide to Uncertainty in Measurement & its Nomenclature document is appended to this report.

Worked Example

A worked example is currently being developed to act as a practical example used within WP2, to demonstrate the application of the guidance. The GRUAN radiosonde temperature product has been chosen for this initial example. Although the GRUAN radiosonde temperature product is not one of the product/technique combinations being studied in GAIA CLIM, it has been chosen because:

- relevant technique technical documentation and uncertainty analysis exist in the literature,

- GRUAN radiosonde products are a widely used reference-grade dataset, which will be used in the traceability closure studies between GAIA CLIM WP2-studied and satellite products.

The Go/NoGo checklist

The Go/NoGo checklist has a short table structure as a way of summarising the existence and reference to the product key documentation needed to verify that it is of reference-grade quality.

The top section of the checklist identifies the precise product specifics to which the checklist refers. Essentially, to make a clear delineation as to the subset of data to which the assessment can be applied. Almost inevitably, the product assessed is only part of data available from a wider network, or larger family of instrumentation, that may include variations in instrumentation or practise not considered in the metrological assessment. The product descriptor therefore minimises any confusion in this regard.

The checklist asks for the following information.

- Has the product traceability uncertainty (PTU) chain/diagram been constructed? A graphical representation of the chain can be appended to the checklist.
- Have the uncertainties associated with the links in the PTU chain/diagram been assessed and combined in accordance with the Guide to Uncertainty in Measurement & its Nomenclature?
- Has the uncertainty assessment been published or made available publically to the user community?
- Has a technical document describing the measurement procedure been produced and published or made available publically to the user community?
- Have the sites using the instruments covered and the time period over which the assessment applies been clearly described?
- Has the traceability confidence assessment, described in §8.1.2 of the Guide to Uncertainty in Measurement & its Nomenclature been completed?

The Go/NoGo checklist is appended to this report.

Future activities

The Guide to Uncertainty in Measurement & its Nomenclature is the best practise document that will form the basis of the D2.6 deliverable, due in M30. The document will continue to be evolved over the remainder of the GAIA-CLIM project, updated with lessons learnt from the practical application of its principles within WP2.

GAIA-CLIM does not sit alone developing best metrological practise within Earth Observation ECVs. NPL will continue to interact with the FP7 QA4ECV, H2020 FIDUCEO and other relevant projects to draw on best practise developed in these sister projects; ensuring consistency between these activities and the wider advice disseminated by these projects. An aligned methodology and reporting of uncertainties, particularly between FIDUCEO & GAIA-CLIM will best facilitate the closure of the traceability loop between ground-based reference and satellite products.

The guide will contain worked examples as an appendix to demonstrate the application of the guidance, so optimise the ease of community adoption outside the GAIA-CLIM consortium into the future. One worked example will be for a GRUAN radiosonde product. Another will be chosen from the task 2.1.x activities.

Guide to Uncertainty in Measurement & its Nomenclature

Version 2.0

GAIA-CLIM
Gap Analysis for Integrated
Atmospheric ECV Climate Monitoring
Mar 2015 - Feb 2018

A Horizon 2020 project; Grant agreement: 640276

Date: 18 January 2017

Dissemination level: PU

Work Package 2; Edited by Paul Green & Tom Gardiner (NPL)

Table of Contents

Version history

Version	Principal updates	Owner	Date
1.0	First release	NPL	18.09.2015
1.1	Revision of Annex A & addition of §7	NPL	05.11.2015
2.0draft	Addition of §8 and minor changes in §7	NPL	18.11.2016
2.0	Addition of §8.1.2	NPL	18.01.2017

The content of sections 1-6 of this guidance note are largely taken from the ‘Intermediate Uncertainty Analysis for Earth Observation: Instrument Calibration Module’ course notes [1]

1 The ISO and BIPM Guide to the Expression of Uncertainty in Measurement (GUM)

The *Guide to the Expression of Uncertainty in Measurement*, known as 'the GUM', provides guidance on how to determine, combine and express uncertainty [2]. It was developed by the JCGM (Joint Committee for Guides in Metrology), a joint committee of all the relevant standards organisations (e.g. ISO) and the BIPM (*Bureau International des Poids et Mesures*). This heritage gives the GUM authority and recognition. The JCGM continues to develop the GUM and has recently produced a number of supplements. All of these, as well as the 'VIM' (International Vocabulary of Metrology, [3]) are freely downloadable from the BIPM website[1].

2 Measurement Traceability and SI

Measurement traceability is defined by the Committee for Earth Observation Satellites (CEOS) as the

> *Property of a measurement result relating the result to a stated metrological reference (free definition and not necessarily SI) through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty.*

Measurement traceability is an unbroken chain (i.e. it is calibrated against X, which was calibrated against Y, which was calibrated against Z, all the way back to SI, or, perhaps, a recognised authoritative reference) Additionally, effective quality assurance requires the documentary evidence that each step is done in a reliable way (ideally audited, at least thoroughly peer-reviewed). Validation of datasets, a prime concern of the GAIA-CLIM project, requires the combination of measurement traceability, quality assurance & process traceability of the reference & target measurement systems; providing an unbroken chain between the measurement systems through a common measurand, be that the target geophysical parameter or a closely related quantity.

Measurement traceability should, ideally, be to the International System of Units, known as the SI from its French name, *le Système international d'unités*. The SI units provide a coherent system of units of measurement built around seven base units and coherent derived units. A coherent system of units means that a quantity's value does not depend on how it was measured. The SI is an evolving system, with the responsibility for ensuring long term consistency with the General Conference on Weights and Measures (CGPM), run through the International Bureau of Weights and Measures, the BIPM, and maintained nationally through the National Metrology Institutes (NMIs). The CIPM Mutual Recognition Arrangement (CIPM MRA) signed in 1999 between the NMIs ensures that measurements made traceably to any NMI within the CIPM MRA are recognised by other NMIs. This is enforced by both formal international comparisons and a process of auditing and peer-reviewing statements of calibration capability. For the user, this means that traceability to the SI can be achieved through any NMI within the CIPM MRA.

[1] http://www.bipm.org/en/publications/guides/

3 Errors, uncertainties and corrections

The terms ‘error’ and ‘uncertainty’ are not synonyms, although they are often confused. To understand the distinction, consider the result of a measurement – the measured value. The value will differ from the true value for several reasons, some of which we may know about. In these cases, we may be able to identify and apply a **correction**. A correction is applied to a measured value to account for known differences, for example the measured value may be multiplied by a gain determined during the instrument’s calibration, or a measured optical signal may have a dark reading subtracted. This correction will never be perfectly known and there will also be other effects that cannot be corrected, so after correction there will always be a residual, unknown **error** – an unknown difference between the measured value and the (unknown) true value.

The specific error in the result of a particular measurement cannot be known, but we describe it as a draw from a probability distribution function. The **uncertainty** associated with the measured value is a measure of that probability distribution function; in particular, the **standard uncertainty** is the standard deviation of the probability distribution, and the equivalent of this for other distributions. There are generally several ‘sources of uncertainty’ that jointly contribute to the uncertainty associated with the measured value. These will include uncertainties associated with the way the measurement is set up, the values indicated by instruments, and residual uncertainties associated with corrections applied. The final (unknown) error on the measured value is drawn from the overall probability distribution described by the uncertainty associated with the measured value. This is built up from the probability distributions associated with all the different sources of uncertainty.

The use of the words ‘error’ and ‘uncertainty’ described here is consistent with paragraph 2.2.4 of the GUM, and described graphically in Figure 1.

Conversely it is worth considering that which is *not* a measurement uncertainty.

- Mistakes made by operators are not measurement uncertainties. They should generally be avoided, and identified thorough checking of the results obtained.
- Tolerances are not uncertainties. They are acceptance limits which are chosen for a process or a product.
- Specifications are not uncertainties. A specification tells you what you can expect from a product or what a user requires from a product. It may be very wide-ranging, including ‘non-technical’ qualities of the item, such as its appearance.

4 The law of propagation of uncertainties

The aim of uncertainty analysis is to estimate the uncertainty associated with the measured value, which may be the result of a process involving several different parameters being controlled and set or measured, and a calculation. To obtain the final uncertainty, uncertainties due to *each and every* element in the process that affect the final result must be combined – i.e. they must be propagated through this process. Ref [1] contains an extended worked example for an airborne EO instrument.

The GUM gives the Law of Propagation of Uncertainty as,

$$u_c^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j), \quad (2)$$

which applies for a measurement model of the form

$$Y = f\left(X_1, X_2, X_3, \ldots, X_i, \ldots\right) \tag{3}$$

where an estimate x_i of quantity X_i has an associated uncertainty $u\left(x_i\right)$. The quantity $u_c^2\left(y\right)$ is the squared standard uncertainty (standard deviation of the probability distribution) associated with the measured value y which comes from a combination of the uncertainties associated with all the different effects, x_i. The square of the standard uncertainty is also known as the **variance**. The second term on the right hand side of eqn. 2 sums the covariance terms. The covariance is a measure of the uncertainty common to the two quantities in the measurement model.
It can help to write the Law of Propagation of uncertainties in terms of **sensitivity coefficients** as

$$u_c^2\left(y\right) = \sum_{i=1}^{n} c_i^2 u^2\left(x_i\right) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c_i c_j u\left(x_i, x_j\right), \tag{4}$$

where the sensitivity coefficient $c_i = \partial f / \partial x_i$. The sensitivity coefficient is a 'translation' from one variable to another. It answers the question: "how sensitive is y to an uncertainty associated with x_i?"

The law of propagation of uncertainties is written in this slightly complex notation of two parts to separate two terms:

- The first term is the sum of the squares of the standard uncertainties $u\left(x_i\right)$ (the sum of the variances) associated with each individual effect multiplied by the relevant **sensitivity coefficient** (the partial derivative). This first term is what is meant by the description '**adding in quadrature**'.
- The second term deals with the **covariance** of correlated quantities. The covariance is a measure of how much the two quantities vary together. If the covariance term is zero, this term becomes zero by definition.

