Publishable Summary for 17NRM05 EMUE
Advancing measurement uncertainty – Comprehensive examples for key international standards

Overview

This project provided a comprehensive set of worked examples illustrating how principles of measurement uncertainty evaluation can support documentary standards and guides. It has promoted uncertainty evaluation according to internationally recognized guides across broad disciplines of measurement. The project has delivered new or improved adaptable examples of and templates for uncertainty evaluation to the Joint Committee for Guides in Metrology (JCGM) as publishers of the internationally acknowledged Guide to the expression of uncertainty in measurement (GUM). It also provided examples to some ten standardisation bodies and other organisations that are specifically related to standards they are developing.

Need

Measurement models describe the relationship between (input) quantities we measure or know something about to the (output) quantities of interest (the measurands). In areas such as energy, environment and health care, these models are frequently non-linear and the quantities measured may have substantial uncertainty. In almost all scientific areas, uncertainties associated with the output quantities must be calculated given the input uncertainties. The traditional approach to uncertainty propagation through a model uses the GUM, but because of the above issues the resulting uncertainty so produced may not always be fit for purpose.

Carefully elaborated examples have been developed that are practical and covering many areas of measurement, capable of delivering reliable results, and as far as possible in a form that can be adapted to actual end-users' data and knowledge. Many end-users learn by example rather than from formal guidance material, using the guidance material to support what they have learned when needed.

The examples have been provided to meet end-users' needs, essential in diverse disciplines including traditional metrology (calibration, testing, comparison and conformance) and the sectors environment, energy, quality of life, and industry and society.

Objectives

The overall objective is to provide a comprehensive set of examples to illustrate uncertainty evaluation according to the GUM suite of documents. The specific objectives of the project are (note that guidance document JGCM 103 is now known as JCGM GUM-6):

1. To develop examples of measurement uncertainty evaluations capable of acting as template solutions that end users can use for related problems. Examples will include measurement model construction using JCGM 103, application of uncertainty evaluation principles for addressing industrial conformity assessments to support JCGM 106 and taking correlations into account as requested by ISO/REMCO, the ISO committee concerned with reference materials.
2. To derive worked examples of uncertainty analyses using the GUM and other methods to assist users to make informed choices on an appropriate uncertainty evaluation method to use. Examples will include an examination of the extent to which the GUM is appropriate for certain applications or whether the Monte Carlo methods of GUM Supplements 1 and 2, or Bayesian methods, have greater efficacy.
3. To collaborate with JCGM/WG1 (the chief stakeholder), and the standardisation, regulatory and accreditation communities (ISO/REMCO, IEC, CEN, OIML, and ILAC) to ensure that the outputs of the project are aligned with their needs, communicated quickly, and in a form that can readily be incorporated into the JCGM Guides and other documents.

Progress beyond the state of the art

Many practitioners, particularly in calibration and testing laboratories, admit having difficulty interpreting and applying measurement uncertainty principles as presented in the GUM suite of documents. They especially have difficulty in quantifying and accounting for correlation between quantities in a measurement model. To help overcome these problems, this project applied the 'learn by example' principle to aid such practitioners with many examples developed to illustrate the principle.

Testing laboratories have activities related to testing, inspection, and certification, which are critical in assuring safety of products and services and in market surveillance. Since conformity assessment according to the latest (2017) edition of the international standard concerned with general requirements for the competence of testing and calibration laboratories requires uncertainty to be used in criteria for decision rules, this project constructed examples illustrating how uncertainty can be accounted for in this area.

The examples in the existing GUM suite of documents have been criticised as not relating sufficiently strongly to modern practice in measurement, and to calibration and testing laboratories. This project has significantly contributed to extending the range and complexity of publicly available validated uncertainty examples beyond the previous simple measurement models and those expressed in a straightforward mathematical manner in the GUM.

The derived examples are supplemented by generic tutorial material, particularly to assist practitioners in setting up their own uncertainty evaluations. These tutorials focus on reporting measurement results, handling correlation, and Monte Carlo and Bayesian methods for uncertainty propagation.

