METROLOGICAL APPROACH FOR EO

GENERAL GUIDANCE ON A METROLOGICAL APPROACH TO FUNDAMENTAL DATA RECORDS (FDR), THEMATIC DATA PRODUCTS (TDPS) AND FIDUCIAL REFERENCE MEASUREMENTS (FRMS) – EXECUTIVE SUMMARY

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) – Executive Summary

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

Version control

Issue	Date	Authors	Reviewed by	Notes
1.0	19.5.22	As above	Authors	First issue
1.1	21.10.24	As above	Authors	Reviewed and updated
1.2	20.6.25	As above	Authors	Reviewed and updated

Project Acknowledgement

IDEAS-QA4EO

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.

Table of contents

1 Introduction

Environmental observations from space-borne, air-borne, ground-based and sea-borne measurements provide essential information about the state of the environment and how it is changing. The information derived from such observations is used by service providers, commercial organisations, public authorities, and national and international organisations to support social and economic development, and to respond to global and local challenges around natural hazards, climate and biodiversity changes, and for food, water and energy security.

The Group on Earth Observations (GEO) was set up over the first three World Summits on Sustainable Development at the start of the 21st Century. GEO's vision, given in its 2016-2025 strategy, is a future wherein "decisions and actions taken for the benefit of humankind are informed by coordinated, comprehensive and sustained Earth observations" through a GEO System of Systems (GEOSS). GEOSS links the many different Earth observing systems, whether orbital or suborbital, from governments and private organisations, into a single interconnected system that facilitates the sharing of data. GEOSS ensures "that these data are accessible, of identified quality and provenance, and interoperable to support the development of tools and the delivery of information services".

The term 'interoperable' covers several concepts all relating to the characteristic of a product or system to work with other product or systems. Interoperable datasets can be combined and compared to obtain broader information than individual datasets. This interoperability requires that data formats are clear and based on common, well-defined concepts. It requires common geographical grids and formats. Additionally, it requires observations to be based on a common measurement scale, referenced relative to a well-understood reference. The International System of Units (SI) provides such a reference and metrological principles such as traceability to SI, documented uncertainty analysis and comparisons between different measurement references, can support the interoperability of Earth observations. It is for this reason that GEO, and its space arm, the Committee on Earth Observation Satellites (CEOS), endorsed in 2010 the Quality Assurance Framework for Earth Observation (QA4EO).

This set of reports provides a theoretical basis and a practical guide to the implementation of the principles of QA4EO, as set out in the QA4EO guidance documents, along with case study examples of the implementation of QA4EO principles.

2 QA4EO and Metrology

QA4EO realises the following principle regarding Earth Observation data quality:

> 'It is critical that data and derived products are easily accessible in an open manner and have an associated indicator of quality traceable to reference standards (preferably SI) so users can assess suitability for their applications i.e. 'fitness for purpose'.'

QA4EO defines high level processes to achieve these objectives, such as may be achieved through well-documented procedures, participation in comparisons, and uncertainty assessments, applicable to all EO data records. 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'. The QA4EO principle stops short of requiring SI-traceability in all circumstances, recognising that the full rigour of linkage to SI may not be viable for all applications and measurements, however, the accompanying guidelines are based on metrological concepts adapted from guidelines of the international metrology community and a metrological approach is strongly implied.

Metrology, the science of measurement, is the discipline responsible for maintaining the SI and the associated system of measurement. The SI ensures that measurements are stable over centuries and that measurement standards are equivalent worldwide. These properties are achieved through the key principles of metrological traceability: uncertainty analysis and comparison.

Since the endorsement of QA4EO, several collaborations between metrologists and the Earth observation community developed detailed examples of how these principles can be applied in practice. Such projects addressed satellite observations (radiometric and active sensors), higher level products derived from satellite observations, and in situ (non-satellite) observations used to calibrate or validate satellite observations and their derived products. The various projects developed tools to document traceability, to perform uncertainty analysis, and to provide summary information on data quality so users can judge fitness for purpose of a data set. A list of projects is available at qa4eo.org/about#projects.