Note that the covariance term covers all pairs of different quantities, e.g. $\left(x_1, x_2\right), \left(x_1, x_3\right), \left(x_2, x_3\right), \ldots$
Since the covariance $u\left(x_1, x_2\right) = u\left(x_2, x_1\right)$, the summation is only over the combinations where $i < j$ (i.e. only half the cases). The 2 in front of this term accounts for the opposite cases.

4.1 Coverage factor k

Having scaled the components of uncertainty consistently, to find the combined standard uncertainty, we may then want to re-scale the result. The combined standard uncertainty may be thought of as equivalent to 'one standard deviation', but we may wish to have an overall uncertainty stated at another level of confidence, e.g. 95 percent. This re-scaling can be done using a coverage factor, k. Multiplying the combined standard uncertainty, u_c , by a coverage factor gives a result which is called the expanded uncertainty, usually shown by the symbol U,

$$U = k.u_C \tag{1}$$

A particular value of coverage factor gives a particular confidence level for the expanded uncertainty. Most commonly, we scale the overall uncertainty by using the coverage factor k = 2, to give a level of confidence of approximately 95 percent. (k = 2 is correct if the combined standard uncertainty is normally distributed. This is usually a fair assumption, but the reasoning behind this is explained elsewhere, in [2].) Some other coverage factors (for a normal distribution) are:

- k = 1 for a confidence level of approximately 68 percent
- k = 2.58 for a confidence level of 99 percent
- k = 3 for a confidence level of 99.7 percent

Other, less common, shapes of distribution have different coverage factors. Conversely, wherever an expanded uncertainty is quoted with a given coverage factor, you can find the standard uncertainty by the reverse process, i.e. by dividing by the appropriate coverage factor.

5 Classifications

Random and Systematic Effects

Correlation will be introduced whenever there is something in common between two measured values that will be combined (i.e. two values that will be averaged, or two quantities used in a measurement equation, or values at different wavelengths that will be combined through interpolation or integration). The simplest way to describe this is in terms of random and systematic effects.

Random effects are those that are not common to the multiple measurements being combined. A common example is noise: two measured values may both suffer from noise, but the effect of noise will be different for each of the two measured values (for example, if noise has increased one measured value, this provides no information about whether any other measured value is increased or decreased by that noise, nor by what extent).

Systematic effects are those that are common to all measured values. If one measured value has been increased as a result of a systematic effect, then we can make a reliable prediction regarding whether any other measured value will be increased, and by how much. For example each time the distance is set for an irradiance measurement using a particular lamp, there will be a (normally small) error in that distance. This will equally affect all measurements of that lamp until the next alignment. If multiple measured values are averaged without realignment, or measured values at different wavelengths are combined in an integral, then the distance error will be common to all those measured values. This is a systematic effect.

When validating EO datasets correlated systematic effects common to both the reference & target instrument systems may exist. For instance, SZA, surface albedo and background atmospheric absorption & scattering processes may be common uncertainty contributors to both measurement systems.

Some effects, such as noise, are always random; other effects can be either random or systematic depending on the measurement process. For example, if three measured values of a lamp are combined in an average and the lamp is realigned between each measurement, then alignment/distance is a random effect. If the lamp is not realigned between measurements, then alignment/distance is a systematic effect.

The error in the measured value due to a random effect will change from one measured value to another. In this case the uncertainty associated with the effect may be the same for each measured value (the probability distribution for the effect is the same for each measured value), but each measured value is independent of each other measured value, as influenced by this effect. The unknown random error at each measured value is an independent draw from the probability distribution, meaning that the error due to the random effect is not only different from, but also independent of, the error at any other wavelength. The standard uncertainty associated with random effects is usually (but not always) determined by calculating the standard deviation of repeated measured values.

Such repeat measurement is difficult if not impossible in the atmospheric domain as the measured quantity is almost invariably non-static such that repeat measurement of the measured quantity is not possible. In a few cases pseudo repeat measurements are possible, that is, if measurements can be taken sufficiently close in time and space and also close in sensitivity, so that the contribution of natural variability to the obtained standard deviation becomes negligible. But those cases are not the rule and in general any estimate of the standard deviation will include contributions from spatial, temporal and sensitivity mismatch.

Another important consideration in the atmospheric domain are influence quantities. Influence quantities do not affect the instrument measurand directly, but affects the derived geophysical measurand through departure from the assumptions of the processing model; e.g., cloudiness in the field-of-view of an instrument can influence the accuracy of its measurement.

The error in the measured value due to a systematic effect will be the same from one measured value to another. The uncertainty associated with the effect is the same for each measured value and the error is the same draw from the probability distribution for all measured values. The standard uncertainty associated with systematic effects cannot be determined by repeat measurements, unless the effect is intentionally altered between repeats (e.g. by realigning a source multiple times using a series of different ‘extreme but acceptable’ alignments in an experiment to characterise the impact of source alignment).

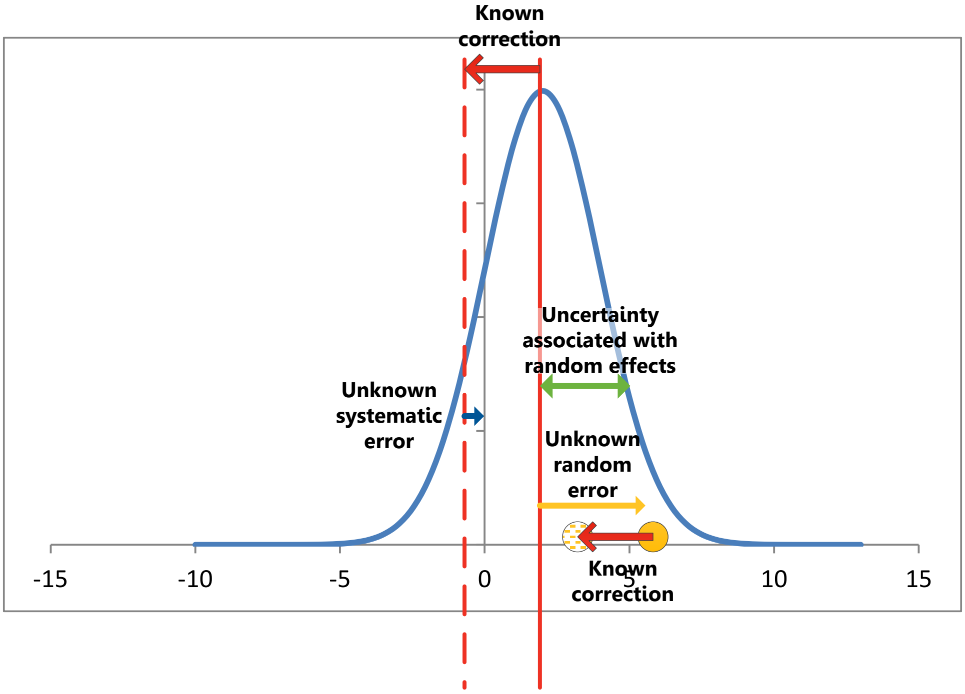


Figure 1: Representing a measurement where there is a known correction, an unknown systematic effect and random effects.

Figure 1 represents a measurement process where there is a known correction, an unknown systematic effect and random effects.

- A measurement is made (obtaining the value represented by the golden circle).
- We know of a correction – a systematic bias due to, e.g. a dark reading – and apply this correction, obtaining the value of the flecked circle.
- There is still an unknown error from the true value of zero. If we make many measurements we obtain the probability distribution function shown in blue. The spread of this, the standard deviation of the normal distribution, is the standard uncertainty associated with random effects – those effects that change from measurement to measurement. Our

measured value is a draw from this probability distribution function. If we take multiple measurements we obtain different draws. The average will tend towards the value at the peak of this distribution.
- When the known correction is applied, the result will be close to the true value, but differ from it by an unknown systematic error common to all the measured values. This comes from its own probability distribution function and all measured values have the same draw from that distribution (not shown in the figure, but this will take the form of a probability distribution centred at the true value with a standard deviation equal to the uncertainty associated with systematic effects).

5.1 Type A and Type B

The terms ‘Type A’ and ‘Type B’ are used with uncertainty analysis. This use comes from the GUM, which defines:

- **Type A evaluation (of uncertainty)** method of evaluation of uncertainty by the statistical analysis of series of observations
- **Type B evaluation (of uncertainty)** method of evaluation of uncertainty by means other than the statistical analysis of series of observations

Type A evaluation uses statistical methods to determine uncertainties. Commonly this means taking repeat measurements and determining the standard deviation of those measurements. This method can only treat uncertainties associated with random effects, for example the uncertainty associated with measurement noise.

Type B evaluation uses 'any other method' to determine the uncertainties. This can include estimates of systematic effects from previous experiments or the scientist's prior knowledge. It can also include random effects determined 'by any other method'. For example we may model room temperature by a random variable in the interval from 19 °C to 21 °C – the temperature range of the air-conditioning settings. Similarly, we may say that a voltmeter with 2 digits after the decimal place has an uncertainty associated with resolution of 0.005 V because we know the rounding range.

It is common to assume that ‘Type A’ evaluation is for random effects and ‘Type B’ evaluation is for systematic effects. This is generally, but not always, the case. For example, a ‘Type A’ method may be used to determine the uncertainty associated with alignment: a lamp may be realigned ten times and the standard deviation of those ten measurements used to determine an uncertainty associated with alignment. In a later experimental set-up, measurements may be taken at multiple wavelengths and these combined in a spectral integral. For that integral, alignment is a systematic effect (the lamp is not realigned from wavelength to wavelength) even though the determination of the associated uncertainty was performed using ‘Type A’ methods. Similarly, the uncertainty associated with a random effect may be estimated from prior knowledge, or a measurement certificate, and thus by a ‘Type B’ method.