Results (indicating prepared examples addressing the objectives)

A selection of the 42 examples developed is outlined. The full set is given in a compendium of examples.

Objective 1: Examples of measurement uncertainty evaluations capable of acting as template solutions

This project covered a broad range of uncertainty example applications and derived worked solutions for generic measurement uncertainty types.

Straight-line calibration: Common practice in calibration is determining a straight-line relationship between two quantities, where measured values of the quantities have uncertainties, as in the international technical specification ISO/TS 28037 on determination and use of straight-line calibration functions. The line can be obtained using appropriate least-squares methods and the associated uncertainty evaluated using the GUM, Monte Carlo or Bayesian methods. These approaches were compared on examples for calibrating a sonic nozzle, measuring haemoglobin concentration, and calibrating a torque measuring system. Known correlations were taken into account. ISO committee TC 69, Application of Statistical Methods, will be upgrading ISO/TS 28037 to full international standard in this area, taking into account these project outcomes.

An uncertainty determination for a complicated measurement model that was improved by this project was that of the Single Burning Item (SBI) test specified within European standard EN 13823. The objective of this test is to classify construction material in the context of the propagation of building fires. Expressions for the rates of heat release (HRR) and smoke production (SPR) during material burning under standardized test conditions are complex functions of the quantities on which they depend. The GUM method for uncertainty propagation was applied to typical data and validated using a Monte Carlo method to provide a worked example of best practice in estimating HRR and SPR and evaluating their associated uncertainties.

Support for calibration and measurement capabilities: National Metrology Institutes (NMIs) are responsible for maintaining traceability to the SI via capability comparison exercises organized by the BIPM. They are required to quantify their performance using calibration and measurement capabilities (CMCs), which are validated by key comparisons in which NMI laboratories compare their measurements of the same artefacts. Their CMC claims are then adjusted as necessary to enable them to be supported by a relevant key comparison. Following a request from the Director of the BIPM, this project has provided a robust new statistical method for making minimal adjustment to the claims and illustrated by an example of measurements of industrial reference gauge block standards.

Conformity assessment for multicomponent materials: The guidance given in JCGM 106, The role of measurement uncertainty in conformity assessment, was extended by this project to treat multicomponent materials involving several quantities. JCGM 106 is acknowledged as the only authoritative international guide in the area that adopts a Bayesian treatment. An example on influenza medication containing four components (pain reliever and fever reducer, cough suppressant, antihistamine, and nasal decongestant) was derived in

which each component had individual tolerances on its content. Risks to the supplier and customer were assessed, and the results incorporated in IUPAC/CITAC Guide, Evaluation of risks of false decisions in conformity assessment of a multicomponent material or object due to measurement uncertainty.

For the calculation of risks of false decisions in the conformity assessment of test results consistent with the provisions of JCGM 106, data were collected for total suspended particulate (TSP) mass concentration in the locality of three stone quarries. The collection was made according to an Environmental Protection Agency method for sampling of ambient air. The studies allowed the involved parties to agree on an acceptance limit to balance the safeguarding of the inhabitants' health and the economic interests of the quarries' owners.

The objective was successfully achieved in all respects.

Objective 2: Worked examples of uncertainty analyses using the GUM and other methods to assist users to make informed choices on an appropriate uncertainty evaluation method to use

This project applied GUM and Monte Carlo methods to uncertainty example applications and derived worked solutions for generic measurement uncertainty types. Some examples are given below as an indication of the project development across various measurement areas.

Environmental sector applications

Quantification of toxic contaminants in the environment: Monte Carlo and GUM methods were applied to evaluate uncertainties associated with the quantification of low masses of benzo[a]pyrene (BaP) spiked on filters commonly used for particulate matter sampling. European Directive 2004/107/EC, *Arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air*, lists BaP as a carcinogenic risk marker for all toxic contaminants in the environment. Uncertainty evaluation associated with the quantification of such micro-pollutants plays a key role in the reliability of their measurement. It was shown that the Monte Carlo methods gave more valid results because of the highly skewed probability distribution for BaP.

