Report to the National Measurement System Policy Unit, Department of Trade and Industry

From the Software Support for Metrology Programme

International activities during the second SS*f*M programme, 2001–2004

By
M G Cox and M J Stevens

February 2004

International activities during the second SS*f*M programme, 2001–2004

M G Cox and M J Stevens
Centre for Mathematics and Scientific Computing

February 2004

ABSTRACT

This document reports on international activities relating to the UK's second Software Support for Metrology (SS*f*M) programme, 2001–2004. In the first SS*f*M programme (1998–2001) a status report on software activity worldwide was produced, structured according to the themes and projects of that programme. This document has a change of focus towards particular international SS*f*M activities of importance to metrology in general, whilst still covering the whole scope of the SS*f*M programme.

ISSN 1471–0005

Authorised by Dr Dave Rayner,
Head of the Centre for Mathematics and Scientific Computing

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

Contents

Background

The UK's Software Support for Metrology (SSfM) programmes, commissioned by the UK National Measurement System Directorate (NMSD), have the aim of improving mathematical and software support for metrology. In addition to the detailed technical work undertaken, support is given by organizing conferences and workshops, providing seminars and advice, undertaking case studies, publishing reports and best practice guides, and operating a club at which issues relating to mathematics and software, across all metrology areas, are discussed. A further feature of the programme is collaboration with relevant organizations worldwide to give added value to the work through complementing its resources and impact.

This report covers international aspects of SSfM-2, the second SSfM programme, 2001–2004, summarizing the status of SSfM-related activity worldwide. In the first programme, SSfM-1, a status report on software activity worldwide was produced [18], structured according to the themes and projects of that programme. The current document has a change of focus towards particular international SSfM activities of importance to metrology in general, whilst still covering the whole scope of the SSfM programme.

Three areas were especially prominent internationally: (a) uncertainty evaluation, (b) key comparison data evaluation and (c) standards relating to statistical calculations for measurement data. These and the other areas indicated are underpinned technically by mathematical and statistical modelling and numerical algorithms.

A reason for the especial importance internationally of (a) is that through their quality systems many laboratories, both the national measurement institutes (NMIs) and industrial laboratories, are required to work to documents such as the GUM (*Guide to the Expression of Uncertainty in Measurement*) [4], published by ISO, and national, regional and sectorial interpretations of the GUM, and ISO/IEC 17025 (*General requirements for the competence of testing and calibration laboratories*) [49]. Many technical issues are raised when implementing the provisions of these documents in the various metrology areas.

Regarding (b), the publication of the MRA (*Mutual Recognition Arrangement*) [3] has required the NMIs to take an approach to key comparisons and the evaluation of the measurement data they provide that is as rigorous as economically possible. Accordingly, much consideration of this topic is necessitated. The *Uncertainties and statistical modelling* project in the SSfM programme has concentrated on (a) and (b) and related topics, and has liaised internationally with many organizations.

Regarding (c), many standards are used in metrology that require numerical and statistical procedures for their implementation.

There are several other technical areas within the compass of SS*f*M such as discrete and continuous modelling, software testing and validation, and internet-enabled metrology. In addition there are metrology clubs and networks and European framework programmes. All these are considered to an extent that is proportionate to the effort involved internationally.

The sections of this report are numbered in accordance with the Theme, Project and Task numbering in SS*f*M-2, to facilitate comparison with the goals of that programme.[1]

1 Modelling techniques

1.1 Extension of empirical models

1.1.1 Spline algorithms and software with automatic knot placement

There is wide interest internationally [9, 10, 11, 12, 13] in the use of discrete empirical models for the analysis of metrology data in the many areas in metrology where adequate physical models are unavailable. Of particular value is data modelling by univariate splines, because of the flexibility of such models to adapt to widely differing types of behaviour. Work within SS*f*M-2 with sub-contractor University of Huddersfield led to improved capabilities in this regard through the provision of spline knot-placement strategies and the evaluation of uncertainties associated with the spline. The work in this area [33] was presented as an invited talk at the international conference *Algorithms for Approximation IV* held in Huddersfield in July 2001 (section 6.5.2). Following an invitation by IAC-CNR, Rome, a further presentation of the work on knot placement in data modelling by univariate splines [34] was given in Rome in December 2001.