Is it worth noting that in the field measurements of atmospheric properties, typically will have a lot of type B uncertainties attached and that a comprehensive uncertainty analysis typically would involve several quantities not quantifiable in a lab setting.

5.2 Absolute and relative uncertainties

The uncertainties given in the law of propagation of uncertainties by the symbol $u(x_i)$ are always standard absolute uncertainties. The term **standard uncertainty** means that it is a single standard deviation of the probability distribution function associated with that quantity. The term **absolute uncertainty** means that it has the same unit as the measurand. In other words, if the signal is in volts, the absolute uncertainty will also be in volts. If the distance is in metres, the absolute uncertainty will also be in metres.

It is common in radiometric calibrations to describe **relative uncertainties**, with units of per cent. The relative uncertainty is the absolute uncertainty divided by the quantity, i.e. $u(x_i)/x_i$.

6 Writing about uncertainties

In casual language we talk about 'averaging a set of measurements' or 'the uncertainty in the measurement is 0.5 %'. In metrology these words are defined carefully to reduce misunderstanding. We cannot 'average a set of measurements' but we can 'average the measured values' obtained from those measurements. The measurement has no uncertainty, there is an uncertainty *associated with* the measured value. For a non-specialist, such definitions can seem pedantic, as with jargon in all fields; but for a specialist, such careful use of words is a source of clarity. The words are defined through the VIM: the international vocabulary of metrology [3].

A *measurement* is made (instruments set up and value recorded) of a *measurand* (a quantity, such as radiance) to obtain a *measured value* (e.g. 0.5 W m^{-2} sr^{-1} nm^{-1}) with an *associated uncertainty* (e.g. 0.5 %). The VIM defines **measurement** as the

> *process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity*

The most important word here is *process*: it defines measurement as the act of measuring. A measurement is not a quantity nor a result. The VIM defines **measurand** as the

> *quantity intended to be measured*

In turn, **quantity** is the

> *property of a phenomenon, body or substance, where the property has a magnitude that can be expressed by a number and a reference.*

Thus quantities are things like length, mass, reflectance, irradiance, instrument gain, etc. When you measure a quantity, that quantity is the measurand of the measurement. The **measurement result** is defined by the VIM as the

> *set of quantity values being attributed to a measurand together with any other available relevant information*

The "other available relevant information" refers to the associated uncertainty, perhaps expressed directly, perhaps as a probability density function, or perhaps implied by the number of digits provided with the result (the latter providing less reliable information). The **quantity value** is a

> *number and reference together expressing magnitude of a quantity*

The reference usually means the unit. The **measured quantity value** (often shortened to measured value) is the quantity value that is the particular measurement result.

A fuller glossary of term is given in Appendix A, see the VIM [5] for the full list of terminology.

7 Framework for the production of metrological robust traceability & process chains

Key to understanding and expressing the robust uncertainty analysis of any atmospheric data product is the ability to clearly display the processing steps taken to produce the dataset. As discussed earlier, to obtain the final uncertainty, uncertainties due to *each and every* element in the process that affect the final result must be combined – i.e. they must be propagated through this process. One method for achieving such a detailed understanding is developing a traceability chain. In metrology, the aim of developing a traceability chain is to demonstrate the series of calibrations which link a measurement to a reference standard. For EO applications, this needs to be developed much further to allow processes to be captured in detail.

Following the procedure of other QA frameworks developed for essential climate variables (ECVs) [6], the total chain is divided into two components that reflects the division between

- Instrument processing chain to L0 instrument raw data – physical model
- Data processing chain from L0 instrument data to final geophysical parameter – processing model.

7.1 Types of Traceability Chains

Regardless of the process being considered (instrumental or data processing), a framework of traceability models is currently being developed within QA4ECV that is being trialled within sister projects, such as FIDUCEO [7]. These are not hard & fast rules that should be blindly followed, but a method conceived to help the user think about all the contributions to the uncertainty budget. As the framework is still being developed, it is hoped that its evolution will be guided via feedback from the user community, including GAIA CLIM. This framework involves considering the traceability in terms of three models.

1. **Physical Model** – This model considers the real-world situation, i.e. what is actually occurring in the real world and the physics driving this.
2. **Processing Model** – This model considers how the raw data collected is processed to provide the end product, through calibration to the final geophysical parameter.
3. **Metrological Model** – This model considers the calibration, or linkage, of a measurement or processed data to a reference.

Separating the types of traceability chain into these three models provides several advantages: the separation essentially provides three angles from which the problem can be approached, it allows for the persons producing the chains to have a clear set of boundaries in which to operate when considering the production of the chains as well as being able to choose the type of model with which they are the most familiar as a starting point. It is noted that there may be significant overlap between the models.

7.1.1 Physical Model

The physical model chains describe the real-world by considering the physics behind each stage of the process which contributes to the measurements taken. This includes all of the physical processes associated with the measurand detection; for a radiometric instrument, this covers the physics of how the EM radiation enters the instrument, how it is modified by the optical system, how it is detected and how it is converted to an electrical signal which makes up the output raw signal.

The aim of the physical model is to be able to describe, reliably, the physical processes which contribute to the generation of the L0 data. Therefore, obtaining a suitable physical model requires an understanding of the detector physics including sources of uncertainty such as noise, the non-linearity of a detector, the Spectral Response Function (SRF) of the detector etc. The model would also include any processing of the signal undertaken by the instrument itself, for example, data compression.

It is unlikely that the physical model chain would incorporate *all* of the possible physical processes occurring in the real-world situation due to the complexity of the real-world. The physical model would essential represent a simplified "best guess" of the real-world. However, in producing the physical model, all contributions should be considered and those processes not included in the model, potentially as they are deemed to have a negligible effect on the data product should at least be documented.

Figure 2 shows an example instrument model for a satellite sensor, showing the main physical processing steps from the incoming radiation to the L0 data.

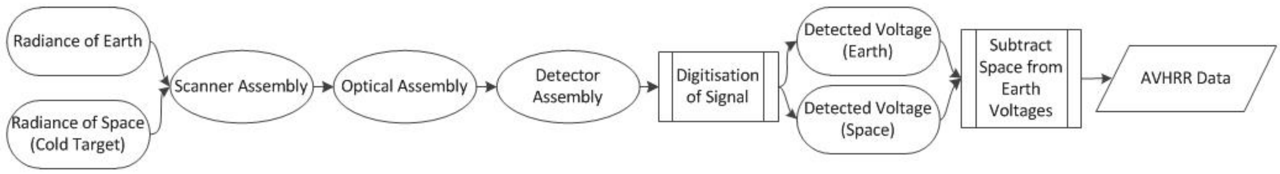


Figure 2: Physical Chain Example – AVHRR Instrument

7.1.2 Processing Model

The processing model chains are intended to describe the input data, processes and output data that contribute to an overall geophysical parameter generation from both Level 0 and ancillary data. This model will include all the processes and assumptions built into the calibration algorithm, as well as any external models or ancillary data used. The processing model will describe a series of calculation steps that the data undergoes to obtain the measurand of interest (i.e. equations and computational models), with inputs derived from the previous step or from pre-set parameters and coefficients, and an output that leads to the next step in the processing chain.

This chain type is conceptually the easiest to understand, particularly within the EO community, where a data producer would intuitively think of a traceability chain as the steps required to produce their product or undertake their process.

One of the key advantages of producing a physical & processing models is the ability to compare these model, so identify differences between the two. This would effectively give the data / product producers details of how their modelled world (represented by the processing chain) differs from the real-world (represented by the physical chain).

An example process chain diagrams of algorithm traceability is shown in Figure 3. Further examples of traceability chains developed within the QA4ECV project can be found at http://www.qa4ecv.eu/ecvs

At a basic level the diagram would contain central boxes representing the processing steps. In addition more detailed information about that step in terms of basic documentation, provenance, assumptions employed and uncertainty analysis should also be provided.

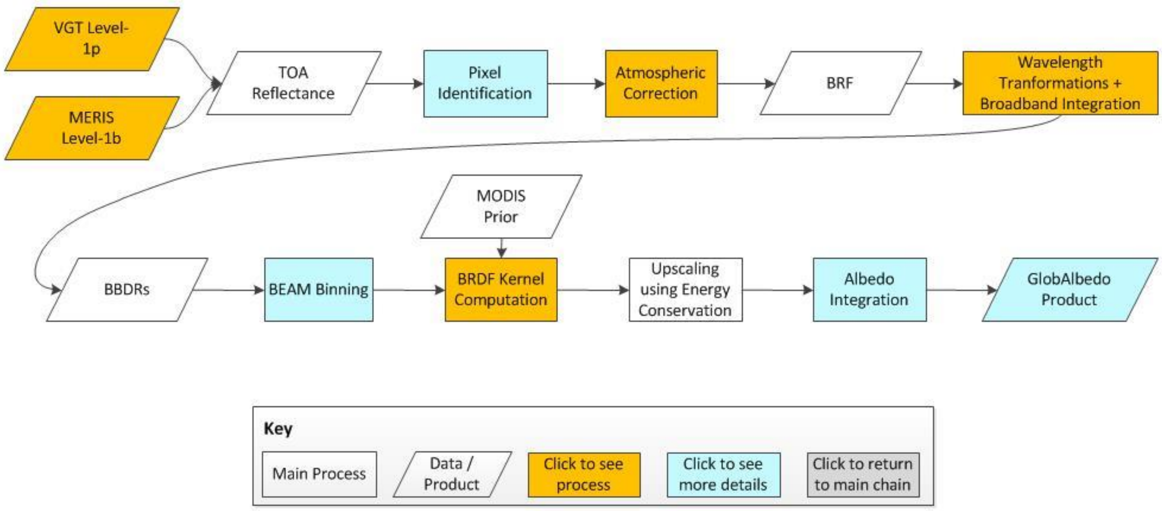


Figure 3. Example traceability diagram for the GlobAlbedo ECV product generation.