Comparison of GUM and IPCC approaches for greenhouse gas (GHG) emissions: Each EU country submits an annual report of its GHG emission estimates, under the Kyoto Protocol, based on contributions for a pollutant from many sources. An Intergovernmental Panel on Climate Change (IPCC) guide, *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*, gives GUM-like and Monte Carlo methods for uncertainty propagation. For actual data for a pollutant from the UK contribution to GHG emissions, the project used GUM and Monte Carlo methods in JCGM guidance documents to obtain the associated uncertainty. Lessons were drawn on the relative merits of the JCGM and IPCC approaches.

Risk and performance analysis for water supplies: Clean water and sanitation are among the sustainable development goals of the United Nations (UN), as in the UN document 'Transforming our world: The 2030 agenda for sustainable development'. Water supply networks involve net balances based on water inflows and outflows from metering equipment in many locations. It was shown that measurement uncertainty is a valuable tool to support the analysis of performance and risk related to these utilities. An example in this area provided simple formulae for the total volume or flow rate at a point in the network and for the uncertainty in terms of the number of measurement sites.

Energy sector applications

Quantification of industrial flow meters: Analysis of a database containing some 16 000 values was used to deduce the uncertainty associated with the orifice discharge-coefficient according to an industry-accepted (Reader-Harris/Gallagher) equation. It is estimated that at least 40 % of industrial flow meters in use are differential pressure-based devices, with the orifice plate being the most popular for accurate measurement of fluid flow. The GUM was used to evaluate the uncertainty, taking all uncertainty sources into consideration. The analysis substantiated previous results, increasing confidence in the use of international standard ISO 5167-2 concerned with the measurement of fluid flow by means of pressure differential devices.

Understanding power transformer losses: European regulations such as Directive 2009/125/EC, *Establishing a framework for the setting of ecodesign requirements for energy-related products*, set requirements regarding energy losses, which play a key role in energy production, transformation, distribution and consumption. Costs of losses generated by power transformers are comparable to product costs and play an important role in evaluating total costs. The example estimated energy losses for typical data and evaluated the associated uncertainty by applying GUM and Monte Carlo methods. These results indicated the reliability of the estimates obtained with an impact on billing.

Improved knowledge of thermal comfort in buildings: International standard ISO 7730 concerned with ergonomics of the thermal environment provides an expression for the degree of satisfaction with that environment. The average thermal sensation of a large group of people in the same environment, quantified by the predicted mean vote (PMV), based on data collected in a controlled climate chamber, is used to quantify thermal comfort. PMV was calculated for typical scenarios using a Monte Carlo (MC) method. The resulting probability distribution was far from normal, justifying the use of MC, the practical repercussions of which have been communicated to the responsible international committee.

Health sector applications

Analysis of medical imaging in cancer diagnosis and treatment: Pixelated images are used in medical imaging to display a section of a tumour, the area of which is estimated by determining the pixels through which the tumour boundary passes. For specific images, the project used Fourier series to reconstruct the boundary of the tumour as a smooth curve, with the number of Fourier harmonics chosen to balance fidelity and smoothness. The area of the section was estimated as that of the region enclosed by the Fourier curve so obtained. The uncertainty associated with this area was evaluated using GUM and Monte Carlo approaches.

Monitoring properties of biological tissues: An imaging technique – magnetic resonance-based electrical properties tomography (EPT) – monitors properties of biological tissues, used as cancer biomarkers. The project determined uncertainties for estimated electrical conductivity based on the analysis of experimental EPT data collected through a magnetic resonance imaging (MRI) scanner using a homogeneous cylindrical phantom. This process required statistical evaluation of repeated MRI phantom scans to evaluate the covariance matrix associated with the EPT input, and then propagated that information through the EPT measurement model using GUM principles.

Critique of guidance for storage of blood products: It is often not possible to establish whether regulations and standards in cold storage systems for blood products or the decontamination of medical devices are being met in a robust fashion. Such determinations are required for accreditation to international standards such as ISO/IEC 17025 and ISO 15189 concerned with requirements for the competence of testing and calibration laboratories and for quality and competence of medical laboratories. Since current guidance is open to interpretation, the examples developed indicated how such documents might reasonably be interpreted. The advice given was well received by Scottish Health, which is making it more widely available.

Industrial sector applications

Non-invasive method for estimating turbofan jet engine nozzle thrust: Acoustic sensor systems for measuring nozzle thrust, jet velocity and other quantities in a jet engine are more resistant to harsh conditions (high temperatures, pressures, etc.) than other measuring systems. For a geometrical configuration of sensors used in practice, estimates of these quantities were obtained for representative test data from Rolls-Royce, and for the first time, a full uncertainty analysis for such a set-up carried out. Further, it was shown that by modifying the sensor configuration geometry, uncertainties could be reduced by 60 %.

Hardness verification in material testing: The project investigated uncertainty evaluation in hardness verification where repetition of tests at the same location is generally not possible. To improve methods provided by international standards such as the Vickers, Knoop, Rockwell and Brinell hardness tests, a harmonized GUM-compliant approach to uncertainty evaluation was developed for application to all hardness tests irrespective of test type. On typical example data sets, the GUM was used for the uncertainty evaluation and the results confirmed using the Monte Carlo method. The international committees responsible for six hardness-testing standards will take account of the work in future editions of these standards.

The objective was fully achieved in all respects.

Objective 3: Collaborate with JCGM/WG1 (the chief stakeholder), and the standardisation, regulatory and accreditation communities to ensure that the outputs of the project are aligned with their needs

All examples and the tutorial material generated within this project have contributed to this objective. In the above text under Objective 1 and Objective 2, some relevant standards, guides and regulations were identified. Further information is given under ‘Impact on relevant standards’ below. A 620-page compendium containing all worked examples and tutorial material was provided to JCGM/WG1 and relevant chapters from the compendium to international standards committees. Further, several standards committees are holding or will

hold meetings to discuss in which ways parts of the compendium will influence the standards they are developing or revising.

The objective was successfully achieved in all respects.

Researcher Mobility Grant

Facilitated by an EMPIR research mobility grant, an Institute of Metrology of Bosnia and Herzegovina researcher developed three examples for measurement uncertainty evaluation in the fields of pressure and gas flow, receiving input from the UK and Dutch national metrology institutes, the United Kingdom Accreditation Service (UKAS), and Sarajevogas from Bosnia and Herzegovina. A main goal was to compare approaches for uncertainty evaluation for the same measurement model. The examples described (1) the calibration of a gas flow measuring instrument by the 'master meter' method, (2) the pressure drop due to gas leakage in a pressurised vessel, and (3) the preparation of calibration gas mixtures of ammonia in nitrogen using permeation. Monte Carlo was applied to two of the examples and validated the GUM approach. For example (3), current practice was improved by an improved modelling of the temperature dependence on the permeation, and other contributions including the calculation of mass flow from weighing data. A Bayesian treatment of the determination of the repeatability and finite resolution contribution of the weighing was developed to model these intertwined effects.

Impact

Over 20 presentations were made on project activities at international conferences. The compendium containing 42 uncertainty examples is freely available on the website. To support these documented examples, a 2-day workshop was held in January 2020 and a 2-day online measurement uncertainty training course in March 2021. The training course was open to all but particularly targeted at Western Balkan countries. It assisted in empowering training courses run by LNE, NPL, PTB and UKAS. An unforeseen benefit of the Coronavirus pandemic was that the course was held online, which permitted an attendance of over 160 rather than the envisaged 25. These activities were extremely helpful in extending the understanding of measurement uncertainty evaluation to a wider circle of end-users. Training was also delivered at eight smaller events. There will be specific longer-term impact in many areas including greenhouse gas emission inventories, energy efficiency and thermal comfort in buildings, neonatology and cancer treatments, and doping tests. The project website includes information on the scope and objectives of the project, and the compendium, as well as uploaded presentations and links to open-access papers published as part of the project.

Four popular articles were published to bring the outputs of the project to a wider audience:

A. Ribeiro and M.G. Cox, EUROLAB supports the EMPIR Special Project "Examples of Measurement Uncertainty Evaluation" (EMUE), EUROLAB Newsbriefing, April 2021.

M.G. Cox, EMPIR project EMUE Examples of Measurement Uncertainty Evaluation, ILAC Newsletter, April 2021.

Francesca R. Pennecchi, Michela Sega and Adriaan van der Veen, EMPIR PROJECT EMUE: EXAMPLES OF MEASUREMENT UNCERTAINTY EVALUATION, CITAC News, April 2021, pages 55-57.

A. Carullo, S. Corbellini and A. Vallan, Misura di temperatura con scheda a micro-controllore (Temperature measurement with micro-controller based board), Tutto Misure https://www.tuttomisure.it/, Issue 2/21, pages 19-25.

Impact on industrial and other user communities

Work on the uncertainty analysis related to the thrust generated by a turbofan jet engine (and other measurands) is regarded as ground-breaking by Rolls-Royce and their academic subcontractor Virginia Tech with substantial improvements possible in the uncertainty associated with exhaust thrust, exhaust temperature and other quantities.

A study of an anti-doping laboratory's evaluation of the uncertainty associated with the measured concentration of the stimulant ephedrine at levels close to the threshold in urine samples was completed. That stimulant is prohibited for use in sport. The evaluation was verified through the laboratory's participation in three rounds of the World Anti-Doping Agency's (WADA) proficiency testing programme. The study concluded with the acceptance of the evaluated standard uncertainty in accordance with WADA requirements.

An example on electric property tomography, a quantitative imaging method based on magnetic resonance imaging (MRI), has been developed in cooperation with Philips, a leading manufacturer of MRI scanners. Philips has been given access to the outcomes including the full quantification of the repeatability of the results.

The work on mobile optical measurement systems (MOMS) and on measurement uncertainty in a multiplexed data-acquisition system has been communicated to practitioners in the automotive and aeronautical industries and in the electronic industries.

Impact on the metrology and scientific communities

The compendium of 42 examples produced in the project has been made available in its entirety to the JCGM, the material within it relating to the further development of the JCGM guides on conformity and interlaboratory comparisons and the JCGM examples document. In addition, the compendium contains supporting tutorial material on Monte Carlo, Bayesian statistics, reporting measurement results, and handling correlation. JCGM/WG1 has formed a sub-group to decide the parts of the compendium to include in three of its guidance documents.

It is expected that use will be made of the compendium, which is fully indexed for reference purposes, by the many metrology and scientific sectors addressed by the project. Its availability through the JCGM will ensure considerable use noting that the number of downloads of the JCGM guides is some 5 000 per month. It is anticipated that the number of downloads of the compendium will be comparable.

Impact on relevant standards

The project made input to over ten national and international standards committees, many of which have made statements of need for improved examples, and to Eurachem, European Accreditation, EUROLAB and UKAS. Several partners are members of ISO/TC 69, the international standards committee concerned with the application of statistical methods, to which presentations on project progress have been made at its annual plenary meetings and its uncertainty working group. ISO/TC 69 is engaged on many standards involving statistical methods for uncertainty evaluation and already technology developed within the project is having an impact, The committee intends to make use of the project's results in upgrading one of its technical specifications, on straight-line calibration, to full international standard.

The coordinator of the project is the liaison officer between ISO/TC 69 and the JCGM. He has regularly reported on JCGM activities, especially those activities to which this project makes input, at TC 69 plenary meetings.

Several presentations have been made to national and international standards committees, which in many cases will have impact on the standards they are producing

Specific input made to standards bodies included an example relating to gas analysis, an example contained in a guide on analytical chemistry for evaluation of risks of conformity assessment of a multicomponent material or object due to measurement uncertainty.

Longer-term economic, social and environmental impacts

In implementing methods of uncertainty evaluation, many practitioners benefit from learning from example. By providing over 40 worked examples in diverse areas of application and in the energy, environment and health care sectors, and in industry and society, in a categorized 600-page compendium, the project will assist end-users in this way. Since correlation between quantities is a topic that many find difficult to understand and handle, the considerable emphasis placed on this topic will be of particular value to the community. The training-course material provided and delivered in the project will be of much interest to many organizations. Particular benefit will accrue to some ten international standards' committees to which project outputs have been provided; many of these committees will use that material in improving their current standards and in developing new standards that take measurement uncertainty into consideration. In particular, the Joint Committee for Guides in Metrology, responsible for the GUM suite of documents, will use a sizeable fraction of the material in the compendium in its guidance documents.

List of publications

J.A Sousa, E. Batista, O Pellegrino, A S Ribeiro, L L Martins. Method selection to evaluate measurement uncertainty in microflow applications, Journal of Physics: Conf. Series, 1379 (2019), 012033 (doi: 10.1088/1742-6596/1379/9/012033).

A.M.H. van der Veen, M.G. Cox. Getting started with uncertainty evaluation using the Monte Carlo method in R. Accred Qual Assur 26, 129–141 (2021). https://doi.org/10.1007/s00769-021-01469-5

A.S. Ribeiro, M.C. Almeida, M.G. Cox, J.A. Sousa, L. Martins, D. Loureiro, R. Brito, M. Silva, A.C. Soares. Role of measurement uncertainty in the comparison of average areal rainfall methods, Metrologia 58 (2021), 044001, https://doi.org/10.1088/1681-7575/ac0d49

J. Pétry, B. De Boeck, N. Sebaïhi, M. Coenegrachts, T. Caebergs, M. Dobre. Uncertainty evaluation in atomic force microscopy measurement of nanoparticles based on statistical mixed model in a Bayesian framework, Meas. Sci. Technol. **32** (2021) 085008, https://doi.org/10.1088/1361-6501/abe47f

J.A. Sousa, E. Batista, S. Demeyer, N. Fischer, O. Pellegrino, A.S. Ribeiro, L.L. Martins, Uncertainty calculation methodologies in microflow measurements: Comparison of GUM, GUM-S1 and Bayesian approach, Measurement 181 (2021), 109589, https://doi.org/10.1016/j.measurement.2021.109589.

This list is also available here: https://www.euramet.org/repository/research-publications-repository-link/

Project start date and duration:	1 July 2018; 36 months	
Coordinator: Dr Maurice Cox, NPL, United Kingdom Tel: 0044 20 8943 6096 E-mail: maurice.cox@npl.co.uk Project website address: http://empir.npl.co.uk/emue/		
Chief Stakeholder Organisation: Joint Committee for Guides in Metrology – Working Group 1 (GUM)	Chief Stakeholder Contact: Dr Walter Bich	
Internal Funded Partners: 1 NPL, United Kingdom 2 BAM, Germany 3 IMBiH, Bosnia and Herzegovina 4 INRIM, Italy 5 IPQ, Portugal 6 LGC, United Kingdom 7 LNE, France 8 NEL, United Kingdom 9 PTB, Germany 10 SMD, Belgium 11 VSL, Netherlands	External Funded Partners: 12 ACCREDIA, Italy 13 LNEC, Portugal 14 UKAS, United Kingdom	Unfunded Partners: 15 AIST, Japan 16 RR, United Kingdom 17 WADA, Switzerland
RMG1: IMBiH, Bosnia and Herzegovina (Employing organisation); NPL, United Kingdom (Guestworking organisation)		