A software implementation containing a number of knot-placement algorithms for the approximation of data by splines was implemented, and is available through EuroMetroS (section 6.1).[2] The software package acts as a 'pre-processor' to NPLFit, software developed by NPL for modelling experimental data using polynomial and spline curves, and also

[1]The programme document for SS*f*M-2 is available at www.npl.co.uk/ssfm/download/documents/ssfm2-prog.pdf .

[2]The EuroMetroS website is www.eurometros.org/ .

available through EuroMetroS. The software takes as input a file of measurement data and allows the user to explore a number of knot-placement strategies to help derive a satisfactory approximation to the data. The output from the software is a set of knot positions that may be used subsequently as input to NPLFit. Software for extracting features (zeros, maxima, minima, inflexions, etc.) from a spline model, was also made available through EuroMetroS, and applied to obtain, e.g., thermo-physical properties from measurement data.

1.2 Algorithms for discrete modelling

1.2.1 Algorithms and software for generalized regression

A report on the classification and solution of regression problems for calibration [29] provides regression approaches for all types of uncertainty structure that can be expected to arise. The report is being used by the Joint Committee for Guides in Metrology in the development of a supplemental guide to the GUM concerned with least-squares adjustment in calibration and other areas (section 1.5.3).

Previously, related work on generalized Gauss-Markov approaches to regression [45] was presented at the international conference *Algorithms for Approximation IV* held in Huddersfield in July 2001 (section 6.5.2).

There is demand for explicit guidance in the form of an international standard on the use of regression for problems with various types of uncertainty structure. The SS*f*M programme has prepared a 'new work item proposal' for a standard in the area that concentrates on the important linear calibration model (section 4.1). This proposal was submitted to the British Standards Institution (BSI), and will subsequently be submitted to the International Standards Organization (ISO) committee TC 69, Application of statistical methods.

An expert on software for uncertainty evaluation and generalized regression, Dr Erwin Achermann, from ETH, Zurich, presented a seminar *With Certainty to Uncertainty* on these topics at NPL in January 2002 (also see section 1.5.3). It was attended by representatives of several NMS programmes.

1.2.2 Non-Gaussian measurement errors

A focus of the work on non-Gaussian measurement errors was a variation of the least-squares approach having robustness properties when applied to metrology data containing discrepant measurements. The new approach can be implemented by a straightforward adaptation of a non-linear least-squares algorithm. This work was presented under the title 'A simple algorithm for robust orthogonal regression' at the *International Congress for Industrial and Applied Mathematics* in Sydney in July 2003 as part of a themed session on regression. (The Wilkinson Prize for Numerical Software was awarded at the same meeting (section 6.5.4).) This work was carried out collaboratively with SS*f*M sub-contractor Huddersfield University.

Prof John Mason, University of Huddersfield, presented another aspect of the research in a paper [52] under the title 'Parametrized approximation estimators for mixed noise distributions' at the international conference *Advanced Mathematical and Computational Tools in Metrology* in Turin in September 2003 (section 6.5.3).

The problem of testing software that addressed approximation problems appropriate for non-Gaussian noise was also considered. The paper 'Validation of CMM form and tolerance assessment software' [21] presented at the international conference *Laser Metrology and Machine Performance V* in Birmingham in July 2001 considered software for fitting geometric elements to data according to the Chebyshev criterion.

1.2.4 Techniques for the efficient solution of large structured calibration problems

A number of large-scale calibration problems in metrology have design matrices containing a distribution of zero and non-zero elements following particular commonly-occurring patterns. Economical solution of such problems requires specially designed numerical algorithms of linear algebra. Presentation of a paper, 'Efficient algorithms for structured self-calibration problems' [43], on the work in this area was made at the international conference *Algorithms for Approximation IV* held in Huddersfield in July 2001 (section 6.5.2).