This set of documents describes the consolidated outcome of all these projects. It is presented here as an executive summary, as well as in greater depth in accompanying documents that outline how to apply these methods. The QA4EO website also contains case study example implementations and training material.

These documents were prepared by the National Physical Laboratory, the United Kingdom's national metrology institute, which has been involved in these various projects.

3 FDRs, TDPs and FRMs

3.1 Definitions

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 FDR 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.

4 FDRs, TDPs and FRMs: a metrological approach

A core principle of FDRs, TDPs and FRMs is that robust metrological approaches for considering traceability and uncertainty analysis are followed. A pure metrological approach has three core components:

1. Documenting the chain of traceability to SI
2. Systematic approach to uncertainty analysis
3. Use of comparisons to validate uncertainty statements.

The uncertainty analysis should follow the principles of the Joint Committee for Guides in Metrology's[2] (JCGM) Guide to the Expression of Uncertainty in Measurement (the GUM). Note that the term 'the GUM' applies to a suite of documents available on the BIPM website and not just the original document in that set. A GUM approach to uncertainty analysis starts with defining the measurand – that is clarifying which quantity is being measured – along with a measurement model (often an equation) that describes how the relationship between the measurand and all input quantities known

[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 JCGM is a committee of many of the world's standardisation bodies, including ISO, IEC, and is hosted by the International Bureau of Weights and Measures (BIPM)

to be involved in the measurement. The uncertainties associated with each term in the measurement model, and with the form of the measurement model (i.e., the extent to which that equation describes reality) are then separately considered to quantify the sources of uncertainty and any correlation structures (see below). Uncertainties are propagated through the measurement model to evaluate a combined uncertainty associated with the measurand. The GUM approach assumes the measurement model is complete. That is, that all known errors have been corrected as far as possible.

Metrologists validate uncertainty analysis and confirm traceability through comparisons. The Mutual Recognition Arrangement (which enables international consistency of SI-dissemination) requires regular, formal international comparisons conducted under strict rules. The QA4EO guideline 4 describes approaches to comparisons in Earth observation based on those formal metrological comparisons. Comparison exercises are commonly performed in FRM programmes to ensure consistency between different in situ observations; FRMs are also used for comparisons of satellite data sets.

4.1 Uncertainties, errors and correlations

There is a difference between the concept of 'uncertainty' and the concept of 'error'. Uncertainty describes the spread of values around the measured value that can be ascribed to the measurand (conceptually the 'true value'). The error is the difference between the true value and the measured value. Known errors can and should be corrected prior to uncertainty analysis. Unknown errors are the consequence of measurement uncertainty.

Although the value of an unknown error cannot be known, it is usually possible to make statements about whether the error in any one measurement is correlated with the error in another measurement. Such correlations arise from these measurements having the same underlying principle or being determined from the same underlying input quantities. For example, if a calibration is performed once every N measurements, then any error in that calibration will be common to all N measured values that use that calibration in their processing. Similarly, if a slope and offset are jointly determined by fitting a straight line to the same set of measured values, then there will be an inherent correlation between the errors in the slope and the offset.

Environmental observations are affected by both instrument uncertainty and natural geophysical variability. Therefore, it is necessary to distinguish correlations due to underlying physical processes from correlations due to measurement errors. Therefore, the term 'error correlation' is recommended. It is 'error correlation' because it is the unknown errors that are correlated. The 'uncertainty' cannot be 'correlated' (or 'random' or 'systematic') because the word 'uncertainty' simply describes the distribution spread and is always positive.

4.2 Propagation of uncertainties and Earth observation processing levels

Two aspects of Earth observation data mean that error correlations must be considered differently from most laboratory measurements. First, data are rarely used in isolation – there is no single comparison between a satellite FDR and a single FRM. Instead, series of comparisons are made over long periods, often in different locations and between different instruments. Observations in FDRs are also combined in spatial averages and compared in long time series. Second, data are often processed through different 'levels' with different engineers and scientists involved in each level's processing. In order to assess uncertainties robustly at higher levels, it is important that sufficient information is shared from lower levels. That means that ideally uncertainty analysis begins at level 0, and is propagated to level 1, then level 2 and so on[3]. The information needs to be sufficient, and yet also provided in a way that does not require higher level processing to understand all the complexity of lower levels.

Therefore, at each level of processing, scientists need to think about error correlations not only between the input quantities in their own measurement model(s), but also in the different output quantities being determined. Each measurement will have measured values at different times and/or locations, and some measurements will also have measured values at different wavelengths/frequencies or perhaps at different viewing angles. The time, location, wavelength, and angle represent different 'dimensions' along which measured values may have correlated or uncorrelated errors. Other 'dimensions' may be appropriate – for example, there may be common (correlated) errors in measured values made by a single type of instrument in an FRM network that has a few types of instruments. In such cases, instrument type would be a 'dimension' for error correlation.

4.3 Long-term data preservation and more immediate applications

Data sets are usually prepared for three types of application: operational use, research, and long-term data preservation. Operational use involves near-real-time data sets produced for time-critical applications (the meaning of 'near real time' will depend on the application and is usually given in the range of a few hours). Such applications often do not need robust uncertainty analysis and simple summary uncertainty information suffices.

The main application of data sets happens with a non-time critical research product. Here, different engineers and scientists specialise in a particular processing level (e.g. level 0 to level 1 is often performed at the Space Agency, while higher level processing is done in academic communities). These processes differ from near real time data in that additional ancillary information can be included and there is some quality control on data (e.g., flagging outliers). Higher level processing can be done either operationally (usually producing datasets that lag real time data by a few days) or for a particular scientific study (where data are calculated for a chosen time period). Such applications may reprocess historical data (even data a decade or longer old) as part of a long time series, but they differ from 'long term data preservation' applications, in that they build on a lower data level produced by other scientists or engineers. These research applications need datasets from lower levels that are provided in a manageable form (usually summarised in some way) both to ensure manageable data volumes, and because the higher-level scientists are not expected to understand all the details of the lower levels. Thus, information about error correlation structures is valuable but should be provided in a summarised and perhaps simplified form to enable the most useful information to be shared efficiently.

Long term data preservation is about storing operational records, and reprocessed research datasets for applications several decades in the future, where scientists may reprocess historical datasets – often working back from the rawest (level 0) data and reviewing all the different levels – to consider new scientific insight. This means that along with the data and uncertainties, all the information

[3] In satellite Earth observation the concept of "levels" is widely used. Level 0 represents the rawest quantity downlinked from the satellite (often a signal in 'counts'). Level 1 processing converts these 'counts' into a physical quantity (e.g. top-of-atmosphere radiance or an altimeter power waveform). Level 2 processing converts the level 1 product into a geophysical quantity (e.g. ground reflectance, surface temperature, retrieved atmospheric composition or altimeter ranges and sea state conditions). Higher level processing performs further geophysical transformations (e.g. to leaf area index) and/or performs spatial and temporal averaging (regridding).

necessary to understand, and reproduce, the development of the summary data sets need to be stored. Long term data preservation requires the recording of all relevant information and decisions, including information about how uncertainties and values were derived. Such information can lead to enormous data volumes, and some compromise between completeness and data volume size is often needed.

