Case studies for the MATHMET Quality Management System at VSL, the Dutch National Metrology Institute

Gertjan Kok[1]

[1] *VSL, Thijsseweg 11, 2629 JA, Delft, The Netherlands*

ABSTRACT

The European Metrology Network MATHMET is a network in which a large number of European national metrology institutes combine their forces in the area of mathematics and statistics applied to metrological problems. One underlying principle of such a cooperation is to have a common understanding of the 'quality' of software, data and guidelines. To this purpose a flexible, lightweight Quality Management System (QMS), also referred to as Quality Assessment Tools (QAT), is under development by the EMN. In this contribution the application of the QMS to several use cases of different nature by VSL is presented. The benefits and usefulness of the current version of the QMS are discussed from the particular viewpoint of a particular employee of VSL, and an outlook for possible future extensions and usage of the QMS is given.

1. INTRODUCTION

Metrology is the science of measurement founded on the SI system of units [1]. The metrological traceability of measurement results is an essential part of metrology. It is defined [2] 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. For this chain to work properly, all constituent parts should be carefully assessed and validated, and the results of the validation should be properly documented. Measurement results play a dominant role in this chain, but that does not mean that hardware and instrumentation are the only things that matter. Mathematical calculations implemented in software can form an essential part of the measurement. These mathematical calculations will almost certainly be implemented in software, which may have been validated using some reference datasets. The whole data analysis procedure may be based on a written guideline. For complying with metrological traceability, it is therefore essential that the used software, data and guidelines are

also under quality control. This requirement means that their working and content is checked for correctness, and that storing meta-information like version control data is properly managed.

As there is worldwide cooperation within metrological applications, it is logical to organize this type of quality control at an international level. The EMN MATHMET [3] is therefore developing a lightweight Quality Management System (QMS) against which the existing procedures at National Metrology Institutes (NMIs) can be benchmarked and which can help to complement them to get more uniformity in assessing the quality of software, data and guidelines by different NMIs. A full description of this QMS is presented in [4], which has recently been published in Acta IMEKO. This article is an accompaniment to it and supports [4].

In this contribution we will report on the application of the QMS to several use cases concerning software, reference data and guidelines by VSL. We will discuss the usefulness, advantages, and disadvantages of the QMS and possible pitfalls in sections 3 to 5, after shortly introducing the QMS in section 2. Finally, in section 6 some overall conclusions will be formulated. Note that all these viewpoints and conclusions are

from the perspective of one employee of VSL only, relate to the particular version of the QMS of March 2022, and they are not necessarily shared by other NMIs or by EMN MATHMET itself.

2. SHORT OVERVIEW OF THE QMS FOR SOFTWARE, DATA AND GUIDELINES

A thorough overview of the QMS for software, data and guidelines is given in [4]. In this section a summary is given, starting with some remarks regarding its scope.

2.1. Goal of the QMS

Originally there was the idea that the EMN itself would 'recommend' software, reference data and guidelines. Assessment of these items by means of the QMS would ensure that the EMN MATHMET recommendations meet the highest quality levels and achieve wide use and substantial impact. For various reasons this is currently not seen as realistic. One important reason is the fact that an EMN is part of the larger entity Euramet [5] and that the decision-making authority, responsibility and liability for such recommendations is not entirely clear. The second reason is the scope of the EMN, which is now seen as a platform to interact with stakeholders and to define future research directions, fostering collaboration and preventing duplication of work. Actual technical work should be done inside other forms of cooperation. Linked to this reason is that the required budget for performing an assessment is not available within the EMN itself.

The QMS might therefore be seen as a tool to help individual NMIs assessing software, data and guidelines, rather than a tool for the EMN itself. Performed reviews will not be published on the MATHMET website but could be put on the website of an individual NMI if an NMI would wish to do so. It therefore seems reasonable to assess the MATHMET QMS from the perspective of a single NMI. Various NMIs in MATHMET have announced that they will indeed discuss the benefits of the MATHMET QMS with people directly responsible for quality control at their respective NMIs. This will be done in the near future at VSL as well. This article presents an initial assessment of the QMS by the author.

2.2. QMS for software

The QMS for software consists of an interactive pdf-file of 5 pages. Based on the calculated risk level, specific fields are visible and need to be filled out. The QMS for software requires that the project team provides information and evidence of documents covering the following aspects and activities:

- Some meta-data
- A risk level analysis resulting in a "software integrity level" that determines the quality interventions needed
- User requirements
- Functional requirements
- Design
- Coding
- Verification
- Validation
- Delivery, use and maintenance.

2.3. QMS for data

The QMS for data consists of an interactive pdf-file with 41 questions, which are again only visible if they are deemed relevant

for the selected risk level. For data the team should provide information regarding:

- General details and responsibilities
- A risk level analysis, resulting in a "data integrity level" that determines the quality interventions needed
- User requirements documentation and approval
- Data life cycle documentation
- Quality planning
- Quality monitoring, control and improvement
- Quality assurance
- Understandability
- Metrological soundness

Most questions to be answered are of general nature. Only the last set of questions explicitly involve some metrological aspects. There are no questions explicitly addressing the mathematical aspect the data may have.

2.4. QMS for guidelines

The QMS for guidelines consists of two different checklists: one checklist for existing guidelines and one for future guidelines. At the moment of writing of this manuscript these checklists are still out for review by the project partners, but preliminary versions have been assessed by VSL. The checklists are quite similar. They ask for information regarding:

- organization generating the document
- independent review and approval available
- appropriate metadata available
- copyright and IP protection
- language
- mentioning of target audience
- relevance for mathematics and statistics in metrology and the target audience
- clearly stated conclusions
- appropriate references
- presentation easy to understand

What is noteworthy, is that the QMS checklist is not asking to perform a thorough review of all mathematics by the user of the checklist, but rather to assess if this has been done (and documented!) already and by whom. For some questions, e.g., regarding the presentation and conclusions, it would of course be beneficial to read through the complete document. However, these questions can also be answered with a reasonable level of confidence by only reading small parts of the document.

3. USES CASE 1: QMS APPLIED TO SOFTWARE

The usefulness of the QMS has been assessed by applying it to two pieces of software.

3.1. Context

The first piece of software was a library of mathematical routines written in Python which can be used to take advantage of redundancy in sensor network data [6]. It was developed in the research project Met4FoF [7]. At a first stage a chi-squared based consistency check is performed to assess the statistical consistency of the sensor data. In the case of consistency, the measurement data is combined into a best estimate of the measurand respecting sensor uncertainties and covariances, whereas in the second case the largest consistent subset of

sensors is constructed. This is not only done for the case that the sensors directly measure values of the measurand, but also if there is a linear relationship between the vector of sensor values x and a vector of values y for the measurand. This vector y reflects the availability of multiple, redundant estimates of the measurand. The relationship takes the form

$$y = a + B x, \tag{1}$$

in which a is a vector and B a matrix.

The second piece of software that was used to evaluate the QMS consisted of a recently developed calculation module based on the written standard ISO 6142-1 [8]. This software is used at VSL in the production of certified reference materials (i.e., gas mixtures) for customers. The input to the software consists of atomic weights, chemical formulas of the mixture components, amount of substance fractions of the components in the parent gases and the added gas mass from each parent gas mixture to the target gas mixture based on weighing of the cylinder. The outputs of the calculation module are the amount of the substance fractions of the components in the target gas mixture, including their uncertainties and covariances.

The quality system at VSL requires that software should be version controlled, documented and validated. However, there are no uniform, detailed procedures or templates to this purpose. In practice different groups assure the quality in different ways.

3.2. Benefits and usefulness of the QMS

The QMS for software was applied to these pieces of software with the aim of assessing the QMS, rather than constructing all required information that might not be readily available.

The following parts of the QMS for software were especially appreciated:

- The templates help to give a uniform description of the software.
- The templates help to avoid overlooking important aspects of software quality.
- At VSL there is a focus on version control, documentation, and validation of software and storing these properly. User and functional requirements as well as software design may be lost after the software has been released. It would be good to properly control these documents as well. Especially the documentation of software design could be useful for future improvements of the software, possibly by new personnel.

The following parts of the QMS for software seemed to be less appropriate for the VSL context:

- The 'review by customer or proxy' may need some flexible interpretation, as VSL does not sell any software to external customers. There could be a 'VSL internal customer' for the software, and/ or the envisaged outputs of the software could be assessed against known requirements of customers. In the case of research projects that are, e.g., funded by the EU, the 'review by customer' is usually difficult to achieve.
- The number of up to three required reviews for some aspects is quite large and can be burdensome, especially for a small NMI like VSL.
- The document asks for requirement and design documents at the moment of filling out the form. At VSL, often a gradual, Agile based, approach is used

for software development. It is not so clear for the author of this paper how the QMS should be used in that context. Should the QMS forms and all implied documentation and reviews be repeated at each 'sprint' (development cycle) or at each new release of the software? Some more guidance and clarity would be beneficial.

As a general observation it would be helpful if the QMS indicates some examples of ICT tools (preferably open source) that could be used in combination with an Agile development process while assuring the traceability (in an administrative sense) of all choices made. In this way, possible requirements of the MATHMET QMS that may not be directly accommodated by the quality system and available software systems at an NMI, could more readily be implemented at an NMI and used in the context of work related to EMN MATHMET.

4. USE CASE 2: QMS APPLIED TO DATA

In this section we will assess the QMS for data by applying it to some mathematical reference datasets that were generated in an earlier project.

4.1. Context

In the TraCIM project [9] reference datasets were produced for various mathematical problems, e.g., for non-linear least squares fitting problems. The precise definition of the computational aims, together with the datasets were stored in a database [10] accessible from the internet for registered users.

An example of such a fit problem is the determination of the best fit parameters a', b' and c' of a function $y = f(x, a, b, c)$ modelling exponential decay to n datapoints (x_i, y_i) $1 \leq i \leq n$ such that

$$D(a, b, c) = \sum_{i=1}^{n} (y_i - f(x_i, a, b, c))^2 \rightarrow \min. \tag{2}$$

In the database [10] reference datasets are available for this and other fit problems.

At VSL there is no specific quality guidance for the generation and documentation of such datasets, other than the requirements mentioned in section 3 for software.

4.2. Benefits and usefulness of the QMS

The interactive pdf file with 37 pages and at most 42 questions was filled out for the application described in section 4.1.

The following parts of the QMS for data were especially appreciated:

- With the help of the data quality management plan template a uniform plan for all applications can be created.
- The questions cover a large range of quality aspects, which might be forgotten if the QMS tool would not be used.

The following parts of the QMS for data seem to be less appropriate for VSL's needs:

- The mathematical aspect of the data is not particularly addressed. There could be more guidance with respect to how to assess the correctness of numerical data.
- Four responsibilities related to data are mentioned: data manager, data administrator, data steward, data technician. These roles do not always seem to exist

at VSL, especially not for data generated in research projects. In many cases the situation seems to be much simpler.

- Similar to what was mentioned in the QMS for software, it would be nice if more guidance could be given regarding how to implement all mentioned quality aspects by means of some ICT tools, preferably open source.

5. USE CASE 3: QMS APPLIED TO GUIDELINES

In this last use case, we will discuss the application of the QMS to a set of mathematical guidelines which was produced in the EMRP project NEW04 [11], and to which VSL contributed.

5.1. Context

In the NEW04 project three best practice guides (BPGs) were produced [12]:

1. A Guide to Bayesian Inference for Regression Problems (BPG1)
2. Best practice guide to uncertainty evaluation for computationally expensive models (BPG2)
3. A guide to decision-making and conformity assessment (BPG3)

Except for the formatting of the title page, BPG1 and BPG2 are very similar in document structure. BPG3 consists of a set of four different loosely connected documents. We applied the QMS checklist for existing guidelines to BPG2. This document of 84 pages provides a summary of current best practice in uncertainty evaluation for computationally expensive models. In the first part of the document the methods are explained. In the second part three case studies are presented.

5.2. Benefits and usefulness of the QMS

The QMS checklist is mainly asking some questions about the existence of specific information like 'version number', 'independently reviewed', 'target audience' and 'appropriate references'. The benefit of this approach is that the assessment can be done fairly quickly without having to read, study and check the document itself. The QMS checklist verifies that some formal quality criteria are fulfilled, and it is not requiring a tedious scientific review of the content. These simple checks can give a good indication of the overall care with which the document has been prepared in a very quick way, which is the main benefit of this QMS for guidelines in our opinion. If the conclusion is that the document hasn't been independently reviewed, then this job is still out to be done, but this is not directly in the scope of the QMS.

The application of the QMS checklist to BPG2 yielded some interesting deficits. BPG2 has no version number, it doesn't say anything about 'copyright', or 'independent review' and there are no 'clearly stated conclusions'. The document simply ends with the last use case. This is particularly interesting, because several of the authors of BPG2 are MATHMET members and even involved in the creation of the QMS. The mathematical content of the BPG may be impeccable, but it doesn't fulfil all quality metrics of the MATHMET QMS.

6. CONCLUSIONS AND OUTLOOK

As over time the scope of the EMN has become clearer, also the place of the QMS in it has been reassessed. Initial ideas about a MATHMET QMS that ensure that 'EMN MATHMET recommendations meet the highest quality levels and achieve wide use and substantial impact' [13] seem to have been replaced in practice by a QMS that can help individual NMIs with their quality assessment, at least from the perception from VSL. In this paper it has been assessed how this worked out for VSL, and which aspects of the QMS proved useful and which less appropriate for the VSL context.

The overall conclusion is that the collaboration within the EMN on the QMS gave useful insights with respect to assuring the quality of software, data and guidelines, and which aspects could matter. At the same time a proper assessment of the different parts and questions of the QMS is needed in order to best align it with VSL's requirements and working field. A discussion with the quality coordinators at VSL is still outstanding.

When more NMIs assess the MATHMET QMS and reflect on its implementation in NMI specific quality procedures, the common ground of most useful aspects of the QMS will become clearer. This may lead to a next step in the development of the QMS, which should lead to a greater uniformity in quality assessment of software, data and guidelines by NMIs, and to a reduction of costs to set-up the system. Also, there might be additional guidance for the usage of modern ICT tools to assure the quality of software, data and guidelines in a more efficient way.

As guaranteeing the quality of software, data and guidelines (cf. the attention paid to research papers with open software and data) is getting nowadays more and more attention, the creation of a common QMS framework by EMN MATHMET for NMIs seems to come at the right moment. This and similar initiatives will help to maintain and increase the trustworthiness in services provided by NMIs.

REFERENCES

[1] EURAMET - European Association of National Metrology Institutes, European Metrology Network (EMN) MATHMET. Online [Accessed 6 June 2023]
www.euramet.org/mathmet

[2] BIPM, SI system of units. Online [Accessed 6 June 2023]
https://www.bipm.org/en/measurement-units

[3] BIPM, VIM3, definition of traceability. Online [Accessed 6 June 2023]
https://jcgm.bipm.org/vim/en/2.41.html

[4] Keith Lines, Jean-Laurent Hippolyte, Indhu George, Peter Harris, A MATHMET Quality Management System for data, software, and guidelines, Acta IMEKO, vol. 11, no. 4, article 8, December 2022, pp. 1-6.
DOI: 10.21014/actaimeko.v11i4.1348

[5] EURAMET - European Asssociation of Metrology Institutes. Online [Accessed 6 June 2023]
www.euramet.org

[6] Metrology for the Factory of the Future, Github software repository. Online [Accessed 6 June 2023] https://github.com/Met4FoF/Met4FoF-redundancy

[7] EMPIR project 17IND12 Met4FoF Metrology for the Factory of the Future, 2018 – 2021. Online [Accessed 6 June 2023] https://www.ptb.de/empir2018/met4fof/home/

[8] ISO 6142-1:2015, Gas analysis — Preparation of calibration gas mixtures — Part 1: Gravimetric method for Class I mixtures, ISO, Geneva, 2015

[9] EMRP project NEW06 TraCIM Traceability for Computationally-Intensive Metrology, 06/2012 - 05/2015; Web site. Online [Accessed 6 June 2023] https://www.tracim.eu

[10] TraCIM database with reference datasets. Online [Accessed 6 June 2023] http://www.tracim-cadb.npl.co.uk/

[11] EMRP project NEW04 Novel mathematical and statistical approaches to uncertainty evaluation, 08/2012 - 07/2015, Web site. Online [Accessed 6 June 2023] https://www.ptb.de/emrp/new04-home.html

[12] EMRP project NEW04 Novel mathematical and statistical approaches to uncertainty evaluation: Best Practice Guides. Online [Accessed 6 June 2023] https://www.ptb.de/emrp/2976.html

[13] Project protocol of EMPIR project 15NET05 MATHMET Support for a European Metrology Network for mathematics and statistics, 2019 - 2023