7.1.3 Metrological Model

The metrological model chains are intended to describe the set of calibrations, or linkages, of a measurement (or of processed data) to a reference standard. The metrological model describes the origins of the input parameters for the processing model such as the origin of the calibration and characterisation coefficients; be those solely laboratory-based, or occasionally / regularly updated in the field. The aim here is to determine what the fundamental reference for the measurement is. In some cases it will be possible to obtain full metrological traceability - that is, an unbroken chain of calibrations back to the International System of Units (SI). In many cases, however, such a complete chain may not be possible. It is important, however, to show what references do exist. The metrological traceability chain could also be documented as a flow diagram with additional information, containing, for example, references to calibration and characterisation results. Dotted arrows can be used where the link is not strong.

The metrological traceability chain is used to estimate the set of uncertainties (both from random and systematic effects) on the outputs. Note that to be a metrological traceability chain, there is a presumption that all processes have been included and have an estimate of an uncertainty. As part of setting up a metrological model, a review of both the physical and processing model must be made to ensure that all processes are included. As to the uncertainties, where possible, evidence for the magnitude and / or probability distribution of the uncertainties must be provided and documented either through measurements or from Monte-Carlo Analysis (MCA). If no measured uncertainty is available for a process then at least an upper limit to its magnitude must be provided with a rationale for its size. Figure 4 shows an example of a metrological model.

The chain is not used to improve understanding of the processes, nor identify sources of uncertainty; these are both covered by the processing and physical model chains. Therefore, the aim of the chain is to purely demonstrate that linkage to a reference standard is achieved.

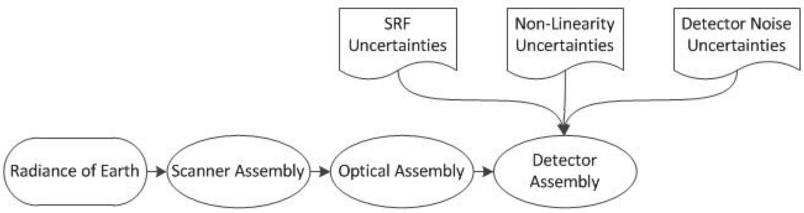


Figure 4. Metrological Chain Example – AVHRR Instrument

7.1.4 Approach to Producing Traceability Chains

- In many cases, the processing model chain is the first type of chain that is produced when describing the traceability of an atmospheric product, as it is the most intuitive type for most users. For many EO satellite applications, the processing model may be the only chain which can realistically be produced in a significant level of detail.

- The physical model involves a more in-depth consideration of the physical processes contributing to the measurement and may be less intuitive for most users.

- The processing and physical model chains are then to be considered iteratively to allow any potential improvements to be made to the processing traceability chain and to ensure that the physical model traceability chain encompasses all relevant elements.

- The metrological model chain should be developed from a combination of the processing and physical models. This chain may have some feedback into the processing and physical model chains; however, this is likely to be limited.

Both the processing and physical model traceability chains will be used for both describing the overall processes associated with an application, as well as being used to describe specific stages. The metrological chain, however, sits alongside the physical & processing chains, and is likely to be used when describing an overall process, rather than the details of individual stages.

The processing, physical and metrological models are then combined to provide an overall model. Alternatively, the overall model can be produced first and split to provide the other models. In either case, it is recognised that producing both the overall model and the set of three other models is not necessary; the production of one or the other is sufficient. **The key aim is ensuring that all relevant data is captured in a systematic manner, whether this be as an overall model, or as three sub-models.** For the technical document deliverable, a single combined chain is required. Figure 5 shows a graphical representation of the sub-model combination. It is noted that the order in which the chains are developed, and the specifics on which each focusses, may vary depending on the application being considered.

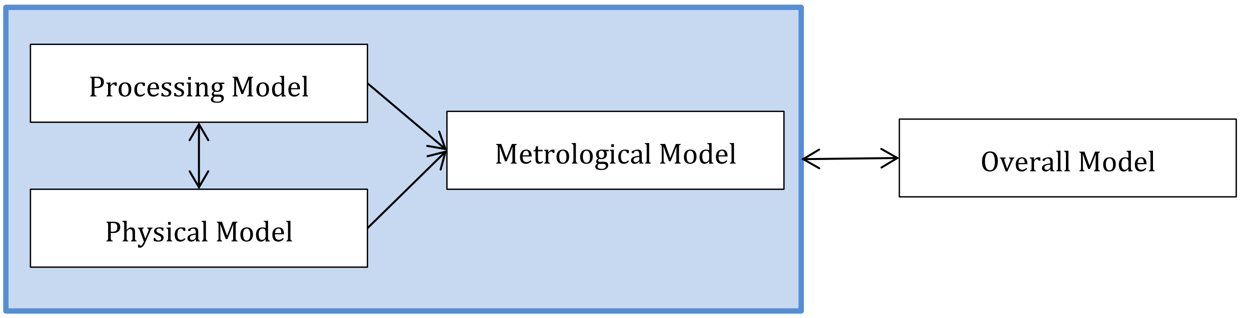


Figure 5: Traceability Chain Production Process

7.2 Representation of the traceability chain

Within the QA4ECV project, a functional prototype of a Traceability and Uncertainty Propagation Tool (TUPT) has been developed[2]. The basic concept of the TUPT is a user-friendly graphical interface that can display (in an electronic interactive format) a visual diagrammatic version of an

[2] http://qa-system-cgi.com.s3-website.eu-central-1.amazonaws.com/#/

algorithm processing step traceability chain of a product[3]. To provide consistency across QA projects, a similar approach is followed in GAIA CLIM.

The chains should be drawn, graphically, as a series of boxes connected to one another via uni- or bi-directional arrows, as seen in Figure 3. Guidance on the types of boxes to be used for each type of model is given at Figure 6. However, it is noted that the underlying information is the important content, so at least in the first iteration, excessive effort should not be spent in formatting the diagrams. The colour scheme is not defined, but should be chosen by the user to best illustrate the commonality in the specific traceability chains. For example, to indicate that further information associated with the box is available, or to group a set of boxes which contribute to a single process.

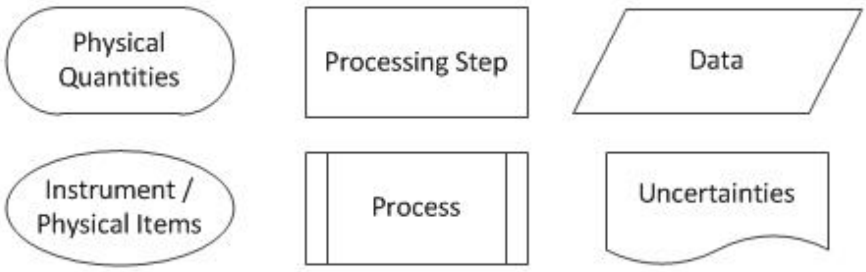


Figure 6. Traceability Chain Shapes and Definitions

7.3 Contributions to consider

The following list is contains some process steps & uncertainty contributions worth considering. This list is by no means exhaustive, the individual users should consider the specifics of their instrument and measurement configuration.

- Incoming radiation
- Field of view / Point spread function.
- Scan mechanism (pointing accuracy, repeatability)
- Input optical system (telescope etc. mirror reflectivity)
- Spectral response function, side-band rejection.
- Stray light (thermal self-emission, RFI etc.)
- Detector assembly (non-linearity, dark signal, read noise, QE)
- Dominant noise sources (thermal, shot, (pre-) amplifier, Generation-Recombination, etc.)
- Digitisation (ADC bits)
- Instrument Housekeeping data (Temperature sensor calibration etc.)
- Calibration source (stability, degradation, emissivity, non-linearity, environmental dependencies)
- Model input datasets & assumptions
- Influence quantities

7.4 Beyond traceability chains

As articulated in the GAIA CLIM Grant Agreement, the vision is to move beyond simple traceability chains (which is effectively understanding the process) towards reference grade products, encapsulated in a 'how to measure' guide and a paper describing the individual produces; for those techniques with sufficient maturity. The ultimate goal is to produce metrologically-

[3] http://ec2-52-39-21-246.us-west-2.compute.amazonaws.com/QA4ECV/TCtool.html

rigorous traceable measurements for the target measurement systems, providing practical coverage factors, applicable in the VO. This may not be possible for all the target measurements within the scope of this project, and will depend on the maturity of the contributing partner technique. However, this ultimate goal should be kept in mind.

The full traceability and uncertainty quantification for each instrument type should mirror the process to define the measurement protocols as described in [9]. The analysis algorithm and error characterisation undertaken should result in a technical document describing the measurement procedure, the existing gaps in the uncertainty assessment, and a publication describing the measurement traceability and its uncertainty.

Useful example publications include:

- Documenting the processing chain and corresponding uncertainties [8]
- General information for reference measurements [9]
- Technical instrument report [10]
- NORS deliverable reports [11], specifically the data user guide & uncertainty budget documents.

8 Producing traceability chains for GAIA CLIM

The breadth of techniques and ECVs covered within GAIA CLIM are extensive, so to try to produce a measurement guide & specific descriptive paper covering all possible permutations far extends the scope of the project in terms of available resources. However, in terms of the VO and the GAIA CLIM aim to describe the process in full as a demonstration of the value of such analysis, rigorous end-to-end treatment of a product uncertainty traceability is essential. Consequently, initially the extent of the GAIA CLIM treated measurement product should be clearly defined. For each product a single traceability chain should be developed which captures all the elements of the system including the physical, processing and metrological aspects. Each participant should therefore:

- Identify the exact measurement product to be quantified within GAIA CLIM
    - The specific technique,
    - The specific measurand,
    - The form of the measurand, i.e. profile/total column.
- Identify the specific dataset, which will go into the VO.