1.3 Promotion of discrete modelling best practice

1.3.1 Maintenance of discrete modelling best practice guide

The first SS*f*M programme produced a best practice guide on discrete modelling [27] and a companion best practice guide on discrete model validation [28]. These guides are being merged and extended within the current SS*f*M programme. They were also contributed for review and comment to the MetroNet cooperative project (COP4) on discrete modelling, data fusion and data fitting (section 6.5.3).

1.5 Uncertainties and statistical modelling

The aspects of uncertainty and statistical modelling on which there is greatest concentration internationally in terms of research and the provision and maintenance of standards and guidance are those indicated in the Background section of this report, viz.,

- Measurement uncertainty evaluation
- Evaluation of key comparison data
- Standards relating to statistical calculations for measurement data.

In recognition of this fact, the second SS*f*M programme gave increased emphasis to these topics, in the project *Uncertainties and statistical modelling*. The activity in the related project *Mathematical and statistical input to metrology standards* was also increased (section 4.1). The SS*f*M programme has representation on the main committees and working groups concerned with generic standards and guides (also see section 4.1).

1.5.1 Statistical modelling

Problems involving correlated data in which *a priori* quantification of the extent of the correlations is required are common in metrology. Such problems have often historically been solved under the assumption that measurement data is uncorrelated. The assumption is a useful one because it leads to problem formulations and solution processes that are typically straightforward to establish and apply. If, however, the correlation effects are ignored, i.e., the data values are taken as mutually independent, the uncertainties evaluated on that basis can strongly underestimate (in by far the most common case of *positive* correlation) those obtained by a more realistic evaluation. It is likely that in some metrology disciplines optimistically small values have indeed been produced.

Within SS*f*M-2 this topic was addressed, partly through international collaboration, in two ways: the use of autoregressive schemes to model correlated data and the use of bootstrap methods to account for correlation effects (section 1.5.2).

Autoregressive schemes, in which the measurement data is analyzed to expose serial correlation effects and hence quantify the associated covariance matrix provide the possibility for realistic evaluation, but there is much to learn in the context of metrology. Activity has involved collaboration with Dr Giovanni Mana (IMGC). Following a meeting held at IMGC, as part of a visit by a member of the SS*f*M team to several organizations in Italy, Dr Mana was subsequently invited to present a paper on the subject at the international uncertainty conference UNCERT 2003 [53] (section 6.5.1). NPL has so far applied the approach to dimensional and chemical measurement.

Collaborative work with NMi on the use of error analysis in modelling a measurement as a precursor to uncertainty evaluation was undertaken, a paper on this work being published in *Metrologia* [69]. The intention is to demonstrate that error analysis has a place in metrology, despite its deprecation in the GUM [4]. The paper shows that an initial formulation using error analysis can assist in providing a foundation for subsequent uncertainty evaluation according to the GUM. It also shows that correlation effects can also sometimes be handled more conveniently.

Some of the statistical modelling principles were applied to the area of time and frequency measurement. Work carried out with experts in the Time and Frequency Metrology programme included the characterization of regular but unevenly spaced two-way satellite time and frequency transfer (TWSTFT) measurements using second-difference statistics [39], and the least-squares analysis of such measurements [47]. The work was presented at the *Fourth Time Scales Algorithms Symposium* held at BIPM in March 2002.

