METROLOGICAL APPROACH FOR EO

GENERAL GUIDANCE ON A METROLOGICAL APPROACH TO FUNDAMENTAL DATA RECORDS (FDR), THEMATIC DATA PRODUCTS (TDP) AND FIDUCIAL REFERENCE MEASUREMENTS (FRM) – UNCERTAINTY ANALYSIS PROCESS

NPL QA4EO Team

JUNE 2025

General guidance on a metrological approach to fundamental data records (FDR), thematic data products (TDP) and fiducial reference measurements (FRM) – Uncertainty Analysis Process

NPL QA4EO team
Lead author: Emma Woolliams
Other contributing authors (alphabetical):
Sajedeh Behnia
Agnieszka Bialek
Pieter De Vis
Jacob Fahy
Nigel Fox
Tom Gardiner
Paul Green
Samuel Hunt
Jonathan Mittaz
Bernardo Mota
Joanne Nightingale
Niall Origo
Anna Pustogvar
Madeline Stedman

National Physical Laboratory

Hampton Road, Teddington, Middlesex, TW11 0LW

This work was carried out in the frame of the Instrument Data quality Evaluation and Assessment Service - Quality Assurance for Earth Observation (IDEAS-QA4EO) contract funded by ESA-ESRIN (n. 4000128960/19/I-NS), and builds on the work of previous projects, see acknowledgments.

Version control

Issue	Date	Authors	Reviewed by	Notes
1.0	19.5.22	As above	Authors	First issue
1.1	30.10.24	As above	Authors	Minor updates for second issue
1.2	20.6.25	As above	Authors	Reviewed and updated

Table of contents

1. Introduction

1.1. This set of documents

This document is part of a set of documents that describe and give practical tools to support a metrological approach for a satellite fundamental data record (FDR), satellite-derived thematic data product (TDP), or fiducial reference measurement (FRM) network or campaign.

Document	Description
Executive Summary	Introduction and overview of what a metrological approach is and what is needed to implement a metrological approach to FDRs, FRMs and TDPs
Metrology Document	A document that describes the metrological principles behind the approach
Process Document	This document, describing the step-by-step processes needed to implement a metrological approach
Templates Document	Templates and examples of uncertainty tree diagrams and effects tables and how to structure an uncertainty report
Toolkit introduction	An introduction to the COMET Python tools

1.2. Scope of this document

This document describes a process for generating an uncertainty budget for an FDR, TDP or FRM. It describes the five steps towards an uncertainty budget.

1.3. General introduction common to all documents

Earth Observation (EO) satellite programmes are operated by a wide variety of space agencies, meteorological agencies and commercial operators and provide observations for a wide range of social, scientific, environmental, and commercial applications. Historical and current EO data provide information about environmental and climate change that is of great value to today's scientists and to decision makers. These data are also a legacy of immense value to future generations. However, for this immediate and legacy value to be realised, EO data sets must be interoperable and temporally stable, so that data from different sensors can be combined. The quality and uncertainty associated with datasets is also needed to assess their fitness for purpose for the desired applications.

Metrology is the discipline responsible for maintaining the International System of Units (SI) and the associated system of measurement. It is core to the SI, that measurements are stable over very long time periods, that measurement standards are equivalent worldwide and that measurements are coherent – that is different types of measurement can be combined because, for example, an electrical watt is equivalent to an optical watt is equivalent to a mechanical watt.

These properties of metrology are desired for EO data records. It is for this reason that over the last two decades there has been considerable research in the collaborative field of EO Metrology. The 2010 endorsement of the Quality Assurance Framework for Earth Observation (QA4EO) by the Committee on Earth Observation Satellites (CEOS) in the frame of the Global Earth Observation System of Systems (GEOSS) set up the basic principle that EO data should be accompanied by a fully traceable indicator of its quality, allowing users to readily assess the fitness for purpose for their applications. Traceability requires that this quality indicator be based on 'a documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards'. QA4EO stops short of requiring robust metrological traceability, but the accompanying guidelines are based on principles adapted from the metrology community.

Since 2010, collaborative EO-metrology projects have been developing robust methods to facilitate broader use of metrological principles in EO applications. In Europe, such projects have been led through European research funding (FP-7, Horizon 2020) and by projects from the European Space Agency, and more recently the broader Copernicus Programme via institutes such as EUMETSAT and ECMWF. In this document we build on this legacy of activity and expand the concepts, nascent in the FIDUCEO project [Mittaz et al 2019], generalising them beyond passive radiometric band sensors, to establish FDRs, TDPs and FRMs.

1.4. Common introduction to FDRs, TDPs and FRMs

The terms Fiducial Reference Measurement (FRM), Fundamental Data Record (FDR) and Thematic Data Product (TDP) were applied initially by the European Space Agency to describe metrologically rigorous observations of specific relevance to space-based observations. The FRM and TDP definitions given here have not yet been formally endorsed by a committee, although they are increasingly being used by the broader Earth observation community and there have been some workshops discussing them.

A fundamental data record (FDR) is a record, of sufficient duration for its application, of uncertainty-quantified sensor observations calibrated to physical units and located in time and space, together with all ancillary and lower-level instrument data used to calibrate and locate the observations and to estimate uncertainty.

Generally, FDRs will be geolocated level 1 products. The FDR provides a record of the physical quantity measured by the sensor, along with the ancillary (additional) information needed to interpret it. Although some applications in reanalyses ingest level 1 products, for many applications FDRs will be used to generate TDPs.

*A **thematic data product** (TDP) is a record, of sufficient duration for its application, of uncertainty-quantified retrieved values of a geophysical variable, along with all ancillary data used in retrieval and uncertainty estimation.*

TDPs provide higher level products that have been processed from FDRs, through algorithms which also often combine information from other FDRs (e.g. from other satellite sensors) or from external information (such as reanalysis models and/or certain non-satellite data), along with such additional information.

Note that the terms 'Fundamental Climate Data Record' (FCDR) and 'Climate Data Record' (CDR) are used for FDRs and TDPs respectively, that are also typically of multi-decadal duration and come from a series of sensors that have been harmonised to a common reference and have value for climate studies.

The definition of FRMs has recently been proposed by CEOS and is available on the CEOS Cal/Val portal: (https://calvalportal.ceos.org/web/guest/frms-assessment-framework).

*A **fiducial reference measurement** (FRM) is a suite of independent, fully characterised, and traceable (to a community agreed reference, ideally SI) measurements of a satellite relevant measurand, tailored specifically to address the calibration/validation needs of a class of satellite borne sensor and that* *follow the guidelines outlined by the GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO).*

Thus, FRMs are the quality-assured observations that can be used to calibrate and validate satellite-based sensor measurements. They will often be in-situ (non-satellite) observing systems, but some planned reference satellites, such as the SITSats[1], could also be considered FRMs.

As ESA states 'these FRM provide the maximum return on investment for a satellite mission by delivering, to users, the required confidence in data products, in the form of independent validation results and satellite measurement uncertainty estimation, over the entire end-to-end duration of a satellite mission.' CEOS has established detailed guidelines for assessing observational systems to evaluate them as 'CEOS-FRMs', with four different classes of FRM possible, depending on to what extent they meet the different criteria.

2. Steps to an uncertainty analysis

The GUM introduction (Guide JCGM 104) and the GUM guide to measurement models (GUM 6) both give a set of steps to go through to support the development of a measurement model and from that an uncertainty analysis. Other metrology documents place slightly different emphases on the different steps, or combine and expand steps in different ways, but overall, there is a consensus in the metrology community about how an uncertainty budget[2] should be developed.

The different EO metrology projects (list given at QA4EO acknowledgements) have adapted these stages to emphasise covariance more strongly and have provided standardised tools for performing some of the analysis. The steps are described as:

- Step 1: Define the measurand and measurement model.
- Step 2: Establish the traceability with a diagram.
- Step 3: Evaluate each source of uncertainty and fill out an effects table.
- Step 4: Calculate the data product and uncertainties.
- Step 5: Record information about the uncertainty analysis for long term data preservation purposes (implicit above) and summarise for today's users.

These steps are discussed in the subsections that follow.

2.1. Step 1: Define the measurand and the measurement model

In environmental observations defining the measurand is often the most difficult step, and it often involves some deep thinking and a challenging conversation between different experts to define exactly what the 'measurand' is. Many observation experts do not like the concept of a 'measurement model' because they don't think they 'measure' the 'measurand'. This is because in environmental observations, there is always what the GUM calls a 'multi-stage measurement model', where measured values are processed through different levels to obtain new measurands, e.g., the processing of top-of-atmosphere radiance into top-of-atmosphere reflectance through a

[1] SITSat stands for 'SI-traceable satellite', which is a class of satellite with considerably lower uncertainties than similar missions and where traceability to SI is clearly documented and validated.

