Validation of SI-based digital data of measurement using the TraCIM system

Daniel Hutzschenreuter[1], Bernd Müller[2], Jan Henry Loewe[1], and Rok Klobucar[3]

[1]Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany
[2]Ostfalia University of Applied Sciences, Salzdahlumer Str. 46/48, 38302 Wolfenbüttel, Germany
[3]Univerza v Mariboru (UM), Slomskov trg 15, 2000 Maribor, Slovenia

Correspondence: Daniel Hutzschenreuter (daniel.hutzschenreuter@ptb.de)

Abstract. The digital transformation in the field of sensors and sensor systems fosters an increasing exchange and interoperation of measurement data by machines. The data of measurement need to be uniformly structured based on The International System of Units (SI) with appropriate information on measurement uncertainty. This work presents a concept for an online validation system that can be used by humans and software to efficiently classify the agreement of XML-structured data with relevant recommendations for measurement data. The system is within the TraCIM (Traceability for Computationally-Intensive Metrology) validation platform which was developed for software validation in metrology where high standards of quality management must be met.

1 Introduction

Validation of the quality of measurement data of all kinds and its ability to be interpreted correctly by different software systems is an emerging need for new technologies in metrology including "digital calibration certificates" (DCCs), networks of sensors and virtual measuring instruments (Eichstädt et al., 2017). These new technologies rely on and foster extensive machine-to-machine communication. While the validation and certification of evaluation algorithms with comparable standards has been recognised as an important and economic supplement to calibration in the quality infrastructure for more than 20 years, the validation of the quality of data exchange regarding comparable standards has not been considered to a great extent.

Presented in this paper is an application of the TraCIM (Traceability for Computationally-Intensive Metrology) online validation tool (Forbes, 2016) to realise a classification of the agreement between measurement data and internationally accepted guidelines including *The International System of Units* (BIPM, 2019), the *International Vocabulary in Metrology* (JCGM, 2012) and the *Guide to the Expression of Uncertainty in Measurement* (JCGM, 2008). In order to apply the classification, it is necessary to provide all data in

Extensible Markup Language (XML) (W3C, 2008) according to the Digital System of Units (D-SI) metadata model for a machine-readable digital data exchange (Hutzschenreuter et al., 2020a).

Figure 1 shows an example of the XML coding for the measurement of a temperature quantity in accordance with the D-SI. The metrological data are fundamentally made up of a numerical value and a unit of measure. Additional information is provided by the time when the measurement was made and a statement on the accuracy of the measurement in the form of uncertainty data. The classification process takes into account a comprehensive analysis of all the data that are provided within the various D-SI fields.

The TraCIM online validation tool was originally developed for the validation of software that calculates output quantities y from given input quantities x through solving mathematical models f, e.g. in implicit form $f(x, y) = 0$. As the D-SI classification introduces a completely new kind of validation, integration with the existing TraCIM system provided a significant development challenge. After an introduction to the fundamental working principles of TraCIM in Sect. 2, Sect. 3 will present the D-SI classification concept and the implementation in TraCIM. The latter section gives further information on the underlying measures to guarantee

```
<si:real>
  <si:label>temperature</si:label>
  <si:value>20.10</si:value>
  <si:unit>\degreecelsius</si:unit>
  <si:dateTime>2021-04-13T10:34:00Z</si:dateTime>
  <si:expandedUnc>
    <si:uncertainty>0.50</si:uncertainty>
    <si:coverageFactor>2</si:coverageFactor>
    <si:coverageProbability>0.95</si:coverageProbability>
    <si:distribution>normal</si:distribution>
  </si:expandedUnc>
</si:real>
```

Figure 1. D-SI XML structure for scalar measurement value of a temperature quantity.

the correct functioning of the TraCIM validation tool as it is software itself that needs verification.

The methodological and technical foundation of the D-SI classification presented here and the respective TraCIM online validation service are outcomes from the joint research project SmartCom in the European Metrology Programme for Innovation and Research (EMPIR) (Bojan et al., 2020). For convenience, we will frequently make use of the term *SmartCom validation* to denote the D-SI classification with TraCIM.

2 TraCIM foundations

The TraCIM system was developed inside the TraCIM project (Traceability for Computationally-Intensive Metrology, 2012–2015) as a joint research project within the European Metrology Research Programme (EMRP) (Forbes, 2016). The project consortium was formed of 14 national metrology institutes (NMIs), universities and industrial partners.