With a narrowed down scope, it should be possible to:

- Identify the specific elements that make up the product chain for this combination of parameters,
- Identify the inputs, the process, the uncertainties and sensitivities of the element to these parameters.
- Characterise the form of the uncertainty, is it random, quasi-systematic or systematic?
    - Independent random effects (noise)
    - Structured random (regular calibration cycles)
    - Systematic effects (long term correlation / fixed parameter)
- Combine the individual elements and associated uncertainty information to create the overall product chain.

It should be reiterated, that although the approach of considering the physical, processing and metrological models may be helpful in ensuring all parts of the chain/tree have been considered, a single chain should be specified for the specific measurand & technique within GAIA CLIM.

8.1 Practical guidance for GAIA CLIM traceability chains

In characterising the uncertainty, reference to previous work/documentation should be made where relevant, but this should not detract from the independent of the GAIA CLIM measurement document. This document needs to be stand alone, so understood if read in isolation from the referenced material.

The traceability chains produced to date should form the basis of this, and require limited additional effort to tailor to the specific case. One concern that should be addressed in the analysis is any differences in site-to-site or user-to-user procedure & observing practice.

- Identify any site-to-site or user-to-user variation in procedure & observing practice from nominally identical instruments so make an assessment of comparability through usage.

The overall measurement equation/chain/tree should consider all contribution factors that feature in the full end-to-end process. However, analysis will likely to be sub-divided in sub-modules as expedient by the specific chain. Each element should have a summary table of knowledge & traceability including an estimate of contribution magnitude. This assessment may be via:

- a formal analytical treatment
- a sensitivity study
- an educated guess

Depending on the level of sophistication, the key is to provide a reasonable estimate with the available information. Once the summary table has been completed for the full chain, it should become clear where further work should be focused to most effectively improve the overall level of knowledge of the process uncertainties.

Table 1. Contribution summary table

Information / data	Type / value / equation	Notes / description
Name of effect		
Contribution identifier		
Measurement equation parameter(s) subject to effect		
Contribution subject to effect		
Element correlation form		
Time correlation form		
Units element correlation		
Units time correlation		
Scales element correlation		

Scales time correlation		
Uncertainty PDF shape		
Uncertainty		
Uncertainty units		
Sensitivity coefficient		
Correlation(s) between affected parameters		
Cross-contribution correlation(s)		
Element/step common for all sites/users?		
Traceable to …		
Validation		

Table 1 shows the summary table to be completed for each process contribution. The notes below add some explanation to the entries.

Name of effect – the name of the contribution

Contribution identifier - unique identifier to allow reference in the traceability chains. Depends on the chain & submodule structure, but A1… Ax, B1 … Bx etc. may be appropriate.

Measurement equation parameter(s) subject to effect – The part of the measurement equation influenced by this contribution.

Contribution subject to effect – The top level measurement contribution affected by this contribution. Either the product, or chain sub-module contribution effected by the contribution.

Temporal correlation form – the form of any correlation this contribution has in time.

Units time correlation – the units of any temporal correlation.

Scales time correlation – factors that scale with the temporal correlation.

Element spatial/spectral correlation form - the form of any correlation this contribution has other than in time, be that spatial, spectral, or any other.

Units spatial/spectral correlation - the units of any spatial/spectral correlation.

Scales spatial/spectral correlation – factors that scale with any spatial/spectral correlation.

Uncertainty PDF shape – the probability distribution shape of the contribution, Gaussian, Rectangular, U-shaped, log-normal or other.

Uncertainty – the uncertainty value

Uncertainty units – units of the uncertainty

Sensitivity coefficient – coefficient multiplied by the uncertainty when applied to the measurement equation.

Correlation(s) between affected parameters – Any correlation between the parameters effected by this specific contribution.

Cross-contribution correlation(s) – identify any cross-correlations within the measurement equation.

Element/step common for all sites/users – Is there any site-to-site/user-to-user variation in the application of this contribution?

Traceable to – describe any traceability back towards a primary/community reference.

Validation – Any validation activities that have been performed for this element?

The summary table, explanatory notes and referenced material in the traceability chain should occupy <= 1 page for each element entry.

Once the summary tables have been completed for the full end-to-end the uncertainties can be combined, allowing assessment of the combined uncertainty, relative importance of the contributors and correlation scales both temporally & spatially. The unified form of this technical document should then allow easy comparison of techniques and methods.

As described in [9] the establishment of reference level observations consists of definition, execution and evaluation phases. This third phase, the systematic evaluation of the performance of those measurement technologies is partially demonstrated by the metrological evaluation activity here within GAIA CLIM.

8.1.1 Temporal and spatial scales in uncertainty assessment

One elucidating aspect of the uncertainty combination would be to consider on a range of temporal and spatial scales, aligned with different user applications, mirroring the random/systematic levels used to classify the uncertainty contribution form. Considered at the level of:

- Instantaneous measurement (smallest unit of reported data) – potentially dominated by random instrumental effects.
- At the calibration cycle/mid-scale temporal averaging scale – where quasi-systematic instrumental effects are treated as random variables.
- At the longer term temporal or spatial averaged scale for a single site/instrument typified by instrument systematic effects
- At network level, incorporating multiple sites/instruments typified by individual site-specific data treated as random variables.

At these different aggregation scales, different uncertainty contributors will dominate with effects on the magnitude of the overall uncertainty and its probability distribution function form. With the information available from the summary tables, this exercise should not be too onerous, but potentially highlight considerations for user applications other than those primarily of the largely instrumentation-orientated teams working with GAIA CLIM.

8.1.2 Product traceability uncertainty summary

A summary table should follow the individual element assessments, in the form given below. The Product traceability uncertainty summary is a summary of the information provided above for this specific product. The purpose of this table is to summarise the assessment and demonstrate at a glance that the dominant contributions to the uncertainty chain have been robustly assessed with adequate traceability.

Element identifier/ name	Uncertainty contribution magnitude	Traceability level (L/M/H)	random, structured random, quasi-systematic or systematic?	Correlated to? (Use element identifier)
1				
2				
3				
4				
….				
TOTAL				

<u>Table category descriptions.</u>

Element identifier/name – The name and identifier should correspond to the relevant contributing element in the product traceability uncertainty chain.

Uncertainty contribution magnitude – the uncertainty estimate for this component, ideally in the product units (although relative uncertainties are acceptable where necessary.)

Traceability level: A description of the traceability associated with this element, following the example set out below.

Traceability Level	Descriptor	Multiplier
High	SI traceable or globally recognised community standard	1
Medium	Developmental community standard or peer-reviewed uncertainty assessment	3
Low	Approximate estimation	10

Although a high level of traceability is desired, this will probably not be the case for all elements. Where that element only makes a small contribution to combined uncertainty, then lower traceability level would be acceptable. The multiplier values provide one possible mechanism to assess this.

Multiplier value assessment: consider the effect on combined uncertainty of applying multiplier to each particular element. If the combined uncertainty is not significantly increased then the traceability level is adequate for that element. If the combined uncertainty does increase significantly, then further work may be required to improve the traceability level.

Note that the reported uncertainties should not have the multipliers included.

Random, structured random, quasi-systematic or systematic? - A descriptor of the form of the uncertainty.

Correlated to? (Use element identifier) – a descriptor as to whether the element is an independent variable, or has correlations to other elements within the product traceability uncertainty chain.

8.2 Resulting Document

The output from this work will be a measurement product technical document which should be stand-alone i.e. intelligible in isolation. Reference to external sources (preferably peer-reviewed) and documentation from previous studies is clearly expected and welcomed, but with sufficient explanatory content in the GAIA CLIM document not to necessitate the reading of all these reference documents to gain a clear understanding of the GAIA CLIM product and associated uncertainties entered into the VO.

9 Further reading

Any further study on uncertainty analysis must start with the GUM itself [2]. The GUM is downloadable from http://www.bipm.org/en/publications/guides/gum.html and this website also contains different supplements to the GUM and an introduction to the GUM.

One JCGM supplement that may be of particular interest within GAIA CLIM is 'Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method' JCGM 101:2008 [4]

NPL offers several good practice guides on measurement and uncertainty analysis, with [5] providing a good introduction. NPL also offers a growing range of training courses, e.g. [1] – both face-to-face and e-learning. See:

http://www.npl.co.uk/publications/good-practice-online-modules/.
http://www.npl.co.uk/learning-zone/training/.

The United Kingdom Accreditation Service (UKAS) Publication M 3003, 'The Expression of Uncertainty and Confidence in Measurement', http://www.ukas.com/library/Technical-Information/Pubs-Technical-Articles/Pubs-List/M3003_Ed3_final.pdf & Publication EA-4/02 of the European co-operation for Accreditation (EA), 'Expression of the Uncertainty in Measurement and Calibration'. http://www.european-accreditation.org/publication/ea-4-16-g-rev00-december-2003 may be of interest.

The best introductory textbook to the concepts of the GUM is arguably *"An introduction to uncertainty in measurement" by Les Kirkup and Bob Frenkel*. It is written in a very straightforward way and provides a good overview of the statistical concepts behind the GUM while remaining pragmatic and practical.

A slightly more advanced and detailed, but still very readable book is *"Data reduction and error analysis for the physical sciences" by P.R. Bevington and D.K. Robinson*. This book discusses the statistical basis of uncertainty analysis, and also describes Monte Carlo techniques and least square fitting.

Annex A – Terminology Glossary

In the 'glossary' below, a few important words are explained, taken from [5]. Precise or rigorous definitions are not given here. They can be found elsewhere, for example in the *International Vocabulary of Basic and General Terms in Metrology*. A useful and correct set of definitions can also be found in UKAS publication M 3003 *The Expression of Uncertainty and Confidence in Measurement* (See Further Reading in Section 16).

accuracy - closeness of the agreement between a measurement result and true value of that measurand. (Accuracy is a qualitative concept only and is not given a numerical quantity value. It is often misused as uncertainty or precision.)

bias (of a measurement) – estimate of a systematic measurement error

bias (of a measuring instrument) - systematic error of the indication of a measuring instrument

calibration - operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication. In other words, the comparison of an instrument against a reference or standard, to find any errors in the values indicated by the instrument. In some cases, calibration assigns a relationship between the input and output of an instrument; for example, calibration of a resistance thermometer could relate its output (in ohms) to an input temperature (in degrees Celsius, or in kelvins).

confidence level - number (e.g. 95 %) expressing the degree of confidence in a result

correction (calibration correction) - compensation for an estimated systematic effect. A number added to an instrument reading to correct for an error, offset, or bias. (Similarly, a reading may be multiplied or divided by a ***correction factor*** to correct the value.)

correlation - interdependence, or relationship, between data or measured quantities

coverage factor - number larger than one by which a combined standard measurement uncertainty is multiplied to obtain an expanded measurement uncertainty, for a particular level of confidence

error - measured quantity value minus a reference quantity value. The offset or deviation (either positive or negative) from the correct value

estimated standard deviation - estimate of the standard deviation of the 'population' based on a limited sample

expanded uncertainty - product of a combined standard measurement uncertainty and a factor larger than the number one. Standard uncertainty (or combined standard uncertainty) multiplied by a coverage factor k, to give a particular level of confidence

Gaussian distribution - (See ***normal distribution***)

influence quantity - quantity that, in a direct measurement, does not affect the quantity that is actually measured, but affects the relation between the indication and the measurement result; e.g., cloudiness in the field-of-view of an instrument can influence the accuracy of its measurement

interval (confidence interval) - interval containing the set of true quantity values of a measurand with a stated probability, based on the information available. The margin within which the 'true value' being measured can be said to lie, with a given level of confidence

level of confidence - number (e.g. 95 %) expressing the degree of confidence in the result

mean (arithmetic mean) - average of a set of numbers

measurand - quantity intended to be measured. The particular quantity subject to measurement

normal distribution - distribution of values in a characteristic pattern of spread (Gaussian curve) with values more likely to fall near the mean than away from it

operator error - a mistake

precision - closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. A term meaning 'fineness of discrimination' but often misused to mean 'accuracy' or 'uncertainty'. Its use should be avoided if possible.

random error - component of measurement error that in replicate measurements varies in an unpredictable manner. An error whose effects are observed to vary randomly.

range - absolute value of the difference between the extreme quantity values of a nominal indication. The interval difference between the highest and the lowest of a set of values

reading - value observed and recorded at the time of measurement

rectangular distribution - distribution of values with equal likelihood of falling anywhere within a range

repeatability (of an instrument or of measurement results) - condition of measurement, out of a set of conditions that includes the same measurement procedure, same operators, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time. The closeness of the agreement between repeated measurements of the same property under the same conditions.

reproducibility (of an instrument or of measurement results) – condition of measurement, out of a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects. The closeness of the agreement between measurements of the same property carried out under changed conditions of measurement (e.g. by a different operator or a different method, or at a different time)

resolution - smallest change in a quantity being measured that causes a perceptible change in the corresponding indication. (e.g. a change of one (1) in the last place of a digital display)

result (of a measurement) - set of quantity values being attributed to a measurand together with any other available relevant information. The value obtained from a measurement, either before or after correction or averaging

sensitivity - quotient of the change in an indication of a measuring system and the corresponding change in a value of a quantity being measured. The change in response (of an instrument) divided by the corresponding change in the stimulus

standard deviation - a measure of the spread of a set of results, describing how values typically differ from the average of the set. Where it is not possible to obtain an infinite set of results (in practice it never is) we instead use the estimated standard deviation.

standard uncertainty - measurement uncertainty expressed as a standard deviation.

systematic error – component of measurement error that in replicate measurements remains constant or varies in a predictable manner. A bias or offset (either positive or negative) from the correct value

true value – quantity value consistent with the definition of a quantity, i.e. the value that would be obtained by a perfect measurement

Type A evaluation of uncertainty - evaluation of a component of measurement uncertainty by a statistical analysis of measured quantity values obtained under defined measurement conditions.

Type B evaluation of uncertainty - evaluation of a component of measurement uncertainty determined by means other than a Type A evaluation of measurement uncertainty

uncertainty budget - statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination

uncertainty of measurement - non-negative parameter describing the dispersion of the quantity values being attributed to a measurand. Alternatively described as a quantity representing the doubt in result of a measurement.

uniform distribution - distribution of values with equal likelihood of falling anywhere within a range

validation - the process of assessing, by independent means, the quality of the data products derived from the system outputs

References

[1] Intermediate Uncertainty Analysis for Earth Observation: Instrument Calibration Module, written within the EMRP-funded Metrology for Earth Observation & Climate (MetEOC) project. http://www.emceoc.org/documents/uaeo-int-trg-course.pdf

[2] JCGM, JCGM 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Report, 2008. www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

[3] JCGM, JCGM 200:2008 International vocabulary of metrology - basic and general concepts and associated terms, Report, 2008. http://www.bipm.org/en/publications/guides/vim.html

[4] JCGM 101:2008 Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf

[5] 'A Beginner's Guide to Uncertainty of Measurement' Measurement Good Practice Guide No. 11 (Issue 2) http://www.npl.co.uk/publications/a-beginners-guide-to-uncertainty-in-measurement

[6] QA4ECV: Gap Analysis of QA4ECV ECV Products D2.2 http://www.qa4ecv.eu/

[7] FIDUCEO project http://www.fiduceo.eu/

[8] Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, Atmos. Meas. Tech., 7, 4463-4490, doi:10.5194/amt-7-4463-2014, 2014. http://www.atmos-meas-tech.net/7/4463/2014/amt-7-4463-2014.pdf

[9] Immler, F. J., Dykema, J., Gardiner, T., Whiteman, D. N., Thorne, P. W., and Vömel, H.: Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products, Atmos. Meas. Tech., 3, 1217-1231, doi:10.5194/amt-3-1217-2010, 2010. http://www.atmos-meas-tech.net/3/1217/2010/amt-3-1217-2010.html

[10] Report on LIDAR measurements by T. Leblanc. http://tmf.jpl.nasa.gov/tmf-lidar/results/ISSI_Team_Report.htm

[11] NORS (Network Of ground-based Remote Sensing Observations in support of the Copernicus Atmospheric Service) deliverable reports. http://nors.aeronomie.be/index.php/documents

GAIA CLIM Product VO Go/No-Go checklist

This document acts as a checklist of the essential requirements of a product needed to attain reference-grade status. Its completion, and independent verification, is a necessary milestone in demonstrating is suitable for inclusion in the GAIA CLIM VO.

Product name: …………………………………………………………………….
Product technique: ……………………………………………………………..
Product measurand: ……………………………………………………………
Product form/range: …… ………………………………………………………
Product dataset: ………………………………………………………………..
Site/Sites/Network location: …………………………………………………
Product time period: ……………………………………………………………….
Data provider: …………………………………………………………………….
Instrument operator: ……………………………………………………………….
Product assessor: …………………………………………………………………..
Assessor contact email: …………………………………………………………….

Checklist version:………………….

Checklist item	Completed Yes/No	Reference/Link/Comment
Complete specific product traceability chain (§2)		
Assess product traceability chain (PTU) according to the GAIA CLIM guidance (§3)		
Produce paper describing contributing uncertainties.		
Produce technical document describing how to make the measurement		
Identify sites using this specific method and the data period covered.		
Has the impact of traceability confidence of the element uncertainties been assessed, according to the guidance?		

Checklist completed by: …………………………………………date ……………………

Checklist reviewed by: ……………………………………..……date ……………………

Accepted for VO	

Completion notes:

All rows in the above table must be completed for a product to enter the GAIA CLIM VO, including links/references to the supporting documents.

The information should then be checked by an appropriate external reviewer. If the form is not accepted the reviewer should provide feedback on the additional input required to address the issues. The review process should then be repeated once a new version of the form, and supporting documents, has been produced.

H2020 GAIA CLIM

WP2 – Product traceability uncertainties & nomenclature guidance document

Paul Green & Tom Gardiner

Types of uncertainty chain

- **Physical Model** – This model considers the real-world physics situation or instrument processing chain to L0 instrument raw data
- **Processing Model** – This model considers how the raw L0 data collected is processed to provide the end product, through calibration to the final geophysical parameter.
- **Metrological Model** – This model considers the calibration, or linkage, of a measurement or processed data to a reference.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Physical model

- Describe the physics of the instrument, considering
  - Incoming radiation
  - Field of view / Point spread function.
  - Scan mechanism (pointing accuracy, repeatability)
  - Input optical system (telescope etc. mirror reflectivity)
  - Spectral response function, side-band rejection.
  - Stray light (thermal self-emission, RFI etc.)
  - Detector assembly (non-linearity, dark signal, read noise, QE)
  - Dominant noise sources (thermal, shot, (pre-)amplifier, Generation-Recombination, etc.)
  - Digitisation (ADC bits)
  - Instrument Housekeeping data (Temperature sensor calibration etc.)
  - …

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Processing model

- Describe the L0 raw data to geophysical parameter processing chain, considering
  - Calibration model (Instrument Gain & Offset)
  - Temporal & Spatial averaging/smoothing
  - Erroneous point removal
  - Ancillary data & associated uncertainties (T, ρ, gas concentration profiles, soil moisture)
  - Model inputs (assumed TOA/BOA reflectance, HITRAN)
  - Geolocation
  - Time of day, SZA, slant path
  - …