[2] The term 'budget' is frequently used in the form 'uncertainty budget'. It describes a table of different sources of uncertainty and is perhaps somewhat out-of-date. 'Uncertainty analysis' is probably preferable.

transformation, and it is common to combine measurements and models in further data processing (e.g., to turn top-of-atmosphere reflectance into ground reflectance using an atmospheric correction radiative transfer code). It is important to recognise that the term 'measurand' can be used at different levels of a process and may include both models and measurements.

Beyond this, it can be difficult to define exactly what the measurand represents – for example, is it a single point measurement, or is it considered to represent an area and a time block? A temperature may be taken at a specific location and at a specific time; but be used to represent the average temperature in a region, over a time window of an hour, for example. Or perhaps several different temperature measurements are averaged to represent that block. In some communities the measurand itself may not be uniquely defined. For example, a concept such as 'air temperature' is fundamentally different depending on the type of housing the thermometer is placed in, and sea surface temperature measured from a satellite (top micron of the surface) is different from sea surface temperature measured in situ (usually a few cm to metres below the surface). A sea level anomaly is calculated relative to a reference surface, itself calculated relative to either the Earth's geoid or ellipsoid. It is very important to be clear what exactly the measurand is (what the measured value represents).

Once the measurand is clear, the next step is to establish the measurement model that describes the relationship with the raw input quantities and the measurand. As described in GUM 6, this can be an iterative process starting with a more basic model (e.g. a linear relationship between sensor counts and a measured radiance) and then including additional quantities to represent the necessary corrections (e.g. nonlinear terms, corrections for temperature sensitivity, etc), whether applied or not, and model approximations (e.g. representing a spectral integral as a trapezium-rule summation).

The measurement model should include all quantities that affect the measurement result. These include quantities that are measured as input quantities, and quantities that represent corrections – e.g. a term to describe an instrument's temperature sensitivity. Sometimes, however, we recognise that there are 'poorly understood effects' that are known to exist, but little can be said about their form or magnitude[3]. Well-understood effects can be directly incorporated as named input quantities in the measurement model; poorly understood effects are formally encapsulated in a random variable, denoted by Δ or ε, say. This variable is treated as having expectation zero (for additive effects) or unity (for multiplicative effects). Including the term ensures that the measurement model used to calculate the measured quantity is the same as that used to assess uncertainties.

Even after including poorly understood effects, it is important to recognise that measurement models are themselves approximations to a more complex reality. Approximations may involve using a low-order empirical model to describe a higher-order phenomenon, or numerical approximations, such as replacing an integral with a summation, or truncating a series expansion. There remains some, for now negligible, uncertainty in the specification of the measurand that should be remembered for future analyses. Such approximations can be accounted for similarly to the poorly understood effects.

Thus, the measurement model could be considered to take the form:

$$Y = f(X_1, \dots, X_N; E_1, \dots, E_M\ ; \Delta_0)$$

where the X_i represent input quantities from well-understood effects, the E_j represent input quantities for poorly understood effects and the Δ_0 represents input quantities relating to the

[3] The term 'poorly understood effects' is used in the GUM document on developing and using measurement models, GUM-6 available at: https://www.bipm.org/en/committees/jc/jcgm/publications

approximations and assumptions associated with the form of the measurement model or the definition of the measurand. While such notation is strictly unnecessary (these are all input quantities), and not currently described in the GUM, separating the concepts in this way can act as an aide memoire to support practitioners in considering all their sources of uncertainty.

Note, that in [Mittaz et al 2019], and many existing analyses, the equation for measured values was written with a "plus zero" term:

$$y = f(x_1, \dots, x_N) + 0.$$

Such a formulation is not mathematically robust but can act as a useful 'aide-memoire' and can perhaps be incorporated more easily into uncertainty tree diagrams (see below), than more correct formulations.

In some environmental observations, it may not be possible to write the measurement model explicitly as an analytical expression. This is, for example, true when inverse models form part of the processing (e.g. in atmospheric retrievals), where quantities are determined through a fit (e.g. fitting the range to an altimeter waveform) or where an iterative process is used to refine an initial estimate with a more complex model (e.g. in determining a model for how lunar irradiance changes with lunar phase, the determined model itself is needed to account for lunar phase changes during the individual daily observations that are used in the model). Where the measurement model cannot be written out explicitly, for the purposes of uncertainty analysis it is still possible to describe it as a function of input quantities, listing the input quantities, but not the functional relationship.

Finally, it is important to realise that most measurement models will often be written as though they are univariate models, obtaining a single measured value, or multivariate models obtaining a single set of measured values (e.g., in altimetry obtaining the set of parameters from the retrieval). However, these equations describe a single observation in what will almost certainly be one of many observations – because the instrument takes such measurements at different times, in different locations or in different spectral bands.

2.2. Step 2: Establish traceability with a diagram

Metrology itself is a multistep process, with metrological traceability being defined in the third edition of the VIM as the

> *property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.*

This implicitly recognises that the input quantities in a measurement model almost always come from their own measurement models (a multistage measurement model). In most metrological measurements, a full uncertainty budget has been calculated at previous steps – that is what is meant by 'a documented unbroken chain…'. In environmental observations, however, this is rarely the case and therefore more work is required to establish the full traceability.

In establishing traceability, diagrams can be very helpful. Such diagrams can both show where terms come from (traceability) and also highlight sources of uncertainty in input quantities and in the approximations and assumptions inherent in the model.

Step 2 is therefore to establish traceability by drawing one, or more, diagrams. The exact choice of diagrams is up to the individual developing the analysis, but there are four types of diagram that have been found to be helpful and may well be many more possibilities. This list is not intended to be exhaustive, nor is it necessary to produce all these diagrams for any particular uncertainty analysis. Note that there is some overlap between these types. Further information, including some ‘rules’ to make sure the diagrams are consistent with what others are doing, are given in Appendix A.

Uncertainty Tree Diagrams. The FIDUCEO project developed the concept of ‘uncertainty tree diagrams’. Such diagrams put the measurement model in the centre and then show the origin of each term of the measurement model, and the measurement models that are used to derive those terms. At the outside of the diagram, the original sources of uncertainty are identified on the leaves of the tree. The first uncertainty tree diagram developed in FIDUCEO was the one for the AVHRR thermal infrared radiance level 1 product repeated here as Figure 2.1. The central equation shows how Earth radiance, L_{E}, is calculated from the measured Earth signal (count, C_{E}) and an onboard calibration with an onboard calibration target, whose radiance L_{T} is itself calculated from its temperature T_{T}. The a vector represents a set of calibration coefficients that are themselves determined through a fitting process based in flight comparisons with other sensors. The upper part of the tree shows this process conceptually, using a simple ‘function of’ notation, rather than writing that relationship explicitly. Uncertainty tree diagrams show the origin of the input quantities in a measurement model, and the sources of uncertainty that affect them.

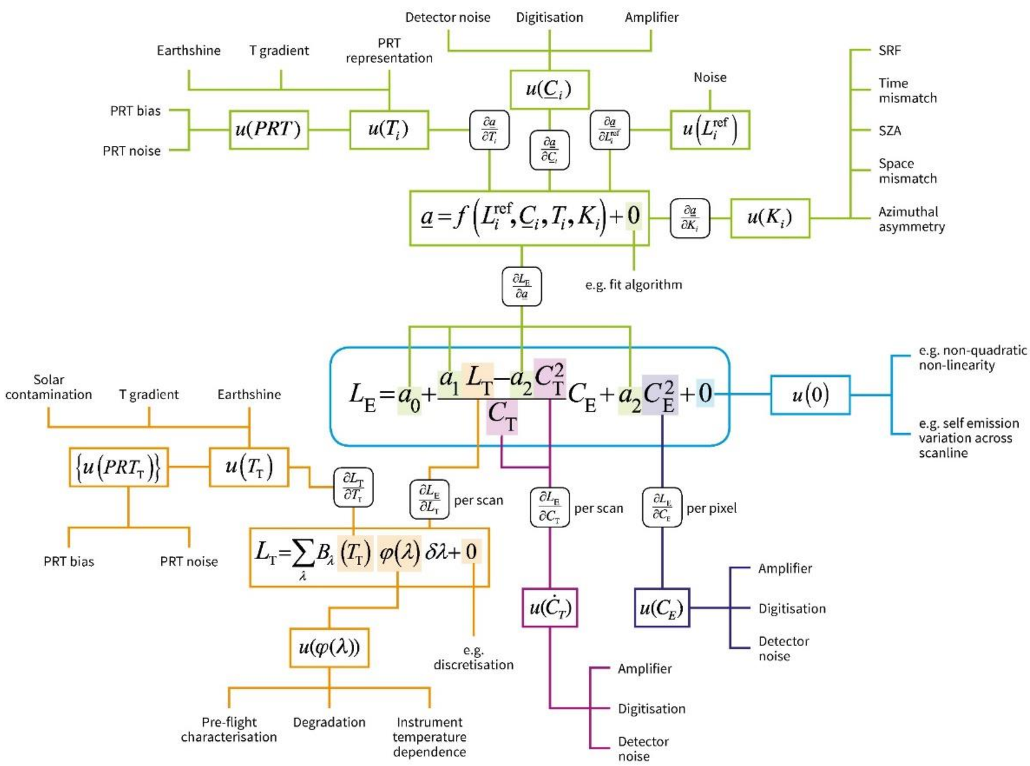


Figure 2.1 The first FIDUCEO uncertainty tree diagram from Mittaz et al 2019

Note that this uncertainty tree diagram used a ‘plus zero’ to represent the hidden assumptions and approximations built into the form of the measurement model. As discussed above, this notation is perhaps not mathematically rigorous, but a useful aide memoire. Alternatively, a ‘plus delta’ could be used, or, more recently, some people have simply created a separate box that emphasises assumptions, for example as shown in Figure 2.2. In some cases, more information is needed to describe those approximations and assumptions than words. In such cases, a 'derivation diagram' (see below) can be helpful.

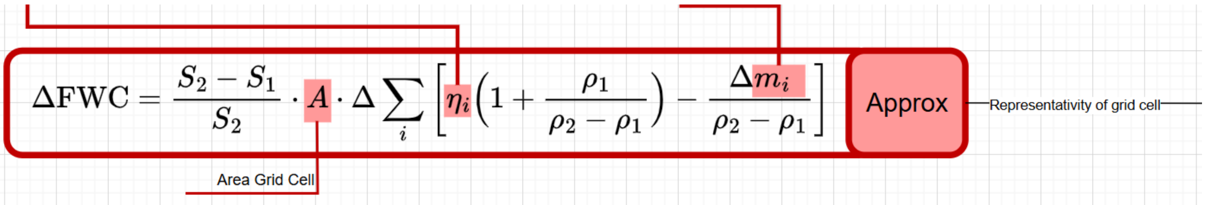


Figure 2.2 Example of a simple presentation of approximations and assumptions in the measurement model that does not use a 'plus zero'.

Processing diagrams. The project QA4ECV developed processing diagrams to show traceability and these have been widely used in other projects since, and are often used separately from a metrological uncertainty analysis. It is difficult on an uncertainty tree diagram to represent sequences of processing steps, or to represent non-analytical processes (such as identifying and removing pixels where there is cloud contamination of the signal). In such cases, a processing diagram, in the form of a flow chart, can be helpful. To use a processing diagram for uncertainty analysis, it is important to identify for each step in the process what auxiliary information is included (and the uncertainties associated with that), and what assumptions and approximations are included in the processing itself. An example processing diagram is given in Figure 2.3. The numbers in red are a hierarchical numbering system to identify each process uniquely. The central processing chain has steps numbered 1, 2, 3 etc. Processing chains for auxiliary and ancillary information has a number that relates to which stage of the chain it is introduced in. For example, there are three 'post measurement corrections' (step 8): the radiative correction combination (8a), smoothing and spike removal (8b) and sonde rotation effects (8c). In turn the radiative correction combination is calculated from two elements labelled 8a1 and 8a2.

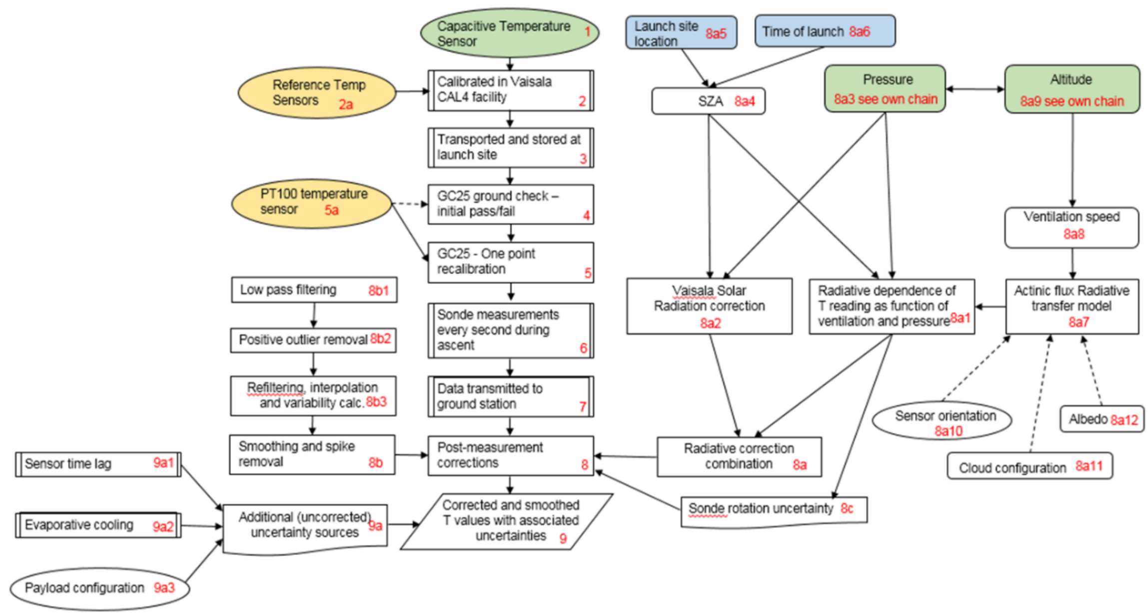


Figure 2.3. Example process chain for a RS92 radiosonde temperature measurement.

Metrological traceability diagrams. For more straightforward instruments, a simple metrological traceability diagram is likely to be the most suitable. Such a diagram represents the 'unbroken chain of calibrations', usually using one shape for an instrument, and another shape for a measurand quantity, and shows each instrument in turn and what it measures. A simple traceability chain for the calibration of the APEX instrument is given in Figure 2.3, and came from earlier training material that may still be available on the MetEOC project website https://www.meteoc.org/training-courses/ although that website is no longer maintained. This type of diagram can usually be helpfully extended into an uncertainty tree diagram, with equations given for each calibration step.

In this diagram, the flow of traceability from SI-calibration to the in-field measurements is demonstrated as a sequence of steps.

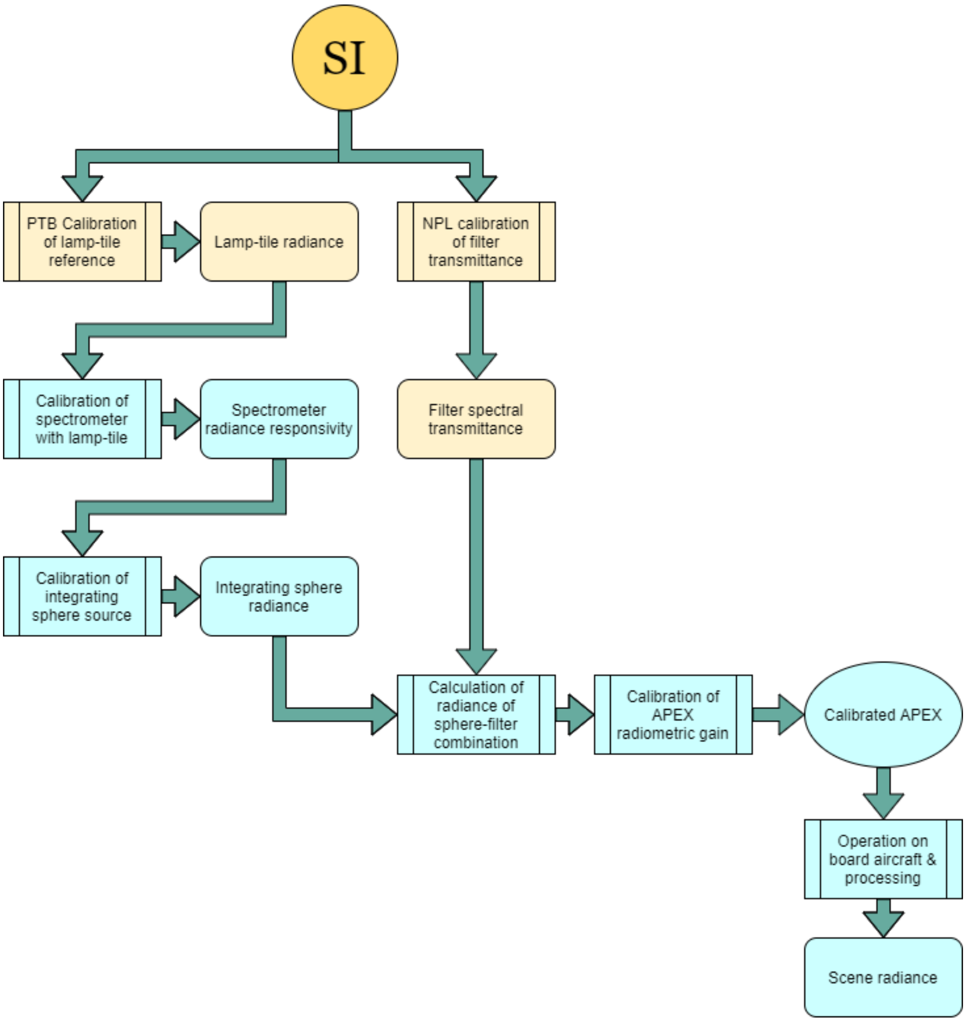


Figure 2.4 A simplified traceability chain for the calibration of the APEX spectrometer carried on the DLR research aircraft

Derivation diagrams. Most recently, for the ASELSU project on sea level rise, a new type of diagram was developed, see: https://www.aselsu.org/diagrams/derivation-diagram. The derivation diagram is like an uncertainty tree diagram in that it provides equations, but its focus is on the assumptions that are built into the derivation of the measurement model. That means that the derivation diagram is a means of describing such assumptions and approximations more completely than as 'plus zero' or 'approximation' terms in the uncertainty tree diagram.

Figure 2.5 shows the ASELSU derivation diagram for the equation used to fit a waveform in sea level measurements from a low-resolution mode radar altimeter. The equation at the top is the descriptive radar equation that describes the interaction of the radar pulse with the ocean surface. Under four assumptions, which are explicitly described on the diagram, it is possible to derive the equation at the bottom, known as the Brown model. Radar altimeter retrackers fit the Brown model to derive the 'range' from the satellite to the ocean surface. Some of the assumptions are partially corrected elsewhere in the uncertainty tree diagram. This derivation diagram, however, clearly presents those assumptions, and can therefore support uncertainty analysis.

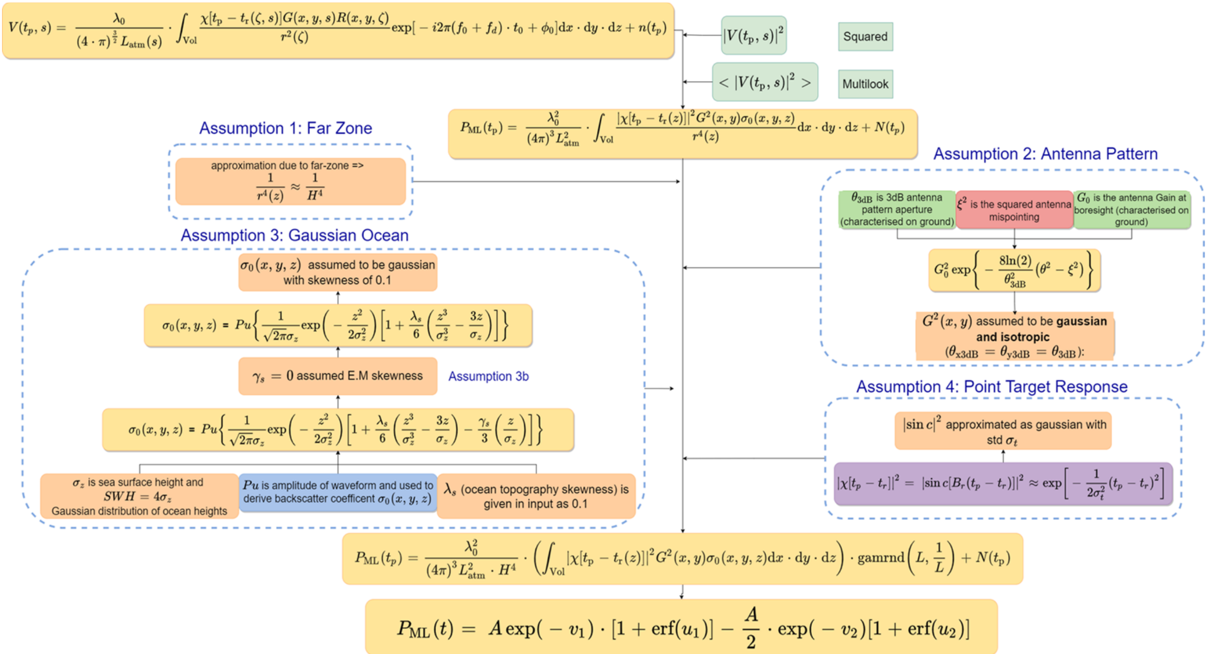


Figure 2.5 Part of the derivation diagram for sea level rise. Although the equations are too small to be seen here, it is possible to see how the different assumptions that are built into the derivation of the form of the waveform are introduced. Identifying these assumptions explicitly allows for further analysis of whether those assumptions are corrected appropriately later in the processing and/or what uncertainties may arise from those assumptions.

There is no single correct approach to Step 2. The diagrams given here are examples that have been found useful in satellite Earth observation analyses. A team working on uncertainty analysis can use any combination of these or develop adaptations of their own. The purpose of the diagrams is to help the team working on an uncertainty analysis to think systematically about the full system they are considering, and to help that team communicate the important aspects of the uncertainty process to other scientists. There are advantages in using diagrams similar to those used in other projects. Doing so, provides an initial recognition that aids understanding and to assist further with this, some rules have been developed on how to draw such diagrams (see Appendix A). However, the important thing is to find the right type of diagram for a particular application and if necessary, to adapt, combine or, sometimes, invent new diagram types.

2.3. Step 3: Evaluate each source of uncertainty and document in an effects table

After the work in step 1 to specify the measurand, and in step 2 to identify where the input quantities of the measurement model all come from, it should be possible to get a list of sources of uncertainty (also known as effects) for each input quantity of the measurement model, and for any assumptions or approximations in the form of the measurement model. Identifying the different sources of uncertainty where they arise in the observational process, is sufficient, because with a full 'multi-stage measurement model' it is then possible to propagate uncertainties through to the measurand of interest.

There are several things that need to be known about each effect so that that propagation can be carried out. FIDUCEO and GAIA-CLIM used the concept of an 'effects table' to document such information systematically. Note that an 'effect' here is a general term describing a 'source of uncertainty'. It describes the underlying process that affects the measured values.

An effects table should have a row or column for each of the following pieces of information:

- Unique identifier of effect (name and/or number code)
- Which term in the measurement model this effect affects
- Magnitude of the uncertainty, and if appropriate, the sensitivity coefficient for the effect
- Where appropriate, the probability distribution function for the uncertainty (unless all are assumed to be Gaussian)
- Whether there is any error correlation between an error in this effect and an error in another effect.
- For each 'dimension of interest' the form of the error correlation and the parameters that define that error correlation form (e.g. length of correlation)
- A maturity indicator that describes how this effect has been evaluated (Is the information in the table based on strong evidence validated with comparisons? Or is it based on expert judgement?)

The tables start with documenting which term in the measurement model this effect affects. It is possible that more than one effect influences each term, and therefore a unique identifier (a name and/or number) is also needed. It may be useful to explain whether this effect affects a term in the primary measurement model (the one used to calculate the measurand), or whether it affects a term in a measurement model for one of the input quantities. GAIA-CLIM used a hierarchical numbering system to make this clear (as seen in the red numbering in Figure 2.3 and described above that diagram).

As with all uncertainty analyses, we also need to identify how large the uncertainty associated with this effect is and how that propagates to the uncertainty in the measurand. If we perform the calculation through Monte Carlo methods, that propagation is handled in the analysis, if we use the Law of Propagation of Uncertainties, then a sensitivity coefficient is needed. The sensitivity coefficient is usually the first derivative of the measurement model with respect to the input quantity the effect affects. A chain-rule of differentiation is used if the uncertainty is an input quantity of a model of an input quantity of the principal measurement model, e.g. for the temperature of the onboard calibration target in Figure 2.1, an uncertainty associated with the temperature affects the determined radiance of the target, and that in turn affects the main measurement model that calculates Earth radiance, thus:

$$\frac{\partial L_E}{\partial T_T} = \frac{\partial L_E}{\partial L_T}\frac{\partial L_T}{\partial T_T}.$$

Evaluating the magnitude of the uncertainty can itself be difficult. Unlike in laboratory experiments, it is difficult to obtain a systematic test of the sensitivity of the instrument to constant or changing conditions as the environmental condition of the instrument and the natural variability of the measurand make repeat measurements almost impossible. However, it is beyond the scope of this document to describe how such evaluations can be done. There are many examples in the publications of the projects given as examples on the QA4EO website.

It is also important to understand and document the error correlation structure of each effect (source of uncertainty). We need to do this for two fundamentally different reasons. First, the different input quantities in a measurement model may be correlated with one another, either due to a common underlying phenomenon, or because they were both derived (e.g. through fitting) from the same raw data. It is important to identify such correlations as this will affect the calculation of the uncertainty associated with the measurand.

Second, beyond such correlations between terms within the measurement model (which are relatively rare in environmental observations), it is also important to think about correlations that will matter not for the calculation of the measurand, but for the work by other scientists at later "levels" of processing. For that we need to identify what the key 'dimensions of interest' are for those future applications.

The concept of 'dimensions of interest' was discussed in the Metrology Document section 3.2. These are the different dimensions for which different observations will be combined or compared at later processing levels. For an imaging radiometric satellite sensor, these will include along-track and cross-track spatial dimensions, a time dimension and a spectral dimension. There may also be a dimension relating to the viewing angle of the surface (whether the sensor views at nadir or at different angles). An effects table should have different rows or columns representing each of the 'dimensions of interest' and should present the error correlation form in each of these dimensions.

To provide a simple way to describe the error correlation form and its size, several formal correlation forms have been proposed. These have been introduced in the Metrology Document. Each correlation form can be parameterised by a small number of parameters. Note, that here we describe the parametrisation of the error correlation form, $r(x_i, x_j)$: the covariance is obtained from $u(x_i, x_j) = u(x_i)u(x_j)r(x_i, x_j)$ and that requires knowledge of the uncertainty itself.

Error correlation form	**Notes**	**How it is parametrised**
random	Fully independent errors, correlation matrix is a diagonal matrix	No parametrisation needed
rectangular_absolute	Systematic effects within a range, correlation matrix is a block of 1s within each range	Start and end of (each) range.
triangular_relative	For simple rolling averages, correlation matrix is a banded-diagonal reducing over a number of scanlines defined by the width	Half-base width

bellshaped_relative	For weighted rolling averages, etc. Assumed that correlation off the diagonal drops as a Gaussian until a cut-off point	Half-base width and Gaussian width
OTHER	Other correlation forms can be and have been defined	As needed

The final thing that needs to be considered when evaluating a source of uncertainty and documenting our findings is how mature our estimate of the properties we have considered is. This is, of course, a qualitative concept and may be somewhat subjective.

However, any practical uncertainty analysis of a complex measurement such as those in environmental observations, will involve combining sources of uncertainty where the magnitude of the uncertainty and the error correlation structures are well known, with other sources of uncertainty where it is not possible to perform a quantitative assessment of these properties. It is important that in documenting uncertainty analysis, indications are also given on how 'mature' the estimate of the magnitude of the uncertainty and the error correlation structures are. Where a source of uncertainty is well understood (has high maturity), then the evidence should also be referenced.

The highest maturity will be achieved when a source of uncertainty has both been evaluated robustly (e.g. through repeat measurements, or through careful modelling) and has been validated in some way by comparison to independent data sets. A moderate maturity is when one of these is in place but not the other (e.g. the uncertainty has been evaluated but not validated or has been estimated through comparison to an independent data set, but not evaluated from more fundamental principles). A low maturity would be where expert judgement has been used to estimate an uncertainty magnitude.

Low maturity estimates may need to be made because the source of uncertainty is fundamentally unknowable. This may be particularly the case with historical datasets where original information does not exist. It may also happen with particularly complex phenomena that cannot be separated from other effects. It is also possible that a low maturity estimate was made as an intermediate step, while the scientists concentrated on other aspects of the uncertainty budget that they believed to be more significant and while they were constrained by budgetary, practical or time limitations.

Example effects tables are given in Appendix B. The CoMET tools provide a means to store effects tables digitally and to use the effects tables to generate an error correlation or error covariance matrix.

2.4. Step 4: Calculate the quantity and the uncertainties and covariances

Once each source of uncertainty has been identified (step 2) and evaluated (step 3), the next step is to calculate the measurand and its associated uncertainty.

As discussed in step 1 (Section 2.1), those who make environmental measurements, particularly at processing levels 1 and 2, will generally develop the analysis described here for a single observation of the measurand they will then apply this model to many observations (in other words, they often write the measurement model as though it were univariate, assuming that it can then be recalculated for each element of the actual multivariate case). It is only at higher levels that these individual observations (at different locations, times and perhaps at different frequencies or angles) are combined and/or compared along the different dimensions of interest. It is also common for observations to be initially processed as 'observational' measurements (giving near-real-time data streams), and then later reprocessed (perhaps multiple times) to take advantage of improved models for e.g., atmospheric corrections, or improved knowledge of the instrument's performance long term. It is in these later reprocessings that full uncertainty information can generally be considered.

The opportunity to reprocess the data also means that it may be possible to improve the operational measurement model with a more sophisticated model that accounts for more environmental and instrumental effects. It is common for a robust uncertainty analysis to highlight areas where the model can be improved, to correct for an effect that is better understood after the uncertainty analysis.

Alongside calculating the measured quantity values at different times, different locations and, where appropriate different wavelengths and angles, the uncertainty associated with the quantity value must also be determined. To do this, each uncertainty effect must be propagated through the measurement model for each observation. For a single observation, such propagation should account for any correlation between different terms within the measurement model. For higher level applications, it is also important to determine error correlation structures for the different quantity values.

There are two main approaches for propagating uncertainties: the Monte Carlo method (MC) and the Law of Propagation of Uncertainties (LPU). These two approaches have been discussed in the Metrology Document, section 2.2.3. The CoMET tools provide software for propagating uncertainties with both methods. Processes such as retrieval algorithms that solve inverse problems using generalised least squares, Bayesian methods and machine learning approaches, usually provide uncertainty estimates directly and thus do not require MC and LPU analyses.

2.5. Step 5: Documenting for different purposes

There are three types of application of an environmental dataset, and they all have different requirements, and therefore different ways in which uncertainties and covariance information needs to be provided.

Operational data are provided for near-real time applications. They inform weather models, operational decisions (e.g., sea surface temperature informing fishing fleets) and other short-term applications. For such applications, timeliness and consistency are extremely important. Uncertainties are often less critical and are usually indicated as 'noise' (random effects) and 'bias' (systematic effects) uncertainties given at a mission level, perhaps with additional 'quality flags' to distinguish 'good' data from 'poor data'. For operational applications, covariance information is rarely calculated or immediately useful.

Research data are provided for other scientists to use for a wide range of applications that are of value today. Research data use a combination of near-real time (not quite as real time as operational data, often with a few days delay to be able to bring in all available information) data and historical data that has been reprocessed. The data are provided for applications at higher levels and uncertainty and covariance information should ideally be provided, but in simplified format. Someone who is working with level 2 data to generate level 3 products, should not need to understand every detail of the level 0 to level 1 processing, but should be given summary error correlation information from earlier levels.

For research data, the uncertainty information needs to be provided in a way that is 'as simple as possible but no simpler'. This usually means combining different sources of uncertainty, for example to provide uncertainties associated with systematic effects, uncertainties associated with random effects and uncertainties associated with partially correlated effects. Or providing information about correlation structures in different dimensions.

Long term data preservation (LTDP) is the record of data for future scientists. Just as today's scientists are reprocessing data that was taken 20 or 30 years ago (by satellites), or 100 or 200 years ago (for in situ observations), it is reasonable to assume that scientists in the future will reprocess and reanalyse today's data. LTDP is the process to ensure the data are available to them (stored on media and in formats they will still be able to access) and that they have all the associated metadata and documentation with them for future scientists to be able to reproduce what has been done, prior to their own reanalyses of that data.

For LTDP it is important to store all information about every source of uncertainty (the full effects tables) for LTDP and to include information about the maturity of the analysis that went into evaluating each source of uncertainty.

3. Putting it all into practice

Establishing an uncertainty budget for an FRM or FDR or TDP requires dedicated effort. The steps presented in this document, along with the theory explained in the Metrology Document, provide a framework for approaching that effort in a systematic way. The QA4EO website also has access to training material, and case study examples, as well as to the CoMET toolkit that can simplify the analysis. The material is under development and will be expanded in the future.

A. Appendix on diagrams

A.1. Introduction

Diagrams are used as visual tools to aid thinking about a measurement or analysis process and to simplify the identification of the sources of uncertainty. In principle, many types of diagram can be used for such a role. However, bringing a consistency in style to the diagrams in different projects and communities can aid with communication. Therefore, these light ‘rules’ are given as guidelines. If it is necessary to alter these ‘rules’ to simplify presentation, that is of course acceptable. Note that the ‘examples’ here and in the main text have not always perfectly followed these rules either – we have collated real diagrams from the communities that have been developed in parallel with any rules.

A.2. Guidelines for uncertainty tree diagrams (from FIDUCEO)

The “Uncertainty Tree Diagram” takes the form shown in Figure 5. The uncertainty tree diagram captures the measurement function and the structure of the dependencies, together with expressions for the sensitivities and short uncertainty contribution descriptors. The central box contains the measurement function, either written out in full, or written conceptually as a function of input parameters. This should include the “plus zero” term. Some terms in the measurement function are directly provided and have a single source of uncertainty (e.g x_3 in the diagram below). These are shown with the sensitivity coefficient between the term and the uncertainty (descriptor).

Other terms, such as x_1 in the diagram, are directly measured but may be influenced by more than one “effect”, each a separate source of uncertainty. Still others, e.g. x_2, are themselves calculated from other input quantities, which have their own sources of uncertainty. We should also document the uncertainties associated with the “plus zero” (assumptions and approximations) – these are the uncertainties associated with the assumptions implicit in the form of the function.

For some sensors such diagrams become extremely complex. In this case, it may not be possible to provide all information on a single figure. This has been resolved by nesting uncertainty tree diagrams (which could be interactive), where sub-chains are represented separately on separate figures.

The rules are as follows:

- The measurement model should be in the centre of the diagram and surrounded by a rectangular, coloured box (FIDUCEO used light blue).
- All lines should be vertical or horizontal only, right angles are used to change direction.
- Each input quantity of the measurement model leads to a ‘branch’ of the tree. Each branch is a different colour and starts from the input quantity highlighted in the same colour.
- The sensitivity coefficients are placed in each branch to show how the uncertainty in the input quantity is translated into an uncertainty in the output quantity. Sensitivity coefficients can be written as partial derivatives, and are in black, rounded-corner boxes within the branch. The chain rule can be assumed, so Effect 2.a.1 relates to the measurand through $\frac{\partial y}{\partial x_2}\frac{\partial x_2}{\partial \chi_1}$.
- Sources of uncertainty (effects) are written on leaves (above simple horizontal lines) connected to the term that these effects affect.

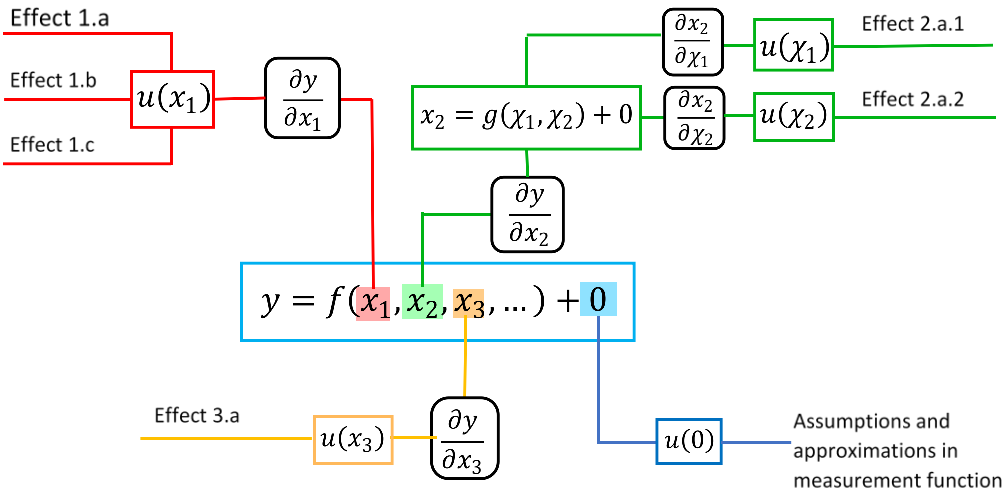


Figure 5 Conceptual Uncertainty Tree Diagram

A.3. Processing diagrams (from QA4ECV and GAIA-CLIM)

The uncertainty tree diagram is a useful approach to documenting a process that involves a single equation. Where corrections must be performed in a specified sequence, then a "processing chain diagram" is more appropriate. The concept of the processing chain diagram has been commonly used in many ATBDs for a long time. These specific rules were developed in the QA4ECV project and improved in the GAIA-CLIM project, where processing diagrams were developed for a broad range of non-satellite (FRM or proto-FRM) observation systems.

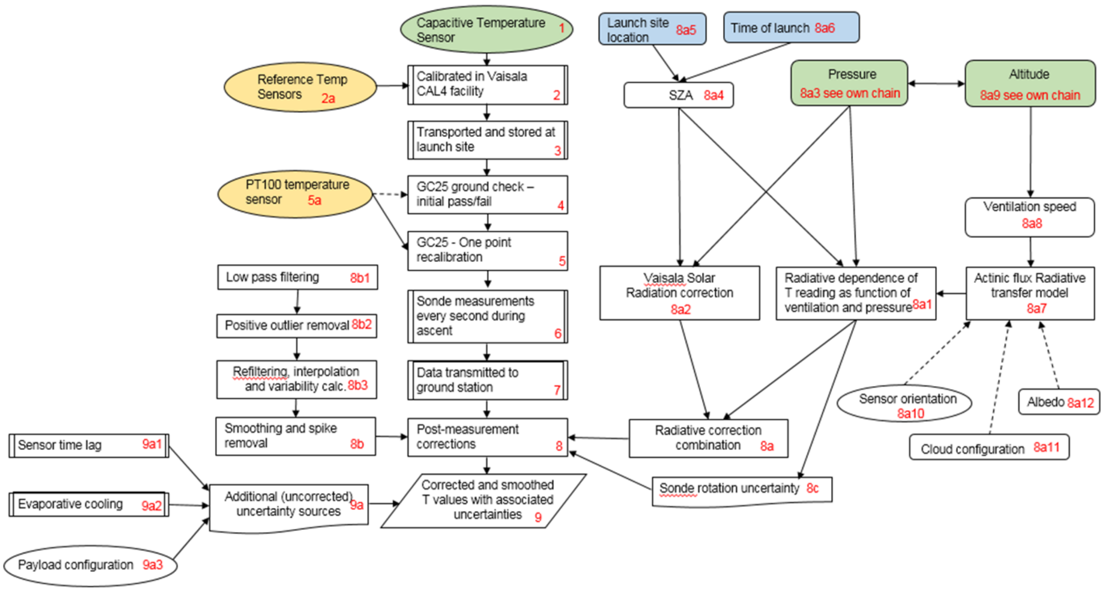
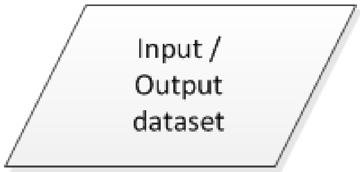
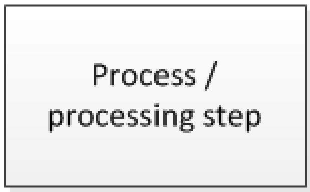
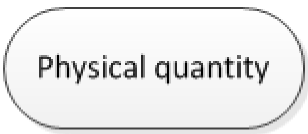
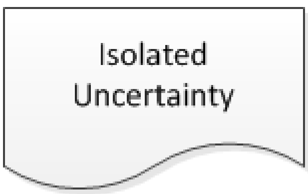


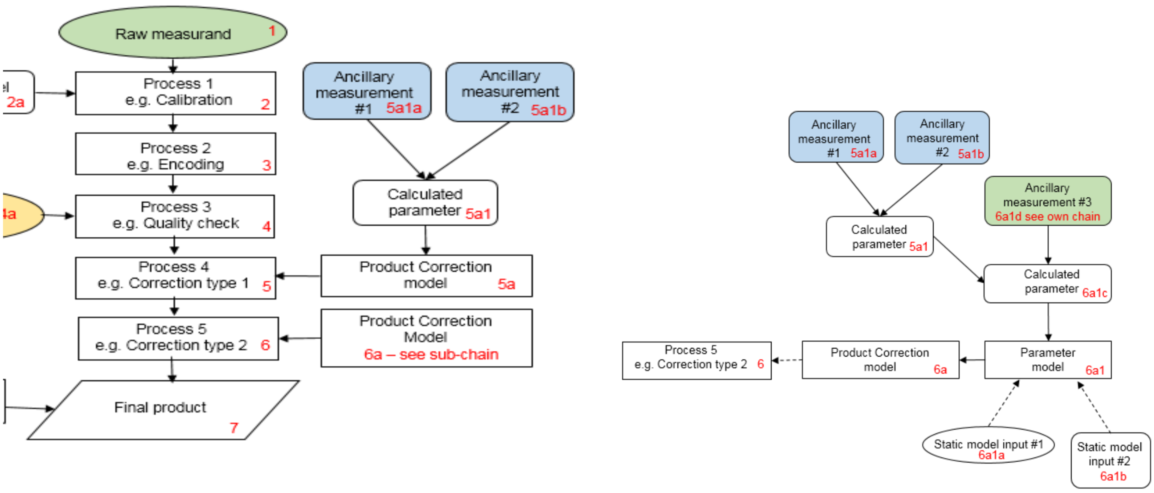
Figure 2. Example process chain for a RS92 radiosonde temperature measurement.

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 2 (An example diagram for a RS92 radiosonde temperature product). Guidance on the types of boxes for each type of chain element is given at Table 1. However, it is noted that the underlying process flow information is the important content, so excessive effort should not be spent in formatting the diagrams.

Table 3.1. Traceability Chain Shapes and Definitions

	Parallelogram	A dataset visible to the user, e.g. initial input, final output product or any intermediate product that is available to the user. Input datasets have been provided from external (auxiliary) sources.
	Rectangle	A process within the chain, used to describe a transformation in the dataset that may or may not have an associated uncertainty. The default box shape. The dataflow within the process is typically invisible to the user.
	Rectangle with side-bars	Essentially identical to the process rectangle. However, sometimes used to represent a sub-chain or major processing block where more granular information is available.
	Ellipse	Name of the instrument, reference material or measurement device from which raw data is obtained. The raw data can also include the data propagated from a previous level. This differs from 'input dataset' in that the instrument is considered part of the main process.
	Rounded rectangle	An ancillary physical quantity dataset or product necessary in the processing chain or to give context to the product.
	Rectangle with wavy bottom	An uncertainty quantity not associated with (isolated from) an element in the traceability chain. Typically used to represent assumptions and known effects that are not directly corrected for (i.e. effects that become part of the +0 term).
	Rhombus	A decision step that may affect whether specific data appears in the output product. Such decisions may impact the probability distribution function of the uncertainty.

If there is a complex sub process, this can be separated out, with an example shown in Figure 3. Arrows in these diagrams represent the direction of the process.



(a)

(b)

Figure 3 (a) extract from a processing chain which includes a reference (6a) to a sub chain, and (b) sub chain for that process.

A.4. Other diagrams

There are no existing formal rules for metrological traceability diagrams or derivation diagrams.

B. Appendix on effects tables

B.1. Introduction

An effects table summarises all the important information about a source of uncertainty (an ‘effect’). Each effect identified in the diagrams, should have an associated effects table (or row or column in an effects table that combines multiple effects). Each effects table should cover all the parameters relating to the sources of uncertainty that were identified in Section 2.3.

An example effects table is given as Table 3.2. The names of the rows would be updated depending on the application, as discussed in the subsections below.

Table 3.2 Example blank effects table

Table descriptor			
Name of effect		Effect name 1.1	Effect name 1.2
Effect identifier		1.1	1.2
Affected term in measurement function			
Maturity of analysis	Maturity of uncertainty estimate		
	Maturity of correlation scale estimate		
	For low maturity, is effect negligible?		
Correlation type and form	Within dimension 1		
	Within dimension 2		
	Within dimension 3		
	Within dimension 4		
Correlation scale	Within dimension 1		
	Within dimension 2		
	Within dimension 3		
	Within dimension 4		
Uncertainty	PDF shape		
	units		
	magnitude		
Sensitivity coefficient			

B.2. Effects table maturity

A robust metrological review of every source of uncertainty may not always be possible, either because information is not available or because project timescales require prioritisations to be made. It is helpful to identify the maturity of analysis so that readers can interpret the analysis. A four-point scale is suggested, based on expert, but qualitative, judgement on the maturity of the evaluation of the uncertainty magnitude, of the error-correlation scale and form, and with an impact statement.

Maturity of analysis	Maturity of uncertainty estimate	0 – Effect identified; no quantification 1 – Estimates only 2 – Some analysis performed to evaluate 3 – Rigorous analysis performed
	Maturity of correlation scale estimate	0 – Not done 1 – Estimated 2 – Scale based on analysis, unsure about correlation shape 3 – Strong evidence for correlation scale and shape
	For low maturity, is effect negligible?	Negligible, Minor or Significant? (or unknown) (Preferably with explanation or evidence) This box allows readers to determine whether a higher maturity would be preferred – it is not worth putting a lot of effort into raising the maturity level of a negligible effect

B.3. Effects table correlation rows

B.3.1. Choice of row headers

The error correlation needs to be considered along the appropriate dimensions. In table 3.2, these were described as "within dimension 1", "within dimension 2" etc.

For the low-earth-orbiting radiometric sensors considered in FIDUCEO, these dimensions are cross track (pixel-to-pixel), along track (scanline to scanline), orbit-to-orbit (which acts as a temporal dimension) and, separately, between spectral bands, thus providing the dimensions: cross track pixel-to-pixel, along-track scanline-to-scanline, temporal orbit-to-orbit and spectral frequency-to-frequency.

For a geostationary radiometric sensor, the dimensions are the same, but instead of an orbit representing the temporal dimension, the temporal dimension is image-to-image. For a radar altimeter, the dimensions could be fast time (within a waveform), slow time (all longer timescales) and spatial (for geophysical corrections). For an in-situ measurement network, they could include instrument-to-instrument, time-to-time and perhaps other dimensions relating to how the measurements are made (e.g. spectral band / viewing angle).

The same dimensions should be used for all effects tables for a particular measurement equation. For the FDR of LEO radiometric sensors, the following table form has been used:

Correlation form	Pixel-to-pixel [pixels]	
	from scanline to scanline [scanlines]	
	Between orbits [orbit]	
	Over time [time]	

Note that the 'between orbits' and 'over time' options could be merged, or kept separate to account for short-term vs. longer term changes.

These dimension descriptors act as the row headers. What is filled in for each row is the correlation 'form' – one of options given in the next section – for that dimension. The same row headers are also used for the correlation scale, where numbers are given to parametrise the correlation form (i.e., if the triangular correlation form is chosen, the 'scale' gives the half base width size). In FIDUCEO, the spectral correlation was handled separately from other types of correlation, through an error correlation coefficient matrix from one channel to the next. This is because in the instruments considered in FIDUCEO, the spectral error correlation was more complex than the spatial error correlation and could not easily be described by standard forms. Also, the spatial dimension was conceptually enormous (very large number of observations in a mission), while the spectral dimension was limited to the number of spectral channels on the instrument. Therefore, for the FIDUCEO instruments, spectral correlation was handled separately. It is important to develop an effects table appropriate to the design of the instrument, and this will vary from example to example.

Channels/ bands	List of channels / bands affected
	Error correlation coefficient matrix

For the TDP, the following dimensions could be included:

Correlation type and form	From level 1

Larger scale temporal
[time]
Larger scale spatial
[geospatial coordinates]

The exact structure and layout of these rows of the effects table will therefore depend on the dimensions of interest. It may also be appropriate to include rows to describe how one quantity in the measurement model has an error correlation with another.

B.3.2. Error correlation forms and scales

The FIDUCEO project defined the following correlation forms. Each correlation form is described in the "correlation type and form" row by one of these names. In the "correlation scale" row, it can then be parametrised by giving quantities to the parameters given here. Other correlation forms may exist and should be defined within documentation. The CoMet toolkit allows correlation forms to be pre-defined and correlation parameters to be stored and used in uncertainty analysis.

Table 4 Parameters defined for different correlation forms

random	none required	For fully random effects there is no correlation with any other pixel
rectangle_absolute	$[-a, +b]$ (rectangle limits). Provide these per pixel/scan cycle/orbit as required. Allow for a way of representing $[-\infty, +\infty]$ to represent fully systematic for all observations in that dimension.	An effect is systematic within a range and different outside that range. For each pixel / scan cycle / orbit in range say number of pixels / etc either side that it shares a correlation with. For fully systematic effects notation to say "systematic with all".
triangle_relative	$[n]$ – number of pixels/scan cycles being averaged in simple rolling average (should be an odd number)	Suitable for rolling averages over a window from $(-n-1)/2$ to $(+n-1)/2$ (i.e. for n pixels/scan cycles being averaged) Assumes a simple mean, not a weighted mean.

bell_shaped_relative	$[n]$ – number of pixels being averaged in a weighted rolling average, from which truncation range and standard deviation for Gaussian representation follow (truncation beyond $\pm n$ pixels, $\sigma = \frac{(n/2-1)}{\sqrt{3}}$ (n should be odd)	Suitable for rolling averages over a window from $(-n-1)/2$ to $(+n-1)/2$ (i.e. for n pixels/scan cycles being averaged). Assumes a weighted mean, for any weights (and thus also includes things like spline fitting). Also suitable for anything else where the assumption is that "closer pixels/scan cycles are more correlated than further pixels".
repeating_rectangles	$[-a, +b, r_{\max}, L, h, i_{\max}]$ per pixel/scan cycle/orbit etc ($(r_{\max}, L, h)$ will be same for different pixels)	Correlation coefficient assumed to be rmax for pixels/scan cycles from $-a$ to $+b$, and h for pixels/scan cycles from $L-a$ to $L+b$ and from $2L-a$ to $2L+b$ and so on ($iL-a$ to $iL+b$) for all integers i up to $i_{\max}$.
repeating_bell-shapes	$[n, \sigma, L, h, i_{\max}]$	Correlation coefficient assumed to drop off as a truncated Gaussian for local pixels/scan cycles etc in the range defined by n and a similar Gaussian with a peak of h and the same width for pixels/scan cycles iL pixels apart on either side, for all integers i up to $i_{\max}$.
Stepped_triangle_absolute	$[-a, +b, n]$ per pixel/scan cycle/orbit etc (n will be same for different pixels)	The step is a rectangular absolute from $-a$ to $+b$ with a correlation coefficient of one, after which the correlation coefficients drops for another $a+b+1$ lines, and then again. n is the number of calibration windows averaged.
Exponential_decay	$[\ell]$	ℓ : Length scale of exponential decay.
Provided_by_pixel	[vector of relative correlation]	

B.3.3. The FIDUCEO approach to spectral correlation

Spectral correlation (from spectral band to spectral band) can be dealt with as in any other dimension. In FIDUCEO, it was considered practically easier to provide that spectral error correlation as an error correlation matrix directly.

B.4. Uncertainty and sensitivity coefficient

The uncertainty rows describe the shape, units and magnitude of the uncertainty and an expression for calculating the sensitivity coefficient. In the supporting documentation, some evidence is required to explain the origin of the values given here.

Uncertainty	PDF shape	Functional form of estimated error distribution for the term	
	units	Units in which PDF shape is expressed (units of term, or can be as percentage etc)	See comment below where uncertainty and sensitivity cannot be separated
	magnitude	Value(s) or parameterisation estimating width of PDF	

Sensitivity coefficient		Value, equation or parameterisation of sensitivity of measurand to term	Where the uncertainty and sensitivity coefficient cannot be separated the sensitivity coefficient should be one and the uncertainty is in units of the measurand.

The uncertainty (“magnitude” row) is the parameter that characterises the dispersion (standard deviation) of values that could be attributed to the measurand based on the measurement. It is always a standard uncertainty (one standard deviation, and never an expanded uncertainty, e.g. for k = 2). The uncertainty will usually have the units of the underlying input quantity to the measurement model, although in some cases uncertainties may be expressed relative terms as a percentage or may be expressed in multiples of the measurement (e.g., uncertainty expressed in mK for quantities in K).

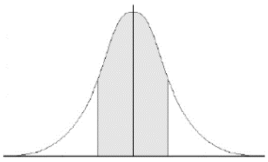
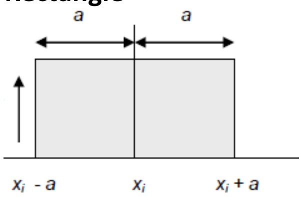
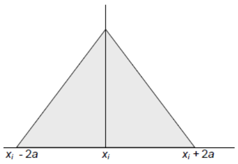
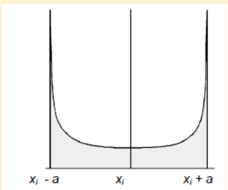
The sensitivity coefficient translates the uncertainty associated with the effect, in the units given in the “uncertainty units” row, into an uncertainty associated with the measurand in the units of the measurand. Such a calculation should consider any translation to the units of the input quantity (e.g., for a relative uncertainty by multiplying by the effect value), and the translation from the input quantity to the output quantity. The Law of Propagation of Uncertainty calculates the sensitivity coefficient as the partial derivative of the measurement function with respect to the term that this uncertainty applies to, $\partial f/\partial x_i$. For uncertainty effects that are shown on the uncertainty tree diagram as a chained series of calculation (multistage measurement model), the sensitivity coefficient is calculated from the chain rule, e.g., $\frac{\partial f}{\partial x_i}\cdot\frac{\partial x_i}{\partial \xi_j}$. The sensitivity coefficient, and any other unit conversion should be written as an analytical expression, or as the value of a calculation of it, in the effects table.

Where the sensitivity coefficient cannot be evaluated analytically, because, for example, the analysis is performed through an iterative software process rather than an analytical expression, it may be evaluated numerically, for example via Monte Carlo Methods. In this case, the uncertainty may be expressed in units of the measurand (as the effect it has on the measurand) and the sensitivity coefficient is 1.

The PDF shape will be one of a defined list of shapes given in Table 3. The actual PDF may not fit perfectly to one of these shapes, but they are likely to be sufficiently close to most actual PDFs, otherwise use the ‘Other’ option.

Table 3 describes common PDF shapes and what the standard uncertainty (the value in “magnitude” under uncertainty in the Effects tables) refers to. Note that these are all for symmetrical PDF shapes. For non-symmetrical shapes, a Monte Carlo analysis is strongly recommended.

Table 3 Parameters defined for different PDFs. For an explanation of these standard uncertainty values, see the GUM section 4.4.

Gaussian 	$u = \sigma$ Uncertainty is the standard deviation	Be careful when using published literature, or a calibration certificate, to provide u. If an expanded uncertainty is quoted, then it's important to divide by k (often $k = 2$ in certificates).
Digitised_Gaussian In early satellite sensors and for some modern applications, the signal is heavily digitised (to reduce the information quantity to downlink). The PDF is therefore a digitised version of a Gaussian, with discrete levels rather than a smooth function	Unknown – treat as Gaussian	The most appropriate standard uncertainty for a digitised Gaussian has not been fully evaluated. Please treat as a Gaussian, but keep this option open for the future
Rectangle 	$u = a/\sqrt{3}$ where a is the half width	Useful for when we know a quantity must be in a range $\pm a$, but it's equally likely to be anywhere in that range, e.g. digitisation
Triangular 	$u = a/\sqrt{6}$ where a is the half base	Useful for where we know there is a range a quantity is in but it's more likely to be in the middle of that range (e.g. when a quantity is the difference between two digitised values)
U-distribution 	$u = a/\sqrt{2}$ where a is the half base	Useful for where we know there is a range a quantity is in but it's more likely to be at the edges of that range (e.g. where there is a feedback loop that switches on and off and encourages drift to the two ends of a temperature range)

B.4.1. Evaluating the uncertainty

There are many ways to do the uncertainty evaluation and the choice will depend on the nature of the uncertainty and the available information. These generally fall into one of the following methods:

- **Provided uncertainties** – if a calibration coefficient is determined through harmonisation or through pre-flight laboratory-based calibration, an uncertainty should be provided with the quantity. It is important to consider the provenance of this uncertainty statement. If it has been rigorously analysed with a "fiducial" QA4EO-compliant method, or is audited to ISO 17025, then it is likely to be directly useable. If it is based on a less rigorous analysis it may be appropriate to review the uncertainty calculation independently (where information is available).

- **Noise estimates** – one of the challenges in EO is that, because the scene is changing, the signal is varying all the time and therefore laboratory approaches of making repeat measurements of a stable source are not possible. However, most satellite sensors have some information about noise performance, from, e.g., for radiometric sensors, a stable scene, onboard calibrator, or deep space views. The Allan deviation can be useful here. For active sensors, noise information is available from the repeatability of the individual waveforms that are averaged to give the final waveform (these are provided in some instances) and from onboard calibration modes.
- **Modelling processes** - sometimes it is possible to estimate the scale of a particular source of uncertainty by modelling the processes on board. In the FIDUCEO project this was done for example for the AVHRR onboard calibration target, where thermal gradients caused by direct solar heating were modelled based on a physical model of the instrument and the available information.
- **Comparison to a reference** - there are occasions when an independent reference measurement is available [e.g., in-situ data], and comparisons to that reference can be used to evaluate the uncertainty. This has been a common method in Earth Observation to evaluate measurement uncertainties and is sometimes the only option. Care needs to be taken to consider the uncertainty associated with the reference, and it is better if this comparison is performed on specific input parameters and not on the resultant measurand. In addition, consideration of the collocation uncertainty – due to any spatial and temporal mismatch between the two measurements should form a part of any comparison exercise.