NPL
National Physical Laboratory

Report

NPL Report
DEM-ES-025

Statistics standards to support metrology 2006–07

Maurice G Cox

NOT RESTRICTED

March 2007

dti
FUNDED BY THE DTI

National Physical Laboratory | Hampton Road | Teddington | Middlesex | United Kingdom | TW11 0LW

Switchboard 020 8977 3222 | **NPL Helpline** 020 8943 6880 | **Fax** 020 8943 6458 | **www.npl.co.uk**

dti

The DTI drives our ambition of 'prosperity for all' by working to create the best environment for business success in the UK. We help people and companies become more productive by promoting enterprise, innovation and creativity.

We champion UK business at home and abroad. We invest heavily in world-class science and technology. We protect the rights of working people and consumers. And we stand up for fair and open markets in the UK, Europe and the world.

The National Physical Laboratory is operated on behalf of the DTI by NPL Management Limited, a wholly owned subsidiary of Serco Group plc

Statistics standards to support metrology 2006–07

Maurice G Cox
Mathematics and Scientific Computing Group

March 2007

ABSTRACT

One of the projects within the third UK Software Support for Metrology programme (SS*f*M-3) is concerned with standards activity relating to the Uncertainties theme in that programme. This report summarizes the work on that project carried out within the third and final year (2006–07) of SS*f*M-3.

© Crown copyright 2007
Reproduced with the permission of the Controller of HMSO
and Queen's Printer for Scotland

ISSN 1744–0475

National Physical Laboratory,
Hampton Road, Teddington, Middlesex, United Kingdom TW11 0LW

Extracts from this report may be reproduced provided the source is acknowledged and the
extract is not taken out of context

We gratefully acknowledge the financial support of the UK Department of Trade and
Industry (National Measurement System Directorate)

Approved on behalf of the Managing Director, NPL
by Jonathan Williams, Knowledge Leader for the Electrical and Software team

Contents

1 Introduction

This report describes standards activity relating to the Uncertainties theme within the third Software Support for Metrology programme (SS*f*M-3) [46]. The relevant project within that theme has the following objectives:

1. To bring SS*f*M expertise in mathematics and statistics, particularly best practice in uncertainty evaluation and related mathematical and statistical modelling, to benefit relevant metrology standards development;
2. To achieve more consistent application of best practice in the guidance, algorithms and procedures standardized for uncertainty evaluation in different contexts;
3. To keep everyone interested in the programme aware of the progress made in the standards committees.

In each of the three years the project is concerned with

1. active participation in relevant BSI, CEN, EA, ISO and JCGM standards committees, particularly focused on uncertainty evaluation,
2. providing an annual report on progress made in these committees, and
3. producing an annual awareness article on the progress made in the committees.

This document constitutes the third annual report, generally subdivided according to the activity carried out by the above standards committees. There was no involvement by SS*f*M within Comité Européen de Normalization (CEN) or with the European co-operation for Accreditation (EA) during the third year of the programme. Advantage was taken to devote greater effort to the ISO and JCGM activity, in which there is increasing demand for UK input.

2 British Standards Institution

2.1 General

SS*f*M provides representation on British Standards committee SS/6, Precision of Test Methods, and SS/6 Panel 3, Measurement Uncertainty, as convenor and secretary. Previously, two meetings of this committee and two meetings of this panel were held each year. To avoid duplication in these meetings as a consequence of considerable commonality of interest, and because the membership in each case is almost identical, it was agreed that two combined meetings would be held each year.

The business in the two combined meetings in 2006–07 included the following items:

- Work on straight-line calibration functions (section 2.2);
- Consideration of the following Standards, Technical Specifications, and Guides:
  - BS 5233, Terms used in metrology [13];

- BS 5702, Guide to statistical process control charts for variables – Part 2: Charts for individual values [14];
- IEC 62387, Passive integrating dosimetry systems for environmental and personal monitoring for external photon and beta radiation using electronic devices for the data evaluation – Part 1: General characteristics and performance requirements for dosimetry systems [24];
- ISO 5725, Accuracy (trueness and precision) of measurement methods and results —Parts 1–4, 6 [26, 31, 32, 33, 37];
- ISO 10576-1, Statistical methods — Guidelines for the evaluation of conformity with specified requirements — Part 1: General principles [34];
- ISO 11843-1, Decision limit, detection limit, capability of detection—Part 1: Terms and definitions [30];
- ISO 11843-5, Capability of detection—Part 5: Methodology in the non-linear calibration case [39];
- ISO/TS 21748, Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation (section 3.2) [35];
- ISO/TS 21749, Measurement and uncertainty for metrological applications—repeated measurements and nested experiments (section 3.3) [38];
- ISO Guide 35, Certification of Reference Materials – General and Statistical Principles [25];
- ISO Guide 98, Guide to the expression of uncertainty in measurement (GUM) [27];
- ISO Guide 99, International vocabulary of basic and general terms in metrology (VIM) [29].

2.2 Straight-line calibration functions

NPL and Eurachem (http://www.eurachem.ul.pt) carried out work under the title *The determination and use of straight-line calibration functions*, support for which was provided under the BSI/DTI Consultancy Drafting Scheme. The objective was to produce a draft ISO Standard in the form of a Technical Specification (TS) that provided guidance on

1. the determination of straight-line calibration functions for commonly encountered measurement uncertainty structures associated with the calibration data,
2. the use of a calibration function so determined to provide predicted values and the uncertainties associated with these values, and
3. validating that the function provides a consistent description of the calibration data.

The work carried out under the BSI/DTI contract was monitored by a steering panel with representatives from LGC Ltd., the Food Standards Agency, and PRI/23, Test methods for rubber and non-black compounding ingredients (within the British Standards Materials and Chemicals Committees).

Appendix A contains the scope of this work.

2.3 Other standards

Work on other Standards, Technical Specifications and Guides constituted document review, administering the UK vote on ISO and IEC New Work Item proposals and final drafts, and tracking related work by other organizations such as EUROLAB and JCGM.

3 International Standardization Organization

3.1 General

SS*f*M provides the convenorship for ISO/TC 69/SC 6/WG 7. ISO/TC 69 is the ISO Technical Committee concerned with Applications of Statistical Methods. SC 6 is the Subcommittee of ISO/TC 69 entitled Measurement methods and results. WG 7 is the Working Group of SC 6 entitled Statistical methods to support measurement uncertainty evaluation.

SS*f*M also provides the convenorship for a new Ad hoc Group on Statistical software and standards established at the June 2006 meeting of ISO/TC 69.

An annual five-day meeting of ISO/TC 69 is held, and, optionally, an interim meeting between adjacent annual meetings.

The Annual Meeting within the period covered by this report was held in Sun City, South Africa in June 2006 and an Interim Meeting in Paris in January 2007.

Relevant matters addressed at the plenary sessions of the June 2006 Annual Meeting were the following:

1. The degree of usability of the documents produced. In particular, as previously reported [16], statistical standards should be made more readable to non-statisticians;
2. The promotion of ISO/TC 69 work. A roadmap and a related presentation are available on the ISO/TC 69 website;[1]
3. Software support for statistical standards. A number of existing standards and standards being developed could benefit from the availability of relevant software;
4. The extension of the standards produced by ISO/TC 69 to cover measurement uncertainty considerations.[2]

Main agenda items at both the June 2006 and the January 2007 meetings relating to the activities of ISO/TC 69/SC 6/WG 7 and the Ad hoc Group on Statistical software and standards were

1. ISO/TS 21748, Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation (section 3.2),
2. ISO/TS 21749, Measurement uncertainty for metrological application—Repeated measurements and nested experiments (section 3.3),

[1] `http://isotc.iso.org/livelink/livelink?func=llworkspace`, and search for TC 69.

[2] ISO/TC 69/SC 5, Acceptance sampling, has already identified such needs in its area. Consideration would be required on the manner in which the necessary interaction between ISO/TC 69/SC 6/WG 7 and the SCs of ISO/TC 69 would be managed.

3. the determination and use of straight-line calibration functions (section 3.4), and
4. guidance for the developers of statistical standards that are to be supported by software (section 3.5).

In addition, consideration was given at the Annual Meeting to the standards listed in section 2.1 and the standard ISO 5725-5, Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method [36].

Prior to the Annual Meeting, ISO/TC 69 had prepared a chart showing the standards it has produced and their interrelationship. A copy of this chart is given as figure 1. ISO standards relevant to this report can be identified there.

3.2 ISO/TS 21748, Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation

A New Work Item proposal was submitted to ISO, together with the current TS as a supporting document, regarding the progression of ISO/TS 21748, Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, to a full standard. A project leader was appointed from LGC Ltd.

The new document title would be *Use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*.

The major concentration in the full standard would be on the provision of a number of practical examples to illustrate the principles of the TS, which, it was agreed, are well expressed there. It is anticipated that the examples would relate to

1. chemical measurement,
2. microbiology, such as plate counting,
3. reproducibility estimates for uncertainty evaluation,
4. the interests of EUROLAB (National Associations of Measurement, Testing and Analytical Laboratories), perhaps concerning aggregates,
5. the interests of ASTM (American Society for Testing and Materials).

3.3 ISO/TS 21749, Measurement uncertainty for metrological applications—Repeated measurements and nested experiments

As for ISO/TS 21748, a New Work Item proposal was submitted to ISO, together with the current TS as a supporting document, regarding the progression of ISO/TS 21749 to a full standard. A project leader was appointed from NIST.

The document title would become *Measurement uncertainty for metrological application—Nested experiments*.

It was considered that the following work was required in developing the full standard:

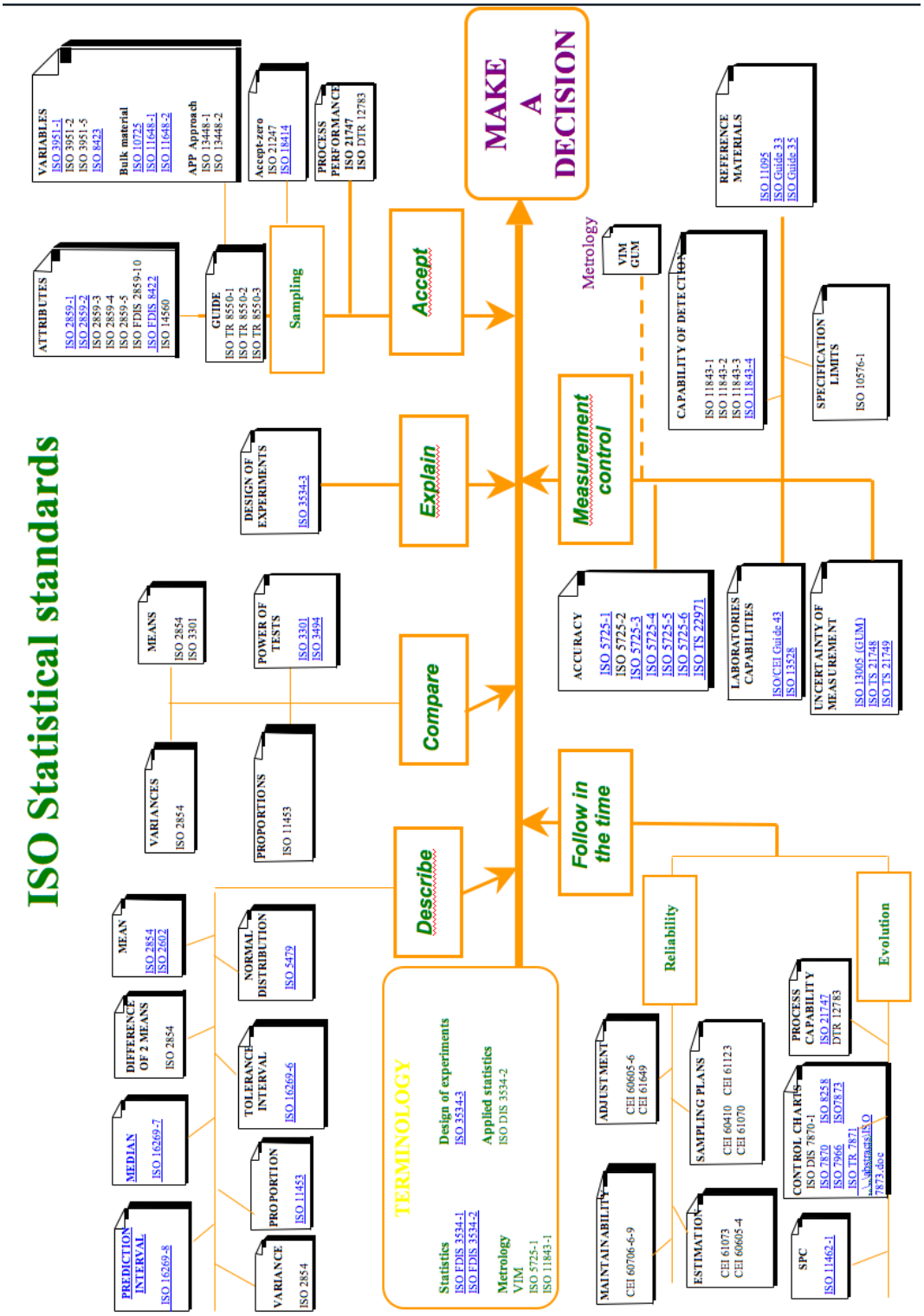

Figure 1: **ISO statistical standards.**

1. Extension of the scope to cover unbalanced designs. Although the document would not cover *strongly* unbalanced designs, material would be added to indicate how, with caution, cases can be handled where there is a small number of missing data. The document would indicate methods (including linear least squares), software and references;

2. Careful attention to the determination of degrees of freedom and the use of the Welch-Satterthwaite formula;

3. A generic way of summarizing an n-level design. The presentation in the TS could then be simplified by giving the generic form, followed by its particularization to 1-level, 2-levels, etc. The relevant formulae would be included in a normative annex;

4. Appraisal of the current material on bias, which was judged not entirely satisfactory. Technical consideration would be given to this issue in readiness for the June 2007 meeting of ISO/TC 69/SC 6/WG 7;

5. Removal of the material on Type B uncertainty evaluation. Since the document is primarily concerned with Type A uncertainty evaluation, the retention of such material is questionable;

6. The inclusion of several new examples to augment or replace the single example in the current TS. It is expected that examples would relate to

   (a) the textile industry (Dupont),

   (b) the determination of creatinine in serum,

   (c) environmental analysis, involving the use of a reference material for validation purposes,

   (d) the interests of UTAC (Union Technique de l'Automobile du Motocycle et du Cycle), France,

   (e) the interests of NPL, and

   (f) the interests of the Tokyo University of Science, Japan.

3.4 The determination and use of straight-line calibration functions

Based on the NPL-Eurachem work (section 2.2), a New Work Item proposal concerning the determination and use of straight-line calibration functions and an accompanying draft of the proposed TS on the subject were prepared, and distributed by ISO for voting in May 2006.

There was a detailed review at the Annual and Interim meetings of the drafts of the TS produced by NPL and Eurachem, account of which was taken of the first review in the work under BSI/DTI contract (section 2.2).

The five participating countries required to progress the work on the TS, should the voting on the New Work Item proposal be positive (which it was), were agreed as France, Japan, Malaysia, UK and US. An expert was nominated from each of these countries. The work will involve developing a Committee Draft (CD) of a TS based on the NPL-Eurachem work.

NPL provided to ISO/TC 69/SC 6/WG 7 a comparison of the scope of this draft TS and ISO 11095 [28], directed at straight-line calibration functions in the context of reference materials.

The next revision of the TS will take account of the following aspects:

1. A better balance regarding a metrological and statistical treatment. The problem addressed is statistically one of "errors in variables" (EiV), whereas the document is more aligned with the GUM: an input-output model with estimates of the input quantities corresponding to the xy-data, with associated uncertainties specified. An exception relates to cases where no uncertainty information is given, but an assumption made, e.g., that the y-data values have identical associated uncertainties. Such cases are those primarily considered in EiV models;

2. Improvement in document usability. Each part of the document dealing with one particular uncertainty structure would be re-presented as a statement of the problem addressed, followed by a recommended solution procedure, with appropriate references to scientific texts and supporting software;

3. Extension of the scope to the case where the uncertainty matrix associated with the x-data is full and there is a common unknown uncertainty associated with the y-data. Such cases are relatively common, and include, for instance, calibration problems where the x-data relates to concentrations obtained by a successive dilution process, and the (constant) standard uncertainty corresponding to the responses has not been evaluated *a priori*;

4. Expansion of the scope to cover calibration functions that can be transformed to straight lines, for example $Y = Ce^{DX}$ or $Y = F + G \ln X$. The manner in which relevant uncertainties would be influenced would be included, as would a table of useful transformations. Consideration would be given to whether such material should be incorporated in the revised document, or contained in a subsequent standard that covered (a) non-linear calibration functions including the above, and (b) polynomial curves;

5. Examples that would be considered include the following:

   (a) Isotope dilution mass spectrometry for pesticide analysis;

   (b) Polymerase Chain Reaction (PCR);[3]

   (c) X-ray fluorescence using independent reference standards.

At least two of these examples would relate to the extended scope in 3.

3.5 Statistical software and standards

The new ISO/TC 69 Ad hoc Group on Statistical software and standards is concerned with the preparation of guidance for the developers of statistical standards that are to be supported by software. Appendix B gives the scope of the work.

[3]PCR is a technique for rapidly producing many copies of a fragment of DNA for diagnostic or research purposes.

The first meeting of the Group was held in Paris in January 2007 as part of the Interim meeting of ISO/TC 69. It was attended by ten representatives, from France, Japan, UK, and USA. The UK presented a document constituting a review of relevant statistical standards and supporting software. This document was discussed in detail and valuable input received concerning revisions to this document, and also concerning a first version of guidelines in this regard for standards' developers, to be produced jointly by NPL and LGC Ltd.

3.6 Other activity

Advice was provided to the International Electrotechnical Commission (IEC) on the possible use of ISO 10576-1 [34], one of the standards produced by ISO/TC 69, in developing a new standard, IEC 62387-1[24], concerned with dosimetry systems for environmental and personal monitoring. The standard ISO 10576-1 is concerned with general tests for conformity with specification limits, whereas IEC 62387-1 works with a one-sided confidence interval, as a result of the nature of the measurand. It was concluded that ISO 10576-1 does not apply, and therefore that IEC 62387-1 should be developed independently of it, and advice to this effect given accordingly.

A proposal was made at the Interim Meeting of ISO/TC 69/SC 6/WG 6 in Paris in January 2007 that a general document be developed that summarized all appropriate methods for measurement uncertainty evaluation, referring to supporting standards and guides for each approach. It was stated that many non-metrologists do not know how to evaluate uncertainty, and, in particular, those involved with interlaboratory studies need to know how to use the results obtained from the studies for this purpose. The document would help to decide the best approach, given the objective and the information available. The Standard ISO 17025 [40], which states that laboratories shall use appropriate methods for evaluating measurement uncertainty, is relevant. LNE (Laboratoire National d'Essais, France) would prepare a draft proposal for comment by ISO/TC 69/SC 6/WG 7. The proposal would then be submitted to JCGM Chairman for consideration. A document being finalized by EUROLAB is relevant and would be consulted.

4 Joint Committee for Guides in Metrology

4.1 Background

The Joint Committee for Guides in Metrology (JCGM) has amongst its tasks (a) the promotion of the "Guide to the expression of uncertainty in measurement" (GUM) [12], and the preparation of further documents for its broad application, and (b) promotion of the use of the "International vocabulary of basic and general terms in metrology" (VIM). Task (a) is the responsibility of Working Group 1 of JCGM (JCGM/WG1) and task (b) that of Working Group 2 (JCGM/WG2). SS*f*M has representation on JCGM/WG1.

The documents produced by JCGM are prepared by the eight Member Organizations of

the JCGM[4] and experts from a number of the National Metrology Institutes (NMIs)[5].

Normally, two meetings are held of JCGM/WG1 each year. Exceptionally, because meetings are not always held exactly six months apart, only one meeting was held within the reporting year (April 2006–March 2007), namely in October 2006. The previous meeting was held in March 2006 and reported on previously [16]. The next meeting will be held in April 2007.

Items on which work was carried out in the reporting period are covered in the following sections.

4.2 The GUM

4.2.1 General

The GUM is considered in the work of the JCGM as the primary document for evaluating uncertainty of measurement. Since its introduction in 1993, with corrections and reprinting in 1995 [12], it has provided the basis for this key aspect of metrology, and will continue to do so.

JCGM/WG1 is developing an electronic GUM (section 4.2.2), a revised GUM (section 4.2.3), and Supplements and other documents to support the GUM (section 4.3).

4.2.2 Electronic GUM

The electronic version of the GUM, which at the time of writing is being prepared by ISO, will add a dimension to the use of that Guide. The current printed version [12] contains an excellent index, but the ability to search for key words and phrases in the electronic form is welcomed. The electronic version will constitute a freely available internet-based document that textually is essentially identical to the printed version, but with corrections made by JCGM to formulae, tables and typographical errors. A list of editorial corrections of this nature was agreed at the October 2006 meeting of JCGM/WG1 and through JCGM submitted to ISO.

4.2.3 GUM revision

Revision of the GUM itself would commence in 2007. This revision would place the GUM on a more consistent probabilistic footing, for example reconciling the current different bases for Type A and Type B evaluations of uncertainty by applying the concepts of Bayesian probability theory [3, 23, 41].

Plans for a revised GUM were made by JCGM/WG1 at its October 2006 meeting and submitted to JCGM for consideration. The target date for finalizing the work on the revised GUM is 2011.

[4]Bureau International des Poids et Mesures (BIPM), International Electrotechnical Commission (IEC), International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), International Organization for Standardization (ISO), International Union of Pure and Applied Chemistry (IUPAC), International Union of Pure and Applied Physics (IUPAP), International Organization of Legal Metrology (OIML), International Laboratory Accreditation Cooperation (ILAC).

[5]DFM, INRIM, NIST, NPL, PTB.

4.3 GUM Supplements and other documents

4.3.1 General

In addition to a revised GUM, documents planned by JCGM/WG1 at the time of writing have the following titles, under the banner *Evaluation of measurement data*:

- Supplement to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method [10];
- Supplement to the "Guide to the expression of uncertainty in measurement" — Models with any number of output quantities [9];
- Supplement to the "Guide to the expression of uncertainty in measurement" — Modelling [8];
- An introduction to the "Guide to the expression of uncertainty in measurement" and related documents [5];
- Concepts and basic principles [7];
- The role of measurement uncertainty in deciding conformance to specified requirements [11];
- Applications of the least-squares method [6].

Scopes for the documents are given in the annexes to a sister report to this report, namely that for 2005–06 [16]. Progress on these documents is reported in the sections below.

4.3.2 Propagation of distributions using a Monte Carlo method

A final committee draft of the first Supplement to the GUM, concerned with the propagation of distributions [10], work led by NPL, was produced during the year. It took account of an exceedingly detailed and most valuable editorial review undertaken by the ISO representative on JCGM/WG1. Technically, this Supplement remains strongly based on SSfM Best Practice Guide No. 6, Uncertainty Evaluation [20], and an SSfM document giving software specifications for uncertainty evaluation [19]. In turn, that document is evolved accordingly, and thus benefits from international input.

This committee draft was provided to JCGM Chairman in October 2006 and through him to General Directors and Presidents of the Member Organizations of the JCGM for approval and editorial comment prior to publication. At the time of writing, most responses have been received. Following account of these reviews, JCGM/WG1 will make any necessary editorial modifications at its April 2007 meeting, and proceed with publication.

4.3.3 Models with any number of output quantities

A revised version of the second Supplement to the GUM, concerned with models having more than one output quantity [9], work led by NPL, was presented to JCGM/WG1 at its October 2006 meeting. The new version contained much general improvement, including closer alignment with the presentation of the first Supplement, and additional technical

material, particularly concerned with coverage intervals and coverage regions. As with the first Supplement, the experience gained in previous SS*f*M programmes is proving invaluable in this work, especially on more general models and on the implementation of the propagation of distributions using a Monte Carlo method.

4.3.4 Modelling

The Supplement concerned with modelling has lower priority than the other areas, and no work was carried out on it in the reporting period.

4.3.5 Introductory document

A new version of the Introductory document was presented to JCGM/WG1 at its October 2006 meeting. This document, the development of which is led by NPL, provides an introduction to measurement uncertainty, the GUM, and documents that support the GUM. The latest version includes new figures and flow diagrams. An advanced draft will be discussed at the April 2007 meeting of JCGM/WG1. It is expected that the document will be finalized and that circulation will start soon after that meeting.

4.3.6 Concepts and basic principles

Work on the supporting document concerned with concepts and basic principles is being progressed by IEC, NIST and PTB. That document, which underpins the other documents being produced by JCGM/WG1, covers information-based probability, modelling of knowledge by probability density functions (PDFs), assignment of PDFs, summarizing a quantity by the parameters of the PDF for the quantity, the (measurement) model of evaluation, and the use of Bayes' theorem to update a PDF given new information. Drafts of the document were reviewed in detail at the October 2006 meeting of JCGM/WG1. A new version will be considered at the April 2007 meeting. Sadly, Professor Gaetano Iuculano, the IEC representative on JCGM/WG1 and who provided valuable input to the work, died in January 2007.

4.3.7 The role of measurement uncertainty in deciding conformance to specified requirements

Work on measurement uncertainty in deciding conformance to specified requirements, led by NIST, was not significantly progressed in the reporting year due to other priorities. Since the JCGM Chairman has stated that the area is becoming more pressing, attention will be devoted to it at and after the April 2007 meeting of JCGM/WG1.

4.3.8 Application of the least-squares method

The document concerned with the application of the least-squares method has lower priority than the other areas, and no work was carried out on it in the reporting period.

4.3.9 Supporting software

Further work has been carried out by NPL on software to support the first GUM Supplement. A previous version of this software was demonstrated to JCGM/WG1 in March 2006 [16]. The new version concentrates on usability of the software. As stated [16], the software will allow all examples in the Supplement to be run with parameters (input values and uncertainties, PDFs, correlation, coverage probability) selected by the user through an intuitive and simple interface. The software would subsequently be made available through a suitable website, such as BIPM's or NPL's.

5 Knowledge transfer

It was mentioned in the previous SS*f*M report on standards [16] that a Special Issue of *Metrologia* [1, 2], commissioned by the editor of that journal, on Statistical and Probabilistic Methods for Metrology, edited by Walter Bich (INRIM) and Maurice Cox (NPL), was published electronically just before that report appeared. The hard copy appeared in August 2006.

Following the publication of this volume the number of downloads from the *Metrologia* website in the subsequent month increased from a little over five thousand (the typical monthly average) to a little over ten thousand. In a paper presented to the Comité International des Poids et Mesures (CIPM), the Editor of *Metrologia* stated

> "The dramatic increase in the number of downloads [of *Metrologia*] ... was caused by publication of the special issue on Statistical and Probabilistic Methods for Metrology. This dramatic increase in the number of those reading and downloading articles from *Metrologia* demonstrates the great interest in the subject matter of this special issue."

Several of the papers in this Special Issue relate to JCGM/WG1 activity [3, 4, 21, 42, 44, 45, 47]. Some of these papers were written by members of JCGM/WG1 and others were spurred to write them by the activity of JCGM/WG1.[6]

Twelve of the 120 full papers published in *Metrologia* in 2006 were cited by the Publisher as "journal highlights". Three of these papers [3, 21, 45] relate to the standards activity reported here and had SS*f*M co-authorship. Collaboration took place over the reporting period, principally through ISO/TC 69 and JCGM/WG1, with authors of *all* the other nine papers.

A paper co-authored with PTB was completed concerning the relationship of the first GUM Supplement to Bayesian analysis [22], and submitted to *Metrologia*.

6 Conclusions

Standards work in the third and final year of the third Software Support for Metrology programme, SS*f*M-3, concentrated on that for the British Standards Institution, the International Standards Organization, and the Joint Committee for Guides in Metrology.

[6] Although they do not strictly address "standards", in addition several papers in this volume [17, 43, 48, 49, 50] relate to the work of BIPM Director's Advisory Group on Uncertainty [15], the Chairman for which is provided by the SS*f*M programme.

SS*f*M continued to provide convenor and secretary for British Standards committee SS/6, Precision of Test Methods, and SS/6 Panel 3, Measurement Uncertainty, which for purposes of economy and efficiency now hold combined meetings. A major task, accomplished by SS/6 through a BSI/DTI contract, was the development of a detailed draft of a standard in the form of a Technical Specification for the determination and use of straight-line calibration functions that takes proper account of the types of uncertainty structure commonly encountered.

SS*f*M continued to provide a member on ISO/TC 69, Applications of Statistical Methods, and the convenor for ISO/TC 69/SC 6/WG 7, Statistical methods to support measurement uncertainty evaluation. The work on the above draft Technical Specification has already fed strongly into ISO/TC 69/SC 6/WG 7 activity. Other main work items of ISO/TC 69/SC 6/WG 7 concerned the planning of the progression to full standards of two Technical Specifications on measurement uncertainty evaluation it had produced previously.

A new Ad hoc Group of ISO/TC 69 concerned with statistical standards and software was formed. SS*f*M provides the convenor for this Group. A document that reviews existing statistical standards in terms of software that supports them was prepared by NPL. A document containing guidelines for producing statistical standards to be supported by software is being developed by NPL and LGC Ltd.

SS*f*M continued to provide a member to JCGM/WG1, Measurement Uncertainty. NPL leads the work on the development of the three planned GUM Supplements and the introductory document. The first Supplement, on the propagation of distributions, was finalized from the viewpoint of JCGM/WG1 and sent for voting and editorial comment only to all Member Organizations. On receipt of comments, assuming a positive vote, the Supplement will be completed and published.

Acknowledgments

The National Measurement System Directorate of the UKs Department of Trade and Industry supported the work of the National Physical Laboratory in this area.

Permission to reproduce figure 1 was given by ISO/TC 69. Correspondence concerning this chart and the work of ISO/TC 69 in general should be directed to M Laurent Parrot (laurent.parrot@afnor.org).

Peter Harris made valuable comments on a draft of this report.

References

[1] W. Bich and M. G. Cox. Foreword to Special Issue on Statistical and Probabilistic Methods for Metrology. *Metrologia*, 43, 2006.

[2] W. Bich and M. G. Cox, editors. *Special issue of Metrologia on Statistical and Probabilistic Methods for Metrology*, volume 43. 2006.

[3] W. Bich, M. G. Cox, and P. M. Harris. Evolution of the Guide to the expression of uncertainty in measurement. *Metrologia*, 43:S161–S166, 2006.

[4] Walter Bich, Luca Callegaro, and Francesca Pennecchi. Non-linear models and best estimates in the GUM. *Metrologia*, 43:S196–S199, 2006.

[5] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — An introduction to the "Guide to the expression of uncertainty in measurement" and related documents. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 104, in preparation.

[6] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — Applications of the least-squares method. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, in preparation.

[7] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — Concepts and basic principles. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 105, in preparation.

[8] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — Supplement to the "Guide to the expression of uncertainty in measurement" — Modelling. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 103, in preparation.

[9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — Supplement to the "Guide to the expression of uncertainty in measurement" — Models with any number of output quantities. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 102, in preparation.

[10] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — Supplement to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 101, in preparation.

[11] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — The role of measurement uncertainty in deciding conformance to specified requirements. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, in preparation.

[12] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. Guide to the expression of uncertainty in measurement. International Organisation for Standardisation, Geneva, Switzerland, 1995. ISBN 92-67-10188-9. Corrected and reprinted.

[13] BSI. BS 5233. Glossary of terms used in metrology, 1986. British Standards Institution, London.

[14] BSI. BS 5702. Guide to statistical process control for variables—Part 2: Charts for individual values, 2001. British Standards Institution, London.

[15] M. G. Cox. The evaluation of key comparison data. *Metrologia*, 39:589–595, 2002.

[16] M. G. Cox. Statistics standards to support metrology 2005–6. Technical Report DEM-ES(RES)007, National Physical Laboratory, Teddington, UK, 2006.

[17] M. G. Cox, C. Eiø, G. Mana, and F. Pennecchi. The generalized weighted mean of correlated quantities. *Metrologia*, 43:S268–275, 2006.

[18] M. G. Cox, A. B. Forbes, P. M. Harris, and I. M. Smith. The classification and solution of regression problems for calibration. Technical Report CMSC 24/03, National Physical Laboratory, Teddington, UK, 2003.

[19] M. G. Cox and P. M. Harris. Software specifications for uncertainty evaluation. Technical Report DEM-ES-010, National Physical Laboratory, Teddington, UK, 2006.

[20] M. G. Cox and P. M. Harris. SSfM Best Practice Guide No. 6, Uncertainty evaluation. Technical Report DEM-ES-011, National Physical Laboratory, Teddington, UK, 2006.

[21] M. G. Cox and B. R. L. Siebert. The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty. *Metrologia*, 43:S178–S188, 2006.

[22] C. Elster, W. Wöger, and M. G. Cox. Draft GUM Supplement 1 and Bayesian analysis. 2007. Submitted.

[23] L. J. Gleser. Assessing uncertainty in measurement. *Stat. Sci.*, 13:277–290, 1998.

[24] IEC. IEC 62387-1. Passive integrating dosimetry systems for environmental and personal monitoring for external photon and beta radiation using electronic devices for the data evaluation. Part 1: General characteristics and performance requirements on dosimetry systems. International Electrotechnical Commission, Geneva, Switzerland, 2006.

[25] ISO. ISO Guide 35 — Certification of reference materials — General and statistical principles, 1989. International Standards Organization, Geneva.

[26] ISO. ISO 5725-1. Accuracy (trueness and precision) of measurement methods and results—Part 1: General principles and definitions, 1994. International Standards Organization, Geneva.

[27] ISO. ISO Guide 98 — Guide to the expression of uncertainty in measurement (GUM), 1995. International Standards Organization, Geneva.

[28] ISO. ISO 11095. Linear calibration using reference materials, 1996. International Standards Organization, Geneva.

[29] ISO. ISO Guide 99 — International vocabulary of basic and general terms in metrology (VIM), 1996. International Standards Organization, Geneva.

[30] ISO. ISO 11843-1. Decision limit, detection limit, capability of detection—Part 1: Terms and definitions, 1997. International Standards Organization, Geneva.

[31] ISO. ISO 5725-3. Accuracy (trueness and precision) of measurement methods and results—Part 3: Intermediate measures of the precision of a standard measurement method, 2001. International Standards Organization, Geneva.

[32] ISO. ISO 5725-4. Accuracy (trueness and precision) of measurement methods and results—Part 4: Basic methods for the determination of the trueness of a standard measurement method, 2001. International Standards Organization, Geneva.

[33] ISO. ISO 5725-2. Accuracy (trueness and precision) of measurement methods and results—Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, 2002. International Standards Organization, Geneva.

[34] ISO. ISO 10576-1. Statistical methods — Guidelines for the evaluation of conformity with specified requirements — Part 1: General principles, 2003. International Standards Organization, Geneva.

[35] ISO. ISO TS 21748. Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, 2004. International Standards Organization, Geneva.

[36] ISO. ISO 5725-5. Accuracy (trueness and precision) of measurement methods and results—Part 5: Alternative methods for the determination of the precision of a standard measurement method, 2005. International Standards Organization, Geneva.

[37] ISO. ISO 5725-6. Accuracy (trueness and precision) of measurement methods and results—Part 5: Use in practice of accuracy values, 2005. International Standards Organization, Geneva.

[38] ISO. ISO TS 21749. Measurement uncertainty for metrological applications—Simple replication and nested experiments, 2005. International Standards Organization, Geneva.

[39] ISO. ISO 11843-5. Capability of detection—Part 5: Methodology in the non-linear calibration case, in preparation. International Standards Organization, Geneva.

[40] ISO/IEC. ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories, 2005. International Standards Organization, Geneva.

[41] R. Kacker and A. Jones. On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent. *Metrologia*, 40:235–248, 2003.

[42] Raghu Kacker, Blaza Toman, and Ding Huang. Comparison of ISO-GUM, draft GUM Supplement 1 and Bayesian statistics using simple linear calibration. *Metrologia*, 43:S167–S177, 2006.

[43] Ignacio Lira. Bayesian evaluation of comparison data. *Metrologia*, 43:S231–S234, 2006.

[44] Ignacio Lira and Wolfgang Wöger. Comparison between the conventional and Bayesian approaches to evaluate measurement data. *Metrologia*, 43:S249–S259, 2006.

[45] M. J. T. Milton, P. M. Harris, I. M. Smith, A. S. Brown, and B. A. Goody. Implementation of a generalized least-squares method for determining calibration curves from data with general uncertainty structures. *Metrologia*, 43:S291–S298, 2006.

[46] National Physical Laboratory. *Software Support for Metrology Programme*. http://www.npl.co.uk/ssfm/index.html.

[47] K. D. Sommer and B. R. L. Siebert. Systematic approach to the modelling of measurements for uncertainty evaluation. *Metrologia*, 43:S200–S210, 2006.

[48] A. G. Steele and R. J. Douglas. Extending chi-squared statistics for key comparisons in metrology. *J. Comp. Appl. Maths.*, 192:51–58, 2006.

[49] R. Willink. Meaning and models in key comparisons, with measures of operability and interoperability. *Metrologia*, 43:S220–S230, 2006.

[50] R. Willink. Principles of probability and statistics for metrology. *Metrologia*, 43:S211–S219, 2006.

A The determination and use of straight-line calibration functions

This annex provides the scope for the work on the determination and use of straight-line calibration functions. It also gives some justification for the work.

A.1 Project scope

This project is concerned with the development of a new standard in the form of a Technical Specification for the determination and use of straight-line calibration functions that describe the relationship between a stimulus variable and a response variable. Given are a number of paired measured values of the stimulus and response variables, as well as the uncertainty structure associated with this data. For a number of common uncertainty structures (including uncertainties/correlation associated with response variable measurement data only, and uncertainties/correlation associated with both response and stimulus data), the Technical Specification will describe:

- Suitable algorithms for estimating the parameters (gradient and intercept) of the calibration function and their associated uncertainties and covariance;
- How to validate whether the calibration function used is appropriate;
- The subsequent use of the calibration function parameter estimates and their associated uncertainties and covariance to provide predicted values of the stimulus variable and the associated uncertainties given measured response values and their associated uncertainties.

A matrix-based approach is used because of the complexity or even impossibility of providing explicit formulae to cover the more complicated uncertainty structures. Where appropriate, formulae or computational procedures would be provided. The Technical Specification would complement ISO 11095 [28], be consistent with the GUM [12] and algorithmically be based on previous SS*f*M work [18].

In some cases, the uncertainties associated with the measured stimulus values can reasonably be assumed to be negligible in terms of their influence on the parameters of the calibration function and the associated uncertainties, relative to those associated with the measured response values. In those cases a simple linear least-squares approach can be used to solve the calibration problem. When this is not the case, or if there is correlation associated with the measured values, an appropriate parameter estimation algorithm must be used.

A.2 Justification

Most standards and other documents concerned with regression and calibration consider problems where only the measured response values are subject to significant uncertainty and offer limited advice regarding the selection and use of parameter estimation algorithms. BSI/SS/6, ISO/TC 69/SC 6 and the SS*f*M programme have identified the need for guidance relating to calibration problems. It is felt that a single generic standard providing reliable algorithms for linear parameter estimation problems arising from

common measurement uncertainty structures (including uncertainties associated with measured values of both variables) will provide an extensive and easily accessible means to treat any linear calibration problem.

The standard would be of use to all calibration laboratories and testing laboratories that calibrate their equipment and relevant accreditation bodies, and would have the significant effect of ensuring self-consistency in results and confidence in the use of calculated uncertainties. It would relate to a wide range of applications in science, not only traditional areas but also rapidly growing areas of measurement such as nanotechnology and biotechnology.

Faithful implementation of the proposed standard would ensure that different organisations that operated with the same calibration data would provide identical results as a consequence of a common underlying solution technology. Solutions could be trusted and not have to be reworked because of possible concerns with the quality of the computational procedures.

B Scope of the activity of ISO/TC 69 Ad hoc Group on Statistical Standards and Software

The Ad hoc Group, Statistical Standards and Software, of ISO/TC 69, Applications of Statistical Methods, will produce guidance for the developers of statistical standards that are to be supported by software. Considerations will include

1. a review of existing standards (ISO and national) for which there exists software support,
2. the manner in which standards should convey information to permit software implementation,
3. forms the software can take,
4. future revisions of such standards; evolution/support of software, and
5. issues regarding legal aspects: copyright, liability, etc.