1.5.2 Non-parametric methods of statistical analysis

The nature of the statistical distributions underlying metrology data is often unknown. Accordingly, the so-called non-parametric methods that are free from or resilient to distributional assumptions are applicable. Some such methods were investigated and applied. The technique known as bootstrap re-sampling is one such method and involves repeated re-sampling, with replacement, from actual data to generate “new” data sets that mimic those that would be obtained by re-measurement. In collaboration with IMGC and CNR-IAC (Rome) the approach was applied to key comparison data evaluation (section 1.5.4). In collaboration with the University of Bologna, the approach was applied to roundness evaluation in dimensional metrology and compared with the use of autoregressive schemes (section 1.5.1). The latter work benefited from a researcher, Dr Annarita Lazzari, from the University of Bologna, working alongside the SS*f*M team at NPL for a period of six months in 2002. A poster on the approach was presented at the BIPM-NPL conference *Impact of Information Technology in Metrology* (section 6.5.1) and an invited seminar in Turin to IMGC and IEN as part of the mentioned visit to Italy (section 1.5.1). The seminar covered both the use of bootstrap re-sampling and autoregressive modelling (section 1.5.1). This application area provides case studies in two of the reports delivered as part of SS*f*M-2.

Collaboration with the Universities of Florence and Bologna has led to a new result concerning the determination of distribution-free confidence regions for a multivariate measurand, which generalizes Chebyshev's inequality for the univariate case [56]. Its extension to correlated variables was also treated. The work [37] was presented at the international conference *UNCERT 2003* [53] (section 6.5.1). This work demonstrates that tighter confidence regions can be obtained by taking account of known correlation effects, supporting the general scientific view that the use of further relevant information is conducive to providing improved results.

1.5.3 Support for GUM and other standards

A major area of international work with a significant impact in the world of metrology is the maintenance of the GUM. This work is the responsibility of Working Group 1 (WG1) of the Joint Committee for Guides in Metrology (JCGM). The SS*f*M programme makes appreciable input to the work of the JCGM and, as a consequence, was asked to present several invited papers on the subject at international meetings (section 6.5.1). The maintenance process currently consists of developing a series of supplemental guides to the GUM, to give added value to that document. Six supplemental guides are at various stages of development, to all of which the SS*f*M programme is contributing.

Supplemental Guide 1 (for which NPL is the lead institute) covers the concept of the propagation of distributions, a generalization of the law of propagation of uncertainty and the use of the Central Limit Theorem. It also covers the validation of methods traditionally used for calculating uncertainty. A large part of the work of JCGM/WG1 has been devoted to this Supplemental Guide. At the time of publication of this report, this document is being made available to the member organizations of the JCGM and the NMIs for pre-publication review.[3]

[3]The member organizations of the JCGM are Bureau International des Poids et Mesures (BIPM), the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), the International Laboratory Accreditation Cooperation (ILAC), the International Standardization Organization (ISO), the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP) and

Supplemental Guide 2 (led by IMGC) treats models where there is more than one output quantity. Such (vector-valued) quantities are almost invariably mutually dependent. The formulae in the GUM for treating a single output quantity are generalized to this vector case. The formulae provided are expressed in matrix-vector notation both for ease of presentation and as a basis for their implementation in software.

Supplemental Guide 3 (led by NIST) treats conformity assessment, an area of importance to manufacturing industry, health and safety. In the presence of inevitable measurement uncertainty, decisions are required concerning quantities that are stipulated deterministically, but measured in a manner that means the actual knowledge of them is probabilistic. This guide provides generic principles in the area that can be tailored to a range of disciplines (also see section 4.2.2).

Supplemental Guide 4 (led by NPL) treats the modelling or formulation stage of uncertainty evaluation. Although there are typically features that relate to a particular application, general guidance is provided on several aspects including (a) the nature of the model (physical, empirical or hybrid), (b) the distributions assigned to the values of the input quantities, and (c) the nature of the uncertainty calculation, viz., whether the quantities involved are real or complex, the model is explicit or implicit, and the output quantity is scalar- or vector-valued. The document is aimed at the NMIs and comparable audiences.

Supplemental Guide 5 (led by NEL) is also concerned with modelling, but with an emphasis at a different level, viz., for industrial use. It is planned to contain many examples of simple models to illustrate conventional GUM-based principles of uncertainty evaluation.

Supplemental Guide 6 (led by PTB) treats the use of least squares and the associated uncertainty evaluation within metrology. It will incorporate some of the concepts in the form of reliable numerical methods, considered in the discrete modelling project of the SS*f*M programme. In particular, some of the regression work (section 1.2.1) will be used.

In addition, a short introduction to measurement uncertainty, the GUM and the supplemental guides is being prepared.

As well as being a member of the project teams concerned with producing all six guides, NPL has lead responsibility for the introductory document. A great advantage of NPL's involvement in developing these documents is that the support provided by the SS*f*M programme enables NPL to devote the effort to the work that it requires and deserves. Some of the other

the International Organization of Legal Metrology (OIML).

organizations are in a less favourable position in that they are not funded by a comparable generic programme. However, in this and other contexts, several countries have expressed interest in developing national programmes that would parallel the SSfM activity (cf. section 6.5.1).

The SSfM best practice guide on uncertainty and statistical modelling [23], the uncertainties software specification [24], and the SSfM best practice guide on discrete modelling [27], all updated within SSfM-2, are used as a basis for the NPL contributions to the supplemental guides.

An extensive document constituting an alternative to the GUM was submitted to the JCGM by an ex-Director of NIM, the Chinese NMI. It describes an approach that is based on different concepts from those of the GUM. The document was reviewed and the Chinese contingent asked to consider ways to improve alignment with the GUM and to highlight advantages of their approach. Several submissions were received and reviewed subsequently, but conceptual differences remain.

Particular emphasis was placed in SSfM-2 on developing further the concept of the propagation of distributions, a generalization of the law of propagation of uncertainty and the application of the Central Limit Theorem, and the basis of Supplement 1 to the GUM (above). Once (an implementation of) the approach has been used to provide (an approximation to) the probability distribution for the value of the output quantity, in the form of a distribution function, it is straightforward to use it to determine (a) an estimate of the value of the output quantity, (b) the standard uncertainty associated with this estimate and (c) a coverage interval for the value of the output quantity. Part of this work was carried out in collaboration with uncertainty experts at PTB. Several papers [22, 30, 31] were published to promote the approach, some jointly with the PTB collaborators.

A further article on the propagation of distributions constituted the published version (in Russian) of an invited paper [35] on the subject presented at the VNIIM-NPL workshop *Mathematics, Statistics and Computation to Support Measurement Quality* in St Petersburg in June 2002 (section 6.5.1). It will subsequently be published also in English in an American journal.

An invited seminar on the general principles of uncertainty evaluation was given in October 2002 at the Universities of Florence and Bologna. A meeting was held on the subject with these universities, and subsequent work carried out, which focused on the use of the principle of maximum entropy to assign appropriate probability distributions to the values of the input quantities to a measurement model. This principle is also staunchly endorsed by PTB [72] and forms an integral part of the formulation stage of the propagation of distributions, the basic concept of the first supplemental guide to the GUM (above).

In conjunction with the above visit to Italy, a related discussion was also held with IMGC, especially on the handling of the correlation effects that arise in temperature measurement.

PTB organized a seminar *The use of Monte Carlo simulation in measurement uncertainty* in Braunschweig in June 2001 (with some 50 attendees). The SS*f*M programme was represented at the seminar, providing an invited presentation, 'The planned supplemental guide to the GUM: Numerical Methods for Propagating Distributions'.

A presentation on measurement uncertainty and the propagation of distributions was made at *Metrologie 2001*, an international metrology congress held in Saint-Louis, France in October 2001.

The collaborative work with PTB on the use of Monte Carlo simulation for uncertainty evaluation was also presented at *NMC 2001*, the UK National Measurement Conference, in Harrogate in November 2001. Following this conference, Dr Bernd Siebert of PTB worked for a period at NPL as a guest worker on Monte Carlo simulation for uncertainty evaluation.

The use of Monte Carlo simulation for uncertainty evaluation was also presented as an invited paper at the INITIATION Workshop *Solving Practical Problems when Implementing a Quality System based upon ISO/IEC 17025 in NