

QA4E  | A QUALITY ASSURANCE
FRAMEWORK FOR
EARTH OBSERVATION

**A guide to “reference standards” in support of
Quality Assurance requirements of GEO**

QA4EO-QAEO-GEN-DQK-003

A guide to “reference standards” in support of Quality Assurance requirements of GEO

| | |
|-----------------------------------|--|
| Author: | Nigel Fox |
| E-mail: | Nigel.Fox@npl.co.uk |
| Editor: | Marie-Claire Greening |
| E-mail: | marie-claire@greeningconsulting.co.uk |
| Issued under Authority of: | QA4EO |
| Issue no: | Version 4.0 |
| Date of issue: | 26 March 2010 |

Summary of Changes since last issue:

Version 3.0 now includes edits and comments made on version 2.0 by GSICS in March 2009.

Version 4.0 incorporates edits and comments made on version 3.0 by WMO and GSICS August – November 2009.

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**

Table of Contents

| | | |
|----------|---|-------------------------------------|
| 1 | ABSTRACT | 3 |
| 2 | SCOPE | 3 |
| 3 | TERMINOLOGY | 5 |
| 4 | INTRODUCTION / CONTEXT | ERROR! BOOKMARK NOT DEFINED. |
| 5 | OUTCOMES | 6 |
| 6 | INPUTS | 7 |
| 7 | STANDARDS AND TRACEABILITY | 7 |
| 8 | ESTABLISHING A REFERENCE STANDARD | 7 |
| 8.1 | INTRODUCTION..... | 7 |
| 8.2 | REFERENCE STANDARD - IDEAL CHARACTERISTICS..... | 9 |
| 8.3 | CHARACTERISING A REFERENCE STANDARD | 9 |
| 8.3.1 | <i>Step 1: Requirement specification / scope of use</i> | 9 |
| 8.3.2 | <i>Step 2: Review existing reference standards</i> | 10 |
| 8.3.3 | <i>Step 3: Identify candidate reference standard</i> | 10 |
| 8.3.4 | <i>Step 4: Sources of uncertainty</i> | 10 |
| 8.3.5 | <i>Characterisation</i> | 10 |
| 8.3.6 | <i>Step 6: Establish operational procedure</i> | 11 |
| 8.4 | USE / REGISTRATION / ENDORSEMENT OF REFERENCE STANDARD | 11 |
| 8.4.1 | <i>Use</i> | 11 |
| 8.4.2 | <i>Registration</i> | 12 |
| 8.4.3 | <i>Endorsement</i> | 12 |
| 9 | CONCLUSION | 13 |
| | ANNEX A: REFERENCE STANDARD DESCRIPTION TEMPLATE | 19 |

A guide to “reference standards” in support of Quality Assurance requirements of
GEO

1 Abstract

This key guideline summarises the processes that should be followed to identify, establish and use a “reference standard” as a means of evaluating performance or compliance for a particular activity as part of an internationally harmonised Quality Assurance (QA) procedure. Such “reference standards” can take many forms, including natural or manmade artefacts or targets and datasets, as dictated by the application. The guideline provides advice to those seeking to find a “reference standard” for a particular application as well as the establishment of new ones. It also provides a template to aid those documenting the characteristics of a “reference standard” and serves as a checklist for those seeking to find one. The guideline sets out the steps needed to register and/or obtain formal endorsement of a “reference standard”. This may be appropriate to facilitate international harmonisation in support of the Group on Earth Observation (GEO)’s Global Earth Observation System of Systems (GEOSS).

2 Scope

A key requirement driving the data quality assurance aspects of the Quality Assurance Framework for Earth Observation (QA4EO) is the need for interoperability and the means to allow all the Group on Earth Observations (GEO)’s stakeholders to be able to readily assess (on receipt) the suitability of a data (or derived) product for their particular application. This requires that all data has associated with it a Quality Indicator (QI), which must be unequivocal and universal in terms of its definition and derivation. In practise there is likely to be a wide range of actual descriptors and terms used (e.g. text or numeric) depending on the specific application or users’ needs. However, all should be based on a statistically derived value and this value should be the result of an assessment of its traceability to an agreed “reference standard” (ideally SI) as propagated through the data processing chain.

This requirement emphasises the critical importance of the “reference standard” in the Quality Assurance (QA) process, i.e., its:

- Definition,
- Scope,
- Characterisation *and*
- Assessment or acceptance of its suitability for a prescribed task.

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**

This key guideline summarises the processes that should be followed to identify, establish and use a “reference standard” as a means of evaluating performance or compliance for a particular activity as part of an internationally harmonised QA procedure. Such “reference standards” can take many forms, including natural or manmade artefacts or targets and datasets as required by the application. They can be relatively simple in nature and / or defined by the user for a particular test. A test may be qualitative in nature or may be more formally defined and subject to detailed calibration as part of a quantitative assessment of performance.

Conceptually, a “reference standard” can be very simple depending on the scope and degree of discrimination required in the task it is being used for. For example, to know that a detector responds to light and dark, a simple light bulb can be used as a reference and this would allow the coarse discrimination required. However, if one wants to determine a threshold level for response, or even a gradation of response as a function of illumination level, then the level of knowledge and characterisation of the “reference standard” increases. All three situations require the user to identify and specify an appropriate “reference standard”. Some applications are more demanding than others and require quantitative characterisation whilst others require formal traceability to SI or be “community agreed”. This document provides guidance on all of these types of “reference standard” but it is necessarily tailored to suit the needs of the more demanding applications. The reader is asked to extract from this guide those details commensurate to their application.

In its purest form, a “reference standard” is a “measurement standard” that provides a direct realisation of a particular SI quantity, usually at a National Metrology Institute (NMI) such as the UK’s National Physical Laboratory (NPL) or the USA’s National Institute of Standards and Technology (NIST). This “primary reference standard” can be used within a traceability chain to establish other “working reference standards”. In some cases, such as those needing the highest accuracy, the instrument itself can be compared directly to the primary standard. The Kilogram is perhaps the most obvious example of a primary “reference standard” as it is currently defined as a physical artefact, i.e. as a “block of metal”, as opposed to any physical process or fundamental constant. However, providing that the quantity (under test) is well defined and the uncertainty of realising the quantity has been fully determined and evaluated by comparison or peer review within the formal governance framework of the meter convention (<http://www.bipm.org/>), any measurement standard (of any description) can be established. This is of particular importance to the Earth Observation (EO) community where there are many situations, particularly in the broad range of measurements that pervade the wider GEO community, that are not easily satisfied by traditional metrological “reference standards”.

A guide to “reference standards” in support of Quality Assurance requirements of GEO

A black body radiator is a good example of a traditional standard that meets the needs of the EO community. Such entities find application as “reference standards” in many fields, e.g., atmospheric chemistry, sea surface temperature, earth radiation budget, etc. In this case, provided the emissivity of the black body is known (measured or calculated) as well as its temperature, its emitted spectral radiance is known. Such a measurement standard can then be used to calibrate another black body or a radiometer directly. In practise, it is not simply the “reference standard” that defines the uncertainty of subsequent calibrated instrumentation but also the method of its use and any interaction with the instrumentation under test.

When considering the choice or use of a “reference standard” to assess traceability and performance, it is not simply the stated uncertainty of the “reference standard” that is important but also its suitability for the particular application. This guideline requires that any “reference standard” must also have an operational procedure associated with it. In an ideal world, NMIs would provide a comprehensive range of “reference standards” and operational procedures to meet the needs of all industrial and technical sectors. In practise this is clearly unachievable, although for most “physical” quantities and for an increasing number of chemical and biological quantities, an hierarchical, fully traceable chain has been established. These traceability chains link specific application needs back to a “measurement service” from an NMI through a series of intermediate, often locally generated, “reference standards” that have been specifically established to facilitate the process.

This key guideline provides advice to those seeking to find a “reference standard” for a particular application as well as the establishment of new ones. It also provides a template to aid those documenting the characteristics of a “reference standard” and serves as a checklist for those seeking to find one. Although not prescribing a detailed characterisation methodology for “reference standards” (given their wide variety), it does indicate the generic principles that should be covered.

3 Terminology

All terms within this document are based on internationally-agreed definitions that are, in many cases, derived directly from formal standardising bodies such as the International Organization for Standardization (ISO). These agreed definitions can be found on the QA4EO website (<http://QA4EO.org/>).

A guide to “reference standards” in support of Quality Assurance requirements of GEO

4 Background and Context

This key guideline is written as part of a set, based on the adoption of existing best practise, to form a Quality Assurance Framework for Earth Observation (QA4EO). The QA4EO was developed to meet the current and aspirational needs of the societal themes of the GEO's GEOSS. It was prepared as a direct response to GEO task DA-06-02 (now DA-09-01-a) to "Develop a GEO data quality assurance strategy, beginning with space-based observations and evaluating expansion to *in situ* observations, taking account of existing work in this arena".

5 Outcomes

This document provides guidance on the characteristics of a particular entity that should be evaluated to assess the suitability of its use as a "reference standard" for a particular application. The result will be a documented template describing the scope of use of a particular "reference standard" (including timescale of validity), its ability (quantitative) to discriminate in terms of performance and evidence to support the above. In terms of evidence, and as appropriate, this may include the traceability chain back to an SI or other internationally-agreed standard, or simply measurement data to support the claimed "performance". The completed template would provide the information necessary for any third party to evaluate the proposed "reference standard's" suitability for a particular application. In some cases, this may lead to a formal endorsement by an organisation accepted by GEO as representative of a "community view". This may subsequently lead to recommendations for its adoption and use by the community as a whole. In addition to the template descriptor of the "reference standard", an operational procedure would also need to accompany its use.

The use of the template contained within this key guideline also serves as a convenient "checklist" for any user of this or associated key guidelines of QA4EO when developing test plans.

**A guide to "reference standards" in support of Quality Assurance requirements of
GEO**

6 Inputs

There are no specific identifiable inputs required by this procedure to enable its outcomes to be achieved. Of course when this guide is used to evaluate a candidate “reference standard”, the inputs will be as required to populate the template.

7 Standards and Traceability

The procedure outlined in this document has no quantitatively assessable outcomes in its own right and there are no appropriate “reference standards” to which traceability should be demonstrated. However, this document will make reference to the use of existing documentary standards (where appropriate) within its guidance, e.g. vocabulary [1], uncertainty analysis [2], and SI traceability [3]. As time progresses, new documentary standards and “best practises” that are applicable to the activities described in this document may be developed and adopted. Therefore, it is recommended that the latest version of this document be reviewed for changes prior to its use.

1. QA4EO endorsed vocabulary: <http://QA4EO.org/>
2. ISO:
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45315.
3. BIPM MRA: <http://www.bipm.org/>

Any comparison carried out following these guidelines will select and make use of appropriate “reference standards” to perform the task that they describe.

8 Establishing a Reference Standard

8.1 Introduction

There are a range of differing requirements for “reference standards”. To determine the:

1. Performance of the item under test to a particular specification,

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**

2. Internal consistency / repeatability of an item in doing a task or between similar items, e.g. in manufacturing production line, or
3. Performance of an item as compared to others doing the similar task, i.e. bias determination.

Not all of the above need to be addressed, but the suitability of a “reference standard” for any one of the three tasks above does not exclude its applicability for any of the others. The discriminating factor between them is often the level of uncertainty required and the level of effort expended in its characterisation. In the context of this document’s structure, there is no real need to discriminate between these “application categories”; any differences will be highlighted in the text at the appropriate point.

As an alternative to the generic definitions above, and perhaps simpler to implement, the following classification is favoured and will be used within this key guideline:

QA4EO interoperability “reference standard”: These are “reference standards” that have been formally endorsed by QA4EO (or another designated body) and are specifically used to evaluate biases between measurements or processes. Such “reference standards” will have undergone significant peer review to ensure that they do not themselves introduce significant unquantified biases when used with differing instruments or techniques. This type of “reference standard” should number at relatively few, although international accessibility and availability will be critical characteristics.

Application-specific performance “reference standards”: Although these “reference standards” can be endorsed by QA4EO (or other), it is not a requirement and in most cases would not be appropriate. However, their selection should be approved by any appropriate responsible body (e.g., funding organisation) prior to use. “Reference standards” of this nature must have characteristics suitable to allow them to be used to evaluate a particular performance characteristic of an instrument or process. This may be with respect to a defined quantity or in terms of repeatability. However, whilst the information it provides may be usable to evaluate differences/biases between instruments or applications, as in the “interoperability reference standard”, its principle objective relates to a single instrument or application.

The practical realisation of a “reference standard” can and will take many forms, ranging from a physical manmade artefact to a natural terrestrial or astronomical target and from an empirical or mathematically constructed dataset to a “law of physics”. The common and discerning feature is that the “reference standard” has been appropriately characterised and documented.

A guide to “reference standards” in support of Quality Assurance requirements of GEO

8.2 Reference Standard - Ideal Characteristics

The ideal “reference standard” should present to the instrument or process under test a characterised input, which is representative of what the “instrument” is intended to view or process. The “reference standard” should ideally enable the provision of a range of values that span the range of those that would be “viewed” by the instrument or process under test when operational. This may lead to a requirement for a number of different “reference standards” to scope the range if the single standard has insufficient flexibility in its characteristics to achieve the full test scope. The key properties of the “reference standard” should be sufficiently well characterised and chosen to provide adequate discrimination of the characteristics of the item under test. The stability of the “reference standard” needs only to meet that required by the application. For example, in some cases stability of a few minutes will be adequate and in others stability of years or decades is desired. It should be noted here that a standard procedure and a description of the observer and conditions of observation may also suffice for some applications, e.g., counting a particular species of animal in a given habitat. In this case this document should be interpreted appropriately with the key requirement still being documentation of the characteristics and reasons for them.

8.3 Characterising a Reference Standard

The following set of steps outlines the generic procedure that should be followed and documented. The template provides a convenient means of doing this. In writing this key guideline, it has been assumed that the reader / user of this document is seeking a “reference standard” for a particular application. However, it can easily be adapted to start from the premise of an “identified potential reference standard” and the process of its documentation.

In selecting a “reference standard” for a particular application, the user can also use this guide as a checklist to confirm that any proposed “reference standard” is suitable for the task, is appropriately characterised and documented and/or identify where any additional checks need to be carried out. Any documentation accompanying a “reference standard” should indicate the temporal validity of any characterisation data and evidence to justify this statement. The template provides a convenient means of doing this.

8.3.1 Step 1: Requirement Specification / Scope of Use

Identify parameter(s) that require evaluation against a “reference standard”. This should include an assessment of the level of performance discrimination required, whether this has a temporal component to it and if so what this is. It should also consider whether the parameter is a relative measure, e.g., spatial uniformity, repeatability, self-consistency,

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**

etc., or whether it requires traceability to a defined quantity. This specification could be derived from a specific customer or application, or it could be based on experience. In many cases it will be “what is achievable” and will be derived from the result of an evaluation of a candidate “reference standard”. In this case, it will be the “cope of use”.

8.3.2 Step 2: Review Existing Reference Standards

Before embarking on the search / development / characterisation of any new “reference standard”, a literature search should be carried out. This should include a review of the QA4EO website (<http://QA4EO.org/>) and any linked community specific portals, e.g., the GEO/CEOS Cal/Val portal (<http://calvalportal.ceos.org/>), to consider what “reference standards” are or have been considered for a particular application and their respective performance specification.

8.3.3 Step 3: Identify Candidate Reference Standard

Following steps 1 and 2, or starting from a candidate “reference standard”, assign an “identifier” to the “reference standard”. If the “reference standard” already exists, this may already have an established “name” and this should then be used. It may simply refer to, e.g., a type of material or a specification, in which case any established generic identifier should be used with something unique to the user’s particular application. If a new “reference standard” is being defined, an appropriate name should be assigned. If this is a method, an appropriate title and reference number should be assigned (note for this type of standard key guideline QA4EO-QAEO-GEN-DQK-002 can be used directly).

8.3.4 Step 4: Sources of Uncertainty

Identify parameters or characteristics that could provide additional sources of uncertainty when using the “reference standard” for a particular application. Where possible and appropriate, establish a simple model to link these parameters to enable the combined sensitivity to the desired primary parameter to be determined.

8.3.5 Step 5: Characterisation

The key characteristics (primary and those identified in step 4: Section 8.3.4) of the “reference standard” and their temporal stability (as needed) must be determined. Where the “reference standard” is being used to characterise instrumentation, it is likely that similar instrumentation will be used for this characterisation process. However, the characterisation instrument will ideally have higher resolution for the parameter under test than will ultimately be required. Any characterisation instrument should itself be calibrated to the level commensurate with its use. In some cases, this may simply be to confirm it is linear when used to measure spatial uniformity or that it is stable when

looking for temporal drift; in others, it will be full traceability to SI units. All of this information should be documented and an uncertainty budget established. This uncertainty budget should contain information on repeatability and reproducibility. QA4EO-QAEO-GEN-DQK-002 can be used to facilitate this process, although the level of detail should be commensurate with the expected usage or criticality of the “reference standard”. For example, one that will be used as a community-approved “reference standard” for interoperability should have relatively comprehensive documentation, whereas something used for a one-off test of an instrument subsystem might be a series of bullets.

8.3.6 Step 6: Establish Operational Procedure

Unless characterised, a “reference standard” is of no use. However, any characterisation will be subject to some level of operational constraints, so an operational procedure must be written to accompany any “reference standard”. This can be quite simple in format, depending on the type of “reference standard” and its criticality within the data processing chain. In addition to the “reference standard’s” own characteristics, the procedure needs to consider its interaction with the item under test and the environment where this is being carried out. An ideal operational procedure will provide guidance on a range of situations rather than have a set of separate documents.

8.4 Use / Registration / Endorsement of Reference Standard

8.4.1 Use

When using a “reference standard”, it is important to ensure that any information related to its characterisation and how it was used is retained and documented as evidence for any future or related activity. If QA4EO-QAEO-GEN-DQK-002 is being used for the activity, there is a prompt within that key guideline for this purpose. It is important to ensure that the characterisation information appropriate to the time of use of the “reference standard” is recorded to avoid confusion with any potential later revisions.

In many situations a “reference standard” will be “stand-alone” and will have a performance checking or assessing function. In such cases, there is no requirement to seek any formal endorsement or registration of the “reference standard”. The only requirement will be to ensure that the organisation or individual for whom the “reference standard” is ultimately being used (to provide evidence of compliance) considers it acceptable. However, for the benefit of the community as a whole (education and cost saving) and/or a means of improving harmonisation, there is a benefit in making available and accessible to all the information on the “reference standard”.

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**

8.4.2 Registration

This can be done through registering (listing) the “reference standard” and its associated characteristics and operational guideline(s) on the QA4EO website (<http://QA4EO.org/>) or other community specific portals as appropriate. Registration will be subject to the provision of a complete set of information (e.g. a completed template) provided by the registrant and it will be their responsibility to ensure that it is correct. It will not be the subject of any formal review (other than any declared within its submitted documentation) and it will be for readers or prospective users to make their own assessment of its suitability for their particular application. Each registered “reference standard” will be assigned a unique identifier. There may be more than one “reference standard” listed for the same task or activity.

8.4.3 Endorsement

For some applications, particularly where interoperability, data harmonisation, absolute accuracy and bias evaluation are key requirements, it may be appropriate for communities to establish and/or recommend specific “reference standards” and associated methods. For example, natural reference targets such as the Amazon rainforest for SAR, the Moon and a set of deserts have been specifically identified for space-based EO. In these cases, such “reference standards” should be formally proposed, subjected to peer review and subsequently endorsed by the relevant representative community body.

All such formally endorsed “reference standards” should then be submitted to the QA4EO secretariat for inclusion on the website as “community-endorsed standards”.

Any individual or organisation can propose a “reference standard” for endorsement for any purpose. Candidates should first be registered with the GEO-recognised community-representative body and then subjected to a peer review process defined by that particular body. This peer review process will consider the evidence supplied by the proposer(s) in support of the “reference standard” and its importance to the community. It will be for each individual representative body to determine what evidence will be required for each “reference standard” and this may vary with time and criticality of need. For example, it may be more important to establish one common “reference standard” quickly than to wait for longer-term studies to ensure it is the best candidate. However, in time such processes are likely to evolve to a common set of requirements in terms of necessary evidence.

Following this review, selected “reference standards” will then be formally community-endorsed. Endorsed “reference standards” will be listed as such on the QA4EO website

A guide to “reference standards” in support of Quality Assurance requirements of GEO

and, where appropriate, the international community should be encouraged to make use of them.

Endorsed standards will be subject to continued review (~ annually) by the relevant organisation and may lose that status in the light of new information or, in the case of a “test site”, lack of maintenance.

In some circumstances, GEO member organisations may request that an “endorsed reference standard” is established to meet a particular urgent need. In such situations, appropriate community representative organisations may be asked to consider if it is aware of potential candidates or where it may “commission” or encourage the development of a bespoke standard.

9 Conclusion

This key guideline summarises the processes that should be followed to identify, establish and use a “reference standard” as a means of evaluating performance and/or compliance for a particular activity as part of an internationally harmonised QA procedure. The steps needed to register and/or obtain formal endorsement of a “reference standard” to facilitate broad uptake and appropriate usage in support of GEOSS have been set out. However, as a “procedural document”, it does not contain sections on performance evaluation, etc.

Annex A

Reference standard description template

The following template provides a summary checklist of generic information required to describe a “reference standard”. It may be more appropriate to provide links to other documents where more detailed descriptions and information exists.

QA4EO guide: QA4EO-QAEO-GEN-DQK-003

Community Approved

Name / identifier of reference standard:

Application / scope of use (parameter being tested):

Key characteristic of reference standard being used

Ability to discriminate “test item” in context of parameter

Temporal stability of reference standard characteristic

Contributory characteristics of reference standard or its use

Evidence to support described characteristics

Measurement report

By whom

When

How

Maintenance of reference standard following characterisation

Need

How

Reference of operational procedure

Reference to more detailed information if available

Contact for further information

**A guide to “reference standards” in support of Quality Assurance requirements of
GEO**