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

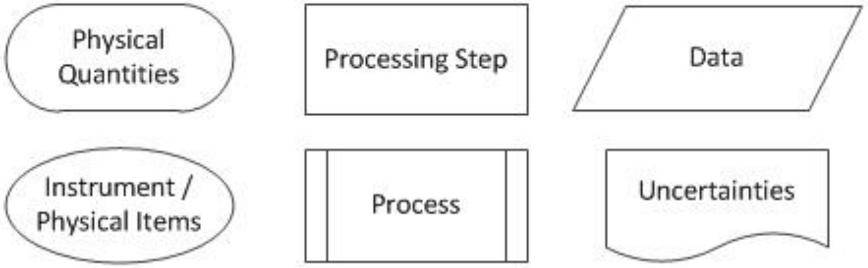
Metrological model

- The metrological model chains are intended to describe the set of calibrations, or linkages, of a measurement (or of processed data) to a reference standard.
- Instrument temperature sensor (±0.5K 2σ Gaussian PDF) calibrated against …
    - Pre-flight check PRT (±0.3K 2σ Gaussian PDF – ±1.0K rectangular usage)
    - Post manufacturing calibration PRT (±0.1K 2σ Gaussian PDF), calibrated against
        - NIST transfer standard (±0.05K 2σ), calibrated against …
            - NIST ITS-90 SI reference (±0.00?K 2σ Gaussian PDF)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

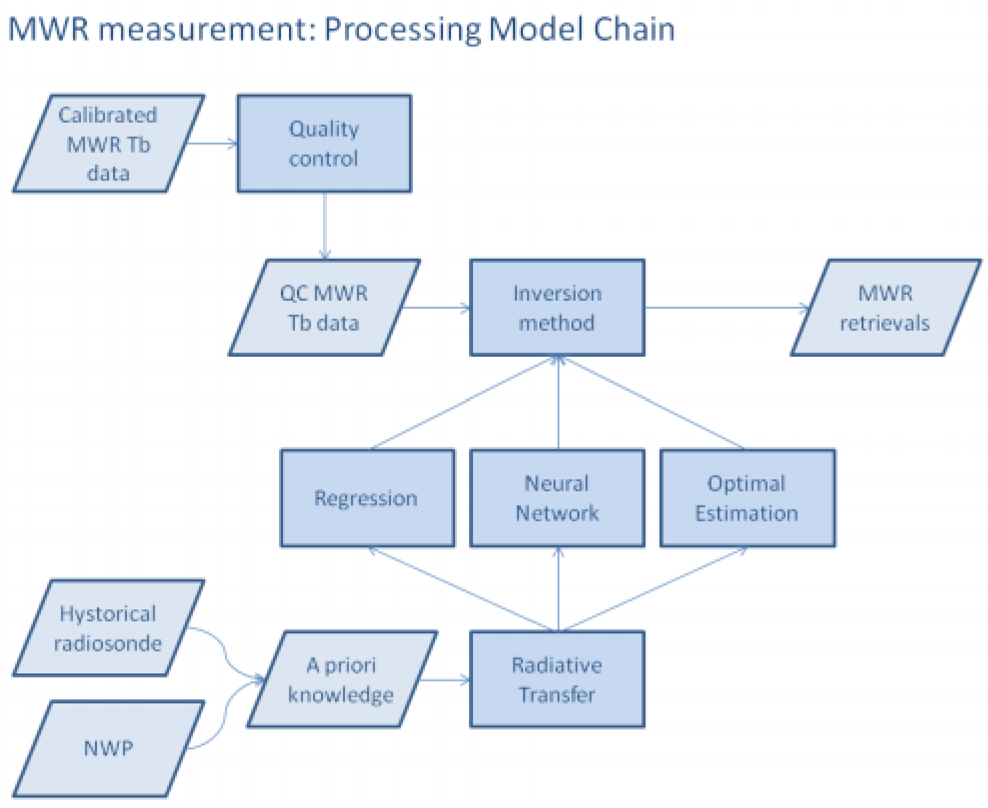
Traceability chain

- The Product Traceability Uncertainty (PTU) chain is a single chain, pulling together all the elements & information from the 3-model method.
- A single PTU chain is to be presented in the product uncertainty document.
- Representation of the chain



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Traceability chain



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Beyond traceability chains

The vision is to move beyond simple traceability chains (which is effectively understanding the process) towards reference grade products, encapsulated in

- technical document describing the measurement procedure
- a publication describing the measurement traceability and its uncertainty

The ultimate goal is to produce metrological-rigorous traceable measurements for the target measurement systems, providing practical coverage factors, applicable in the VO.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM traceability chains

1. Identify the exact measurement that we will quantify within GAIA CLIM VO; specific technique, specific measurand, profile/total column.
2. Identify the specific dataset to be addressed in GAIA CLIM, that will go into the VO
3. Identify the specific product chain for this combination of parameters
4. Identify inputs, (their inputs) the process, the uncertainties and sensitivities of the product to these. To form a measurement equation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM traceability chains

Product name: In-situ radiosonde RS92
Product technique: Capacitive temperature sensor
Product measurand: Temperature
Product form/range: profile (ground to 30km, 1sec sampling)
Product dataset: GRUAN Reference level sonde dataset
Site/Sites/Network location:

- Lindenberg, Germany 52.2100 °N, 14.1200 °E, 98.0 m
- Sodankylä, Finland, 67.3700 °N, 26.6300 °E, 179.0 m

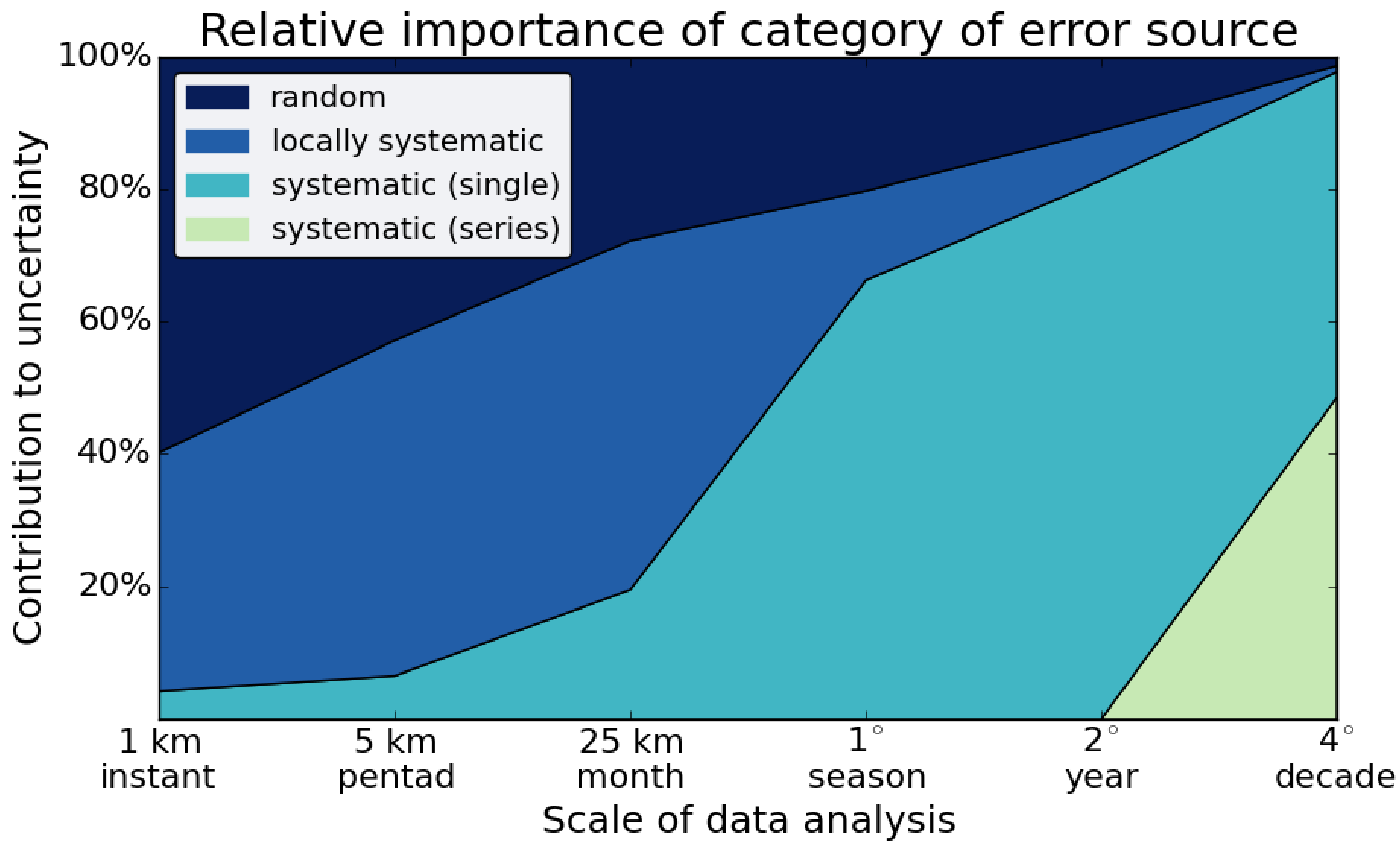
Product time period: Jan 1 – Dec 31, 2014
Data provider: GRUAN
Instrument provider: Meteorologisches Observatorium Lindenberg & Ilmatieteen laitos
Product assessor: Paul Greem NPL
Assessor contact email: paul.green@npl.co.uk

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM traceability chains

5. Characterise the form of the uncertainty, is it random, quasi-systematic, systematic, on what time/spatial scale?
    - Independent random effects (noise)
    - Structured random (regular calibration cycles)
    - Systematic effects (long term correlation / fixed parameter)
6. Can we assign a probability distribution function (PDF)?
7. Ensure that the use of nominally identical instruments from site-to-site and user-to-user is unified in procedure & observing practice. Need an assessment of comparability through usage.
8. Reference to previous work/documentation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.



http://www.fiduceo.eu/content/why-worry-about-all-sources-errors

Product traceability uncertainty chain assessment

- Reference to previous work/documentation should be made where relevant, but this should not detract from the independent of the GAIA CLIM measurement document. The PTU document needs to be stand alone, so understood if read in isolation from the referenced material.
- Identify any site-to-site or user-to-user variation in procedure & observing practice from nominally identical instruments so make an assessment of comparability through usage.
- Each element should have a summary table of knowledge & traceability including an estimate of contribution magnitude.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Element contribution table

Information / data	Type / value / equation	Notes / description
Name of effect		
Contribution identifier		
Measurement equation parameter(s) subject to effect		
Contribution subject to effect		
Element correlation form		
Time correlation form		
Units element correlation		
Units time correlation		
Scales element correlation		
Scales time correlation		
Uncertainty PDF shape		
Uncertainty		
Uncertainty units		
Sensitivity coefficient		
Correlation(s) between affected parameters		
Cross-contribution correlation(s)		
Element/step common for all sites/users?		
Traceable to …		
Validation		

...n the ...rch and innovation programme under grant agreement No 640276.