The main goal of the TraCIM system was

– to automate the validation of software used in metrology and measuring machines that requires computationally extensive data evaluation,

– to guarantee the metrological traceability of such validations to reliable numerical standards provided by trusted organisations such as national metrology institutes (NMIs), and

– to establish a public and long-term available online validation service for use by customers and machines with low maintenance requirements.

2.1 Basic concepts of TraCIM validations

In the following, we detail the basic requirements of the TraCIM system. The requirements were mainly formulated by PTB (Physikalisch-Technische Bundesanstalt), with the implementation carried out by Ostfalia University of Applied
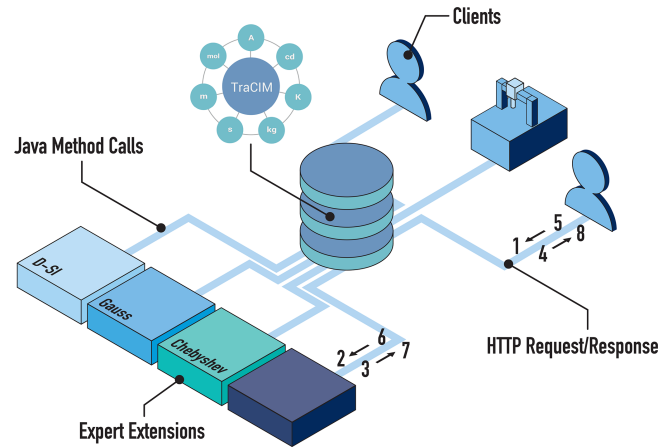

Figure 2. TraCIM architecture and message flow.

Sciences. The TraCIM system has been designed to be universally usable to validate different software systems that are mainly used in measuring machines. Typical algorithms implemented by such software systems include Gaussian and Chebyshev geometric element fitting (Härtig et al., 2015; Wendt et al., 2016). Because of the wide range of different algorithms, the TraCIM system core was designed to be application agnostic using the minimum possible knowledge about the underlying algorithms. It performs the function of a management system to control the validation process while so called *expert extensions* or experts for short take over the real validation process. Figure 2 depicts the general message flow.

After undertaking a number of basic administrative tasks, including registering as a user and placing and paying for an order for specific data sets, a user can then download the relevant data sets. Each order has a unique order key and an individual expert assigned to it; for example, in the example in Fig. 2, an expert is the Chebyshev extension. The individual steps for a complete validation process are as follows:

1. A client makes a request for a test data set. As the order key is included, the TraCIM server can establish a mapping from the order key used to the expert involved.

2. The server stores the request and hands it over to the expert.

3. The expert returns the test data set to the TraCIM server. The data can be generated on-the-fly, read from a database or supplied by other means. Optionally, the expert returns the expected result for the test data set.

4. The server stores the test data set and an eventually available expected result in the database and hands the test data set over to the client.

5. After the client has computed the test results based on the test data set, results are submitted to the server.

6. Because the test result contains a link to the relevant expert, the server sends the test result and an eventually available expected result from the database to the appropriate expert.

7. The expert validates the submitted result either by a comparison algorithm or by comparison with the expected result. Typically, the validation result is either *passed* or *not passed*, which is sent to the server.

8. The server stores the validation result in the database and forwards it to the client.

The process described above is in use at PTB since 2015 for Gauss and Chebyshev geometric element fitting tests in the domain of coordinate metrology (Härtig et al., 2015; Wendt et al., 2016). More than 90 software certifications have been executed by users from 21 countries worldwide in this time.

2.2 Quality of TraCIM validations

The TraCIM validation system was developed for an application in the industrial field of metrology where developers of measurement software rely on independent institutes such as NMIs to provide services for the certification of their computational capabilities. Trust in the validation services depends on the quality of each validation – a fundamental requirement of all national and global measurement service infrastructures. To establish the necessary levels of quality for validations and to assess their maturity, the TraCIM project defined quality rules that need to be met by each validation service. The set of operational quality rules takes into account the high-quality management needs for validation services from NMIs such as PTB in Germany. Mathematical quality rules were designed to guide the development of TraCIM validations with one correct and unambiguous result. Today, these quality rules are maintained and further developed under the umbrella of the TraCIM association that was founded in 2015 alongside the TraCIM system development. For completeness, the current list of all quality rules is provided in Appendix A. The quality rules will not be further discussed in this work. The interested reader can refer to a detailed discussion of the quality rules in respect of Chebyshev fitting from Hutzschenreuter (2019).

3 SmartCom D-SI document classification

The aim of the SmartCom validation process is to classify data from measurement according to its fulfilment of requirements for digital metrological data. Underlying requirements have been collected and summarised in the universal, unambiguous, safe and easy-to-understand *Digital System of Units* metadata model (D-SI) (Hutzschenreuter et al., 2020a). It takes into account principles from the most important international guidelines in metrology, including the following:

1. SI – The International System of Units – representation of units of measure for physical quantities by seven SI base units (BIPM, 2019);

2. VIM – International Vocabulary in Metrology – common language for metrological terms like quantity, number, unit and uncertainty (JCGM, 2012); and

3. GUM – Guide to the Expression of Uncertainty in Measurement – recommendations for evaluation of measurement uncertainty and reporting of measurement uncertainty (JCGM, 2008).

The SmartCom validation process was developed for data provided in XML format, e.g. in a document, where metrological quantities are given by the D-SI XML reference structure (Hutzschenreuter et al., 2020b). The validation procedure runs through the provided XML document and selects all XML elements from the D-SI namespace for testing. User-specific data in the XML document which are not part of a D-SI element are not considered. To run the validation, it is essential that all data are valid XML. Documents with invalid XML are rejected. For each of the following five D-SI XML elements, independent classifications will be made whenever the test procedure finds these element types in the document under test.

– real: scalar quantity values with SI units and univariate measurement uncertainty.

– constant: scalar values of physical and mathematical constants with SI units and univariate measurement uncertainty.

– complex: Cartesian and polar coordinate complex quantities with SI units and bivariate measurement uncertainty.

– list: vectors of real and complex types with multivariate uncertainty.

– hybrid: machine-readable adaption of quantities with non-SI units.

The classification assigns each D-SI element a quality ranking or "medal" ranging from PLATINUM (best agreement with metrological needs) to IMPROVABLE (incompatible with metrological needs). Section 3.1 explains the different kinds of medals. After all individual D-SI elements have been classified, an overall result is evaluated by selecting the lowest class that was achieved among all individual D-SI elements.

3.1 Classes at D-SI validation

The D-SI validation procedure associates the XML elements with the following classes that are ordered in a medal ranking from PLATINUM (best) to IMPROVABLE (worst).

- PLATINUM is achieved when all units of measure are provided by only using SI base units, the unit for dimension 1, hour, minute, degree, arcminute and arcsecond. Furthermore, all metrological information in the data is complete and reasonable according to their mathematical definition (e.g. feasible boundaries for intervals).

- GOLD is achieved if the data satisfy the requirements of the PLATINUM class and where SI prefixes (e.g. kilo) and SI-derived units with their own symbols (e.g. newton) are used with the units of measure.

- SILVER is achieved if the data satisfy the requirements of the GOLD class and where additional SI-allowed non-SI units (e.g. dalton) are used with the units of measure.

- BRONZE is achieved if the data satisfy the requirements of the SILVER class and where units of measure are used from the eighth edition of the SI system (BIPM, 2006), which were deprecated in the ninth edition of the SI system.

- IMPROVABLE is the ranking of a D-SI XML element if units of measure other than those allowed in the eighth and ninth editions of the SI system are used (e.g. feet) and if even no unit is stated in the data. Furthermore, data get the attribution IMPROVABLE if the metrological information is incomplete or not reasonable, e.g. if the value of an uncertainty is negative while it must be a positive value by definition, if dimensions of lists are not compatible with dimensions of associated uncertainty statements, etc.

The XML elements of the D-SI data model are supported by an XML Schema Definition (XSD). The validation of XML data against a XSD is a standard approach in computer science, and whenever possible it was used for the D-SI validation procedure. However, in many situations concerning the evaluation of the units, compliance of dimensions and reasonable uncertainty, additional software is needed to supplement the XSD validation. In our test procedure, the validation with the XSD is the initial step to exclude basic incompatibilities with the D-SI model. If this validation is successful, the XML data are read by a software that performs further comprehensive testing to evaluate the correct data classes.

3.2 From validation to classification – the TraCIM way

At a first glance, the TraCIM concept and SmartCom classification are very different. However, after some investigation, it turns out that the second half of the TraCIM validation process can adequately accommodate the SmartCom procedure. Applying SmartCom within the TraCIM framework requires no test data set to be delivered to the client. The clients can

start by defining XML documents which adhere to some D-SI structure and then send it to the TraCIM server for validation in step five of the TraCIM procedure of Sect. 2.1. This implies that the first four steps have to be omitted. There is no validation of software with respect to some input–output behaviour but validation of the XML data from the client going through steps six to eight of the TraCIM procedure. Specifically, at step seven in Sect. 2.1, the D-SI expert extension for the SmartCom classification acts as follows:

- The D-SI expert evaluates the class achieved by all XML data.

- The test is *passed* if one of the medals form Sect. 3.1 is obtained. Tests with invalid XML are considered *not passed*.

- The validation result with detailed information of the kind of medal is sent back to the server.

A clear description of the validation details and how the data exchange works is provided by a user manual (Lin et al., 2020). Currently, at the time of writing, a client can request a public SmartCom validation for XML data against the D-SI classification without charge. With each validation, a comprehensive validation report is delivered to the client. It describes the validation scope and validity range of the validation results. If the validation outcome is a classification below PLATINUM, the validation report will provide an additional appendix. This appendix will document all locations in the XML where a ranking below PLATINUM was detected and the reasons for the classification. These details are to enable clients to identify opportunities for data improvement and to remove errors.

3.3 Reliability of the TraCIM D-SI classification software

Taking sufficient measures to verify reliable and correct outcomes of TraCIM validations is an essential aspect of the underlying quality rules. In the case of "traditional" TraCIM algorithm validation, three independent reference software implementations should be used to verify the correctness of the reference data for tests (see Appendix A, rule 13). In the case of SmartCom classification, this requirement is transferred to a validation of the D-SI expert software with independent example data provided by three different members of the SmartCom project. The example data consist of XML documents with over 500 D-SI example elements for the different classes PLATINUM to IMPROVABLE from Sect. 3.1. These examples are stored within the D-SI XML scheme repository (Hutzschenreuter et al., 2020b).

In the first stage of the application, the example data are used to establish standard unit and integration tests for the software at compilation and deployment. It supports the verification of the software at development and source code editing.
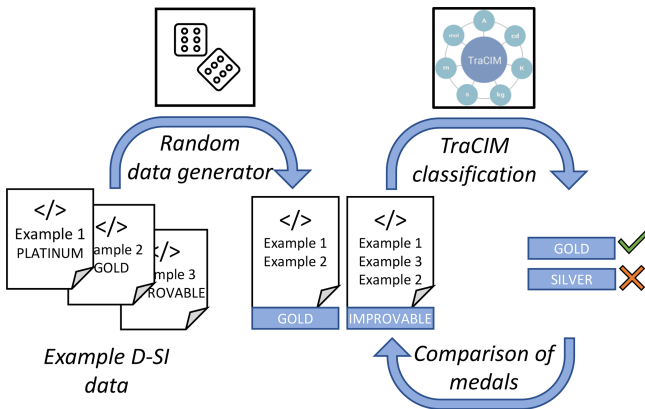
Figure 3. Verification of the D-SI classification software running in TraCIM.

In the second stage of the application, the D-SI classification is verified when it is running in the TraCIM system environment. This stage takes into account the correct function of the test in practical situations that

- cover the relevant part of the message flow of the TraCIM architecture (e.g. Sect. 2.1),

- include data sets made up by two or more different D-SI elements with different medals, and

- allow for more comprehensive numbers of different data sets than the number given by the pure example data

Figure 3 outlines the concept of the verification process that we have implemented for the second stage. The starting point is the D-SI example data sets. A data generator randomly assembles comprehensive XML documents by putting together several D-SI examples in new XML files. Thereby, for each new file the expected medal (result) of a TraCIM classification is the medal with the lowest ranking from the underlying D-SI examples (ranking according to Sect. 3.1). The comprehensive XML documents are sent with an internal component of software (client) to the TraCIM system that is running the D-SI validation subject to verification. The test results from TraCIM are collected by the client. The TraCIM verification is successful if all medals reported in the TraCIM result match the expected medals of the XML documents.

The reliability of the verification outcome will increase with the number of data sets that are used. Thus, features of automation are considered at several points to support an efficient handling of many data sets. The data generator can create batches of data sets with common properties, e.g. the kind of medals and numbers of D-SI elements in each XML file. Once a database with different XML files has been created this way, the verification client is able to perform the verification with all data in one run.

4 Conclusions

With the development of the D-SI classification process, a significant step forward away from classical algorithm validation towards the new field of general data analysis with TraCIM has been made. The challenges were to find a concept for integrating the D-SI classification of data into the TraCIM requiring minimum refactoring of the system and ensuring sufficient compliance with the TraCIM quality rules. Following an analysis of the requirements, a solution was found that provided the necessary functionality requiring a suitably minor change to the original TraCIM design.

PTB is currently implementing the required changes to the TraCIM system, and a D-SI classification service is planned for customer access in 2021. An example service for a proof of the concept is already available online (Lin et al., 2020). An important part of the remaining work is to create XML files for the verification of the D-SI test service with the random data generator. A plan to develop a data sampling concept where all possible combinations of medals as well as small, medium, and large size files are equally present is also underway.

Finally, D-SI classification of XML-formatted data is only one example of services that will be needed in the future for a trusted verification and certification of the digital data quality. One of the most important developments in the metrological field is machine-readable digital calibration certificates (Brown et al., 2020). Here, we expect the need for verification of complex data structures on community specific requirements beyond the functionality provided through XML data schemes.

Appendix A: TraCIM quality rules

The TraCIM association provides universal quality rules for software testing, fulfilling needs from metrology. Listed below is the 2019 set of these quality rules for long-term available and highly automated online test services with minimum need for human interaction.

A1 Operational quality rules

§1. Each institution running a test service using the TraCIM trademark must be a member of the TraCIM association.

§2. It is recommended that all quantities related to a test be provided with a unit in the SI.

§3. A clear description of the test procedure and its data exchange formats must be provided, which comprises

- the access to the test (registration, order) and work flow of testing,

- the data provided by customer and data provided by test system,

- the data exchange format of the server and the format of the test data, and

- using well-established data formats for machine-readable communication.

§4. It is recommended to provide a public test.

- The public test shall include a set of public reference pairs if appropriate.

§5. Each institution running a test service using the TraCIM trademark has to operate an integrated quality management system. It has to cover the following aspects:

- Test reports and the corresponding reference pairs must be stored for at least 10 years.

- Test data (and software) shall be available and their integrity shall be ensured for at least 10 years.

- Inevitable maintenance of server, expert modules, data sets and structures must be considered.

§6. A test report must provide as a minimum the following administrative information:

- administrative core data
  - service provider name, address, and responsible person;
  - customer name and address;
  - tested software name, version, and provider with address;
  - unique identification of all parts belonging to the certificate;
  - date of testing and date of issue of certificate;
  - signature.

- description of test type, scope of tested algorithms, and validity range of test result;

- uncertainty of test evaluation shall be stated with the results of the test if relevant;

- description of evaluation method and decision rules for the test;

- description of method and plan for the sampling of the underlying test data;

- additional national required information if relevant.

§7. The software used for rating the performance of the algorithms under test (i.e. the expert modules) must be maintained under a revision control system. Each test report must be traceable to the software revision used for its creation.

A2 Mathematical quality rules

§8. A clear and unambiguous mathematical description of the underlying computational aim must be provided, which includes

- the clear description of all computational tasks (a reference to a reviewed computational aim data base is recommended),

- the description of input and output quantities,

- any other information necessary for correct solutions like rounding rules or the amount of solutions.

§9. The input data are assessed error free.

- Input quantity values have no uncertainty.

§10. Reference parameters must be provided and also if relevant their numerical uncertainty.

- Determination of reference parameters and uncertainty must be disclosed.

§11. A test shall provide only one correct result.

- Each result must be unambiguous.

- All possible results must be covered by the reference parameters.

§12. A test shall reflect common practical situations and not academic situations.

§13. The validity of reference pairs must be proven.

- They shall be validated by at least three independent implementations.

- They can in addition comprise checks of mathematical plausibility requirements.

Author contributions. BM and DH contributed with the idea, performing the work to establish the TraCIM D-SI classification and preparing the article. JHL contributed to the implementation of the D-SI expert and unit classification. RK contributed significant parts of the verification work.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

References

BIPM: The InternationalSystem of Units(SI) – 8th edition, Publications of the Bureau Internationaldes Poids et Mesures (BIPM), available at: https://www.bipm.org/documents/20126/41483022/si_brochure_8.pdf (last access: 15 November 2021), 2006.

BIPM: The InternationalSystem of Units(SI) – 9th edition, Publications of the Bureau Internationaldes Poids et Mesures (BIPM), available at: https://www.bipm.org/documents/20126/41483022/SI-Brochure-9-EN.pdf (last access: 15 November 2021), 2019.