4.4 A pragmatic approach

The definitions of FDRs, FRMs and TDPs given above can be daunting, and the strict uncertainty analysis described in these guidelines may not be possible to be applied. This can be because of practical issues – such as limited resources, expertise, or funding, or for scientific or engineering reasons – a source of uncertainty is not sufficiently well understood or accessible from the measurements that are taken, or perhaps relevant information is simply not provided by lower-level processes. The guidelines here provide an ideal approach, but practically, real FRMs, FDRs or TDPs may not be able to complete a perfect uncertainty analysis. Sometimes comparisons will be used to estimate uncertainties, or the combination of several sources of uncertainty, rather than to validate independent uncertainty analysis. Sometimes a likely or maximum uncertainty value is estimated by expert judgement. If such decisions are well documented, particularly in documentation aimed at long-term data preservation purposes, it is acceptable to 'be pragmatic' about uncertainty analysis. Datasets should, however, have statements of what assumptions and approximations have been made and should document, perhaps through a 'maturity' statement, which aspects have been evaluated in detail, and which have only been estimated. Ideally, such information would be stored in an agreed and formal way to ensure machine readability and ongoing understandability over the long term.

5 Methodology

There are five steps towards a metrological uncertainty analysis. These steps are described in further detail in other documents in this series and tools (document templates, guidelines for diagrams, format specifications and Python code modules) are available to assist with these steps.

5.1 Step 1: Define the measurand and the measurement model

Defining exactly what is being measured and provided in a dataset is often more difficult than it first appears. As an example, consider a thermometer 'measuring' air temperature. The thermometer does not measure temperature per se, but the expansion of a liquid in a tube, or the change in resistance of a wire. Second, the thermometer does not measure air temperature per se – it measures its own temperature, and it is the responsibility of the person designing the observing system to ensure that either the thermometer is in equilibrium with the air temperature of interest, or that differences between those two temperatures can be modelled and corrected for. Finally, the 'air temperature' measured by the thermometer is often assumed to represent something else (e.g., the average temperature over a particular spatial region). This example shows up several different problems in defining a measurand even for something as 'simple' as air temperature measured with a thermometer: the measurand may not be uniquely defined, and it may relate to a more fundamental 'measurement' through a complicated forward, or indeed often inverse model.

Similarly, for a satellite observation, the measured signal, often in 'counts' needs to be converted to a physical quantity (e.g., top-of-atmosphere radiance within a spectral band). Processes such as orthorectification alter the perception of the measurand. To provide a useful definition of a measurand, it is necessary to define (or assess) whether an observation represents an average value within a spatial region defined by, e.g., a pixel, or a peak value within a footprint. When satellite and in situ data are compared, they are likely to measure different things. For example, satellite-based measurements may relate to sea surface temperature as the top micron of the water, measured over a satellite footprint, whereas in-situ measurements may relate to sea surface temperature at a single point at a depth of a few tens of centimetres.

Furthermore, there may be questions of reference – is a range measured relative to the Earth's ellipsoid or to its geoid, for example. At higher levels of processing, where measured values are combined with models, the measurand may be even more difficult to define. However, defining the measurand is important both to describe the dataset to users and to enable clear thinking in the uncertainty budget. Often, rather than considering a single measurand, it is necessary to consider a sequence of linked measurands: the outputs of different processing levels for example. From a GUM perspective, this would be called a 'multi-stage measurement model' with different measurands at each stage. It is necessary to perform uncertainty analysis for each measurand (for each stage of the multi-stage measurement model) and propagate uncertainty between these stages/levels.

The measurement model itself may be able to be written as an equation with an analytical function. Sometimes, however, the analytical expression cannot be written, and it is approximated through numerical processes involving iterations, non-linear fitting or machine learning techniques. Whether or not it can be written as an equation, the processes by which input quantities are combined to determine the measurand, is known as the measurement model. It is important to realise that there will be uncertainties associated with the form of the measurement model (whether the process it describes accurately describes reality) as well as with the input quantities that are used within it.

5.2 Step 2: Establish traceability with a diagram

A visual representation of how a measurement and its traceability is achieved, along with visually representing the different sources of uncertainty, is highly valuable in assessing performance. Diagrams as described below are extremely useful tools to help understand and communicate how a measurand is derived and to consider and share what the sources of uncertainty are. Diagrams show where terms come from and thus highlight sources of uncertainty in input quantities and in the approximations and assumptions inherent in the model.

There are different types of diagrams that can be helpful for different purposes. The guidance documents describe rules for and give examples of such diagrams. In real uncertainty analysis several of these diagrams may be used (and indeed new forms of diagram may be helpful). The 'rules' are there as a guide only; but following them provides a consistency between data products that can help the community's overall understanding.

Processing diagrams are usually given in the form of a flow chart. They describe the different processes that are carried out to derive the measurand and are particularly useful either as an introduction to different steps or levels (that are themselves analysed through other types of diagram), or where it is important for processing steps to be carried out in a specific order. These are commonly used for higher level processing (TDPs rather than FDRs). For a processing diagram to be considered metrological, rather than a standard flow chart, it is important to include sources of uncertainty in the diagram, for example by considering the assumptions and approximations at each step or the origin of and uncertainties associated with auxiliary information introduced at each processing step. (Note the more detailed document in this series: the 'Process document' provides examples of processing diagrams).

Uncertainty tree diagrams are based on equations. The measurement model is written in the centre of the diagram, often with a $+0$ or $+\delta$ term added to account for the recognition that the model itself has uncertainty[4], and each quantity in the measurement model has a 'branch' showing the origin of that term (often further equations). At the edge of the diagram there are 'leaves' which list all the sources of uncertainty. An example uncertainty tree diagram is given in Figure 5.1.

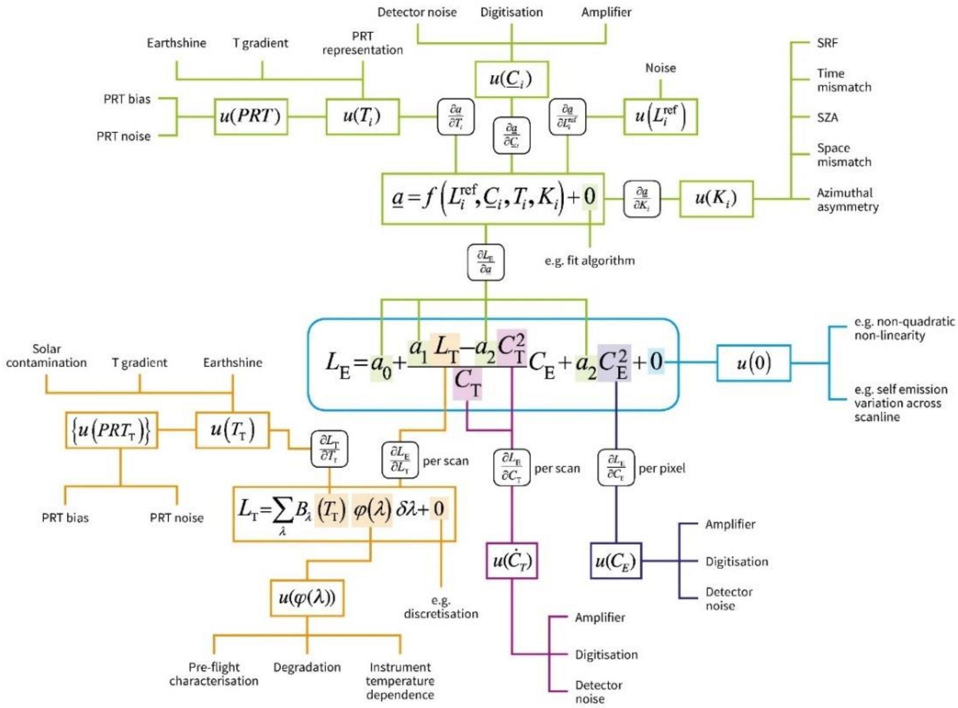

Figure 5.1 The first FIDUCEO uncertainty tree diagram from Mittaz et al 2019. Examples of other types of diagrams are given in the Process Document.

Measurement traceability diagrams are diagrams used to show metrological traceability where instruments are calibrated against other instruments. These use two different types of symbol used to represent instruments and the quantities that the instrument measures. They are used alongside an uncertainty tree diagram to give a more intuitive overview of the measurement process (e.g., amplifier, detector, aperture, rather than equations that give amplifier gain, detector responsivity and aperture area). (Note the more detailed document in this series: the 'Process document' provides examples of such traceability diagrams).

Derivation diagrams are diagrams that show how a measurement model is derived. Such derivation processes can themselves introduce approximations, and derivation diagrams can show up those approximations. A good example of this is the waveform model used to fit the waveform from a satellite altimeter over open ocean. The model is fitted to the measured signals to derive quantities such as the range and significant wave height. The derivation model can highlight the assumptions built into the form of the model – e.g., in this case, that the ocean waves have a particular distribution, and that the antenna pattern can be assumed to be gaussian. The actual fitting of the model would make an uncertainty tree diagram, but the derivation diagram can help define what uncertainties are

[4] The original uncertainty tree diagram used $+0$ as an aide memoire to emphasise that the form of the measurement model itself is uncertain, but that we assume it is correct (that nothing is added). Some scientists find this a helpful aide memoire and that using 'zero' makes a clear point about the nature of the effect and distinguishes it from known corrections. Others prefer the more rigorous mathematical notation of $+\delta$.

associated with the $+0/+\delta$ term in the measurement model. (Note the more detailed document in this series: the 'Process document' provides examples of derivation diagrams).

5.3 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'). There are several things that need to be known about each effect and the FIDUCEO and GAIA-CLIM projects used the concept of an 'effects table' to document, systematically the information that needs to be known about each effect.

The exact rows of an effects table will depend on the application, but there are several common requirements for each source of uncertainty:

- Which quantity in the measurement model it affects
- The magnitude of the uncertainty (value assigned to the standard uncertainty)
- The shape of the probability distribution function for the uncertainty
- How the uncertainty associated with this effect is propagated to the measurand (the sensitivity coefficient)
- The error correlation shape and scale for all 'dimensions' (see section 4.2) that are relevant both for determining the measurand and for subsequent 'higher level' processing or applications that perform averages and/or comparisons.

Additionally, it is valuable to document whether the analysis in the table is mature (based on sound analysis with evidence and validated through independent comparison) or very immature (based on a best guess) or somewhere in between.

The 'effects table' provides a common method for recording what is known about each source of uncertainty. This is valuable to think through the uncertainty analysis and for recording for long term data preservation purposes. Using effects tables that follow the documentary templates and examples given in the guidelines will lead to consistency within the community. The tools available also include ways of storing effects tables digitally.

5.4 Step 4: Calculate the FRM/FDR/TDP and associated uncertainty

The fourth step involves processing the FRM, FDR or TDP through the measurement model and determining the associated uncertainties. It is common for the analysis performed in the previous steps leads to refining the measurement model, and thus how the FRM, TDP or FDR is calculated. The analysis may also improve quality control and verification processes and lead to improved harmonisation of long-term data sets that combine data from different sensors.

The uncertainty may be calculated at the full resolution of the dataset (i.e., for each individual observation at different times, locations, and perhaps wavelengths, angles etc), or processed for some subset of the data, which are then used to parameterise a look-up-table that enables faster processing of the full data set. The choice will depend on processing times and data volumes.

There are two ways of processing uncertainties that are described in the GUM. Uncertainties may be processed using Monte Carlo methods (as in an ensemble analysis), or through the Law of Propagation of Uncertainties (a Taylor expansion linearising the measurement model often recognised as 'the square root of the sum of the squares', although when there is error correlation a full covariance matrix is needed). Monte Carlo can provide better results for non-linear models and is the often the preferred option where the processing cannot be written analytically (e.g., in neural networks or iterative processes), however, it is computationally expensive and does not provide easy access to the importance of different sources of uncertainty. A hybrid approach can use Monte Carlo analysis to evaluate sensitivity coefficients that are propagated through the law of propagation of uncertainties or used in look up tables.

The CoMet toolkit provides a practical Python implementation of the Monte Carlo and Law of Propagation of Uncertainties. The toolkit aims to enable users to easily propagate their uncertainties and abstract away the complexity of dealing with measurement error-covariance information. QA4EO key document 5 provides further details on the CoMet toolkit.

5.5 Step 5: Documenting for different purposes

For operational use, the outcome of the uncertainty analysis may be a simple look up table. For research applications and higher-level processing, simplified information needs to be given, e.g., providing combined uncertainties for (a) effects that lead to uncorrelated (random) errors, (b) effects that lead to fully correlated (systematic) errors and (c) that lead to 'in between' errors that are partially correlated (structured). For long-term data preservation, all information should be stored along with complete documentation.

5.6 Iterating the five steps and uncertainty validation

The 'Five Steps to an Uncertainty Analysis' are not a recipe that needs to be followed only once. They represent a sequence of steps to document a current understanding of the traceability, measurement model and uncertainties in a systematic manner that can be helpful both to the scientists producing the uncertainty analysis and for those using the outcome of it.

However, there is almost always a need to iterate these five steps as new information arises. Such information may come from an improved understanding of the physics of measurement, or an insight from work by others, particularly from those working at an earlier level of processing. It may also arise from the insight gained from comparisons between FRMs, between FRMs and satellite FDRs/TDPs, or between different satellites.

Such comparisons can also support the validation of uncertainty analyses. Further validation can come through peer review and formal auditing of data sets. The systematic presentation of uncertainties that are encouraged by the five-step method can support such peer review and auditing.

Figure 2 shows the CEOS Five Steps within an iterative framework.

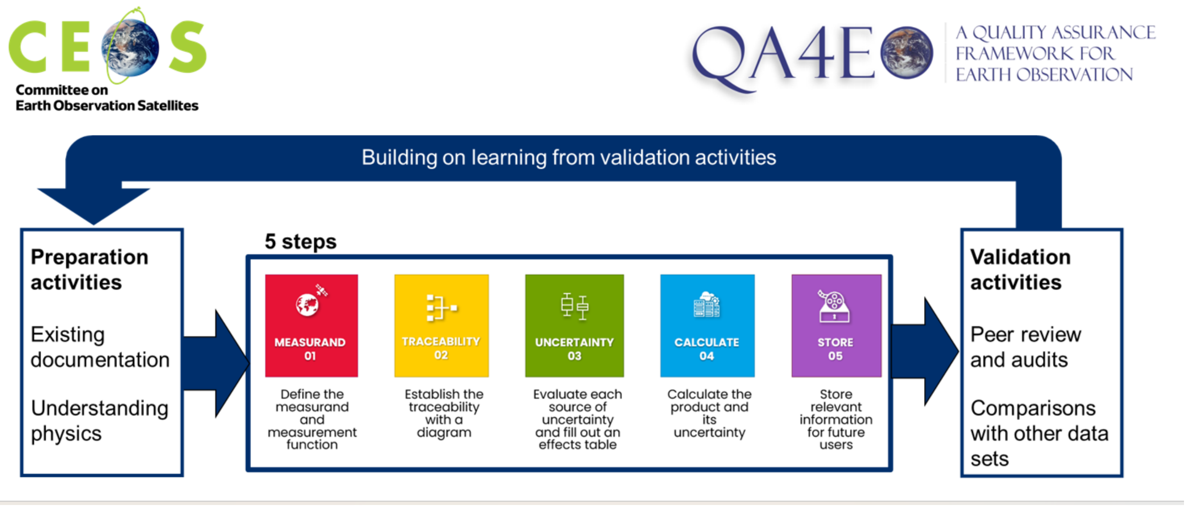

Figure 2 An iterative framework for the CEOS Five Steps