Element contribution table explanation

- **Name of effect** – the name of the contribution
- **Contribution identifier** - unique identifier to allow reference in the traceability chains. Depends on the chain & submodule structure, but A1… Ax, B1 … Bx etc. may be appropriate.
- **Measurement equation parameter(s) subject to effect –** The part of the measurement equation influenced by this contribution.
- **Contribution subject to effect** – The top level measurement contribution affected by this contribution. Either the product, or chain sub-module contribution effected by the contribution.
- **Temporal correlation form** – the form of any correlation this contribution has in time.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Element contribution table explanation

- **Units time correlation** – the units of any temporal correlation.
- **Scales time correlation –** factors that scale with the temporal correlation.
- **Element spatial/spectral correlation form -** the form of any correlation this contribution has other than in time, be that spatial, spectral, or any other.
- **Units spatial/spectral correlation -** the units of any spatial/spectral correlation.
- **Scales spatial/spectral correlation –** factors that scale with any spatial/spectral correlation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Element contribution table explanation

- **Uncertainty PDF shape** – the probability distribution shape of the contribution, Gaussian, Rectangular, U-shaped, log-normal or other.
- **Uncertainty** – the uncertainty value
- **Uncertainty units** – units of the uncertainty
- **Sensitivity coefficient** – coefficient multiplied by the uncertainty when applied to the measurement equation.
- **Correlation(s) between affected parameters** – Any correlation between the parameters effected by this specific contribution.
- **Cross-contribution correlation(s)** – identify any cross-correlations within the measurement equation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Element contribution table explanation

- **Element/step common for all sites/users** – Is there any site-to-site/user-to-user variation in the application of this contribution?
- **Traceable to** – describe any traceability back towards a primary/community reference.
- **Validation** – Any validation activities that have been performed for this element?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GRUAN RS92 temperature example

5.2 Calibrated in Vaisala CAL4 facility (2)

The CAL4 contains PTU reference sensors that are recalibrated at regular intervals against standards that are traceable to NIST (for pressure and temperature) and its Finnish equivalent, MIKES (for humidity). The respective operating ranges and accuracies of the PTU sensors are 3 (±0.6) to 1080 (±1) hPa, −90 (±0.5) to 60 (±0.5) °C, and 0 (±5) to 100 (±5) % RH, respectively (Vaisala, 2007).

See [1]

Information / data	Type / value / equation	Notes / description
Name of effect	Vaisala CAL4 facility calibration	
Contribution identifier	2	
Measurement equation parameter(s) subject to effect	Temperature product	
Contribution subject to effect	Temperature product	
Element correlation form	none	
Time correlation form	none	
Units element correlation	K	
Units time correlation	K	
Scales element correlation	1	
Scales time correlation	1	
Uncertainty PDF shape	Normal	
Uncertainty	±0.15	
Uncertainty units	K	
Sensitivity coefficient	1	
Correlation(s) between affected parameters	None	
Cross-contribution correlation(s)	None	
Element/step common for all sites/users?	Yes	
Traceable to …	2a - reference T sensor	
Validation	Inter-comparison studies.	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Product traceability uncertainty chain assessment

- This assessment may be via:
  - a formal analytical treatment
  - a sensitivity study
  - an educated guess
- Once the summary table has been completed for the full chain, it should become clear where further work should be focused to most effectively improve the overall level of knowledge of the process uncertainties.
- The summary table, explanatory notes and referenced material in the traceability chain should occupy <= 1 page for each element entry.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

PTU chain summary table

Element identifier/ name	Uncertainty contribution magnitude	Traceability level (L/M/H)	random, structured random, quasi-systematic or systematic?	Correlated to? (Use element identifier)
2	±0.15K	H	random	none
3	$(\Delta T_{GC25}/3)^2$	H	Systematic (over ascent)	5a
2&3	$\sqrt{u_c(\text{cal})^2 + (\Delta T_{GC25}/3)^2}$	H	Systematic (over ascent)	none
8b2	±0.05K	M	quasi-systematic	2
9a1	< 0.03K	M	quasi-systematic	2
1	Std deviation	H	random	none
8a1	$2 \cdot \Delta T/\sqrt{3}$	M	random	none
8a2	$\frac{1}{2\cdot\sqrt{3}}\lvert I_a^{\text{clear sky}} - I_a^{\text{cloudy}}\rvert$	M	Systematic (over ascent)	none
8a2	$\Delta T \cdot u_c(I_a)/I_a$	M	Systematic (over ascent)	none
8a4	±1 m/s	M	quasi-systematic	Altitude product
8a4	$\Delta T \cdot u(v)/v$	M	quasi-systematic	Altitude product
9a2	<0.2K	M	Systematic (over ascent)	none
TOTAL				

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Impact of traceability confidence calculation

- Traceability level: A description of the traceability associated with this element

Traceability Level	Descriptor	Multiplier
High	SI traceable or globally recognised community standard	1
Medium	Developmental community standard or peer-reviewed uncertainty assessment	3
Low	Approximate estimation	10

- The multiplier values provide one possible mechanism to assess this.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Impact of traceability confidence calculation

Element identifier/ name	Uncertainty contribution magnitude	Traceability level (L/M/H)	random, structured random, quasi-systematic or systematic?	Correlated to? (Use element identifier)
2	±0.15K	H	random	none
3	$(\Delta T_{GC25}/3)^2$	H	Systematic (over ascent)	5a
2&3	$\sqrt{u_c(cal)^2+(\Delta T_{GC25}/3)^2}$	H	Systematic (over ascent)	none
8b2	±0.05K	M	quasi-systematic	2
9a1	< 0.03K	M	quasi-systematic	2
1	Std deviation	H	random	none
8a1	$2 \cdot \Delta T/\sqrt{3}$	M	random	none
8a2	$\frac{1}{2\cdot\sqrt{3}}\lvert I_a^{\text{clear sky}} - I_a^{\text{cloudy}}\rvert$	M	Systematic (over ascent)	none
8a2	$\Delta T \cdot u_c(I_a)/I_a$	M	Systematic (over ascent)	none
8a4	±1 m/s	M	quasi-systematic	Altitude product
8a4	$\Delta T \cdot u(v)/v$	M	quasi-systematic	Altitude product
9a2	<0.2K	M	Systematic (over ascent)	none
TOTAL				

www.gaia-clim.eu

ed funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Impact of traceability confidence calculation

- The output from this work will be a measurement product technical document which should be stand-alone i.e. intelligible in isolation.
- Reference to external sources (preferably peer-reviewed) and documentation from previous studies is clearly expected and welcomed, but with sufficient explanatory content in the GAIA CLIM document not to necessitate the reading of all these reference documents to gain a clear understanding of the GAIA CLIM product and associated uncertainties entered into the VO.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM Product VO Go/No-Go checklist

GAIA CLIM Product VO Go/No-Go checklist

This document acts as a checklist of the essential requirements of a product needed to attain reference-grade status. Its completion, and independent verification, is a necessary milestone in demonstrating is suitable for inclusion in the GAIA CLIM VO.

Product name: ……………………………………………………………………………
Product technique: ………………………………………………………………………
Product measurand: ………………………………………………………………………
Product form/range: …… …………………………………………………………………
Product dataset: …………………………………………………………………………
Site/Sites/Network location: ……………………………………………………………
Product data range: ………………………………………………………………………
Providing Institution: ……………………………………………………………………
Contact name/email:…………………………………………………………………………

Checklist version:……………………

Checklist item	Completed Yes/No	Reference/Link/Comment
Complete specific product traceability chain (§2)		
Assess product traceability chain (PTU) according to the GAIA CLIM guidance (§3)		
Produce paper describing contributing uncertainties.		
Produce technical document describing how to make the measurement		
Identify sites using this specific method and the data period covered.		
Has the impact of traceability confidence of the element uncertainties been assessed, according to the guidance?		

Checklist completed by: ………………………………………………date ……………………

Checklist reviewed by: …………………………………………………date ……………………

Accepted for VO	

www.gaia-clim.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM Product VO Go/No-Go checklist

- Product name:
- Product technique:
- Product measurand:
- Product form/range:
- Product dataset:
- Site/Sites/Network location:
- Product time period:
- Data provider:.
- Instrument operator:
- Product assessor:
- Assessor contact email:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

GAIA CLIM Product VO Go/No-Go checklist

Checklist item	Completed Yes/No	Reference/Link/Comment
Complete specific product traceability chain (§2)		
Assess product traceability chain (PTU) according to the GAIA CLIM guidance (§3)		
Produce paper describing contributing uncertainties.		
Produce technical document describing how to make the measurement		
Identify sites using this specific method and the data period covered.		
Has the impact of traceability confidence of the element uncertainties been assessed, according to the guidance?		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.

Questions?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 640276.