Bojan, A., Weber, H., Hutzschenreuter, D., and Smith, I.: Communication and validation of metrological smart data in IoT-networks, Adv. Prod. Eng. Manag., 15, 107–117, https://doi.org/10.14743/apem2020.1.353, 2020.

Brown, C., Elo, T., Hovhannisyan, K., Hutzschenreuter, D., Kuosmanen, P., Olaf, M., Mustapää, T., Nikander, P., and Wiedenhöfer, T.: Infrastructure for Digital Calibration Certificates, 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT, 3–5 June 2020, Roma, Italy, https://doi.org/10.1109/MetroInd4.0IoT48571.2020.9138220, 2020.

Eichstädt, S., Bär, M., Elster, C., Hackel, S. G., and Härtig, F.: Metrology for the Digitalization of the Economy and Society, Physikalisch-Technische Bundesanstalt, PTB Mitteilungen, Fachverlag NW, Carl Schünemann, Bremen, Germany, https://doi.org/10.7795/310.20170499, 2017.

Forbes, A.: Final Publishable JRP Summary for NEW06 TraCIM Traceability for computationally-intensive metrology, EURAMET European Metrology Research Program (EMRP), Euramet, available at: https://www.euramet.org/ (last access: 8 December 2021), 2016.

Härtig, F., Wendt, K., Brandt, U., Franke, M., Witte, D., and Müller, B.: Traceability for Computationally-Intensive Metrology Validation of Gaussian Algorithm Implementation, available at: https://tracim.ptb.de/tracim/resources/downloads/ptb_math_gauss/ptb_math_gauss_manual.pdf (last access: 15 November 2021), 2015.

Hutzschenreuter, D.: Validation of evaluation algorithms in multidimensional coordinate metrology, Physikalisch-Technische Bundesanstalt PTB-F-63, Fachverlag NW, Carl Schünemann, Ph.D. thesis, technical University Carolo-Wilhelmina, Braunschweig, Germany, 2019.

Hutzschenreuter, D., Härtig, F., Wiebke, H., Wiedenhöfer, T., Forbes, A., Brown, C., Smith, I., Rhodes, S., Linkeová, I., Sýkora, J., Zelený, V., Ačko, B., Klobučar, R., Nikander, P., Elo, T., Mustapää, T., Kuosmanen, P. M., Maennel, O., Hovhannisyan, K., Müller, B., Heindorf, L., and Paciello, V.: SmartCom Digital System of Units (D-SI) Guide for the use of the metadata-format used in metrology for the easy-to-use, safe, harmonised and unambiguous digital transfer of metrological data – Second Edition, Zenodo, https://doi.org/10.5281/zenodo.3816686, 2020a.

Hutzschenreuter, D., Härtig, F., Wiedenhöfer, T., Hackel, S. G., Scheibner, A., Smith, I., Brown, C., and Heeren, W.: SmartCom Digital-SI (D-SI) XML exchange format for metrological data version 1.3.1, Zenodo [data set], https://doi.org/10.5281/zenodo.3826517, 2020b.

JCGM: Evaluation of measurement data – Guide to the expression of uncertainty in measurement – JCGM 100:2008, Publications of the Bureau Internationaldes Poids et Mesures (BIPM), available at: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf (last access: 15 November 2021), 2008.

JCGM: International vocabulary of metrology – Basic and general concepts and associated terms (VIM) – JCGM 200:2012, Publications of the Bureau Internationaldes Poids et Mesures (BIPM), available at: https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf (last access: 15 November 2021), 2012.

Lin, S., Hutzschenreuter, D., Loewe, J. H., Härtig, F., Müller, B., and Heindorf, L.: Traceability for computationally-intensive metrology – Test for communication interfaces used for the exchange of metrological data, Zenodo, https://doi.org/10.5281/zenodo.3953555, 2020.

W3C: Extensible Markup Language (XML) 1.0 (Fifth Edition), World Wide Web Consortium (W3C) Recommendation 26 November 2008, available at: https://www.w3.org/TR/xml/ (last access: 15 November 2021), 2008.

Wendt, K., Brandt, U., Lunze, U., and Hutzschenreuter, D.: Traceability for computationally-intensive metrology User manual for Chebyshev algorithm testing, available at: https://tracim.ptb.de/tracim/resources/downloads/ptbwhz_math_chebyshev/ptbwhz_math_chebyshev_manual.pdf (last access: 15 November 2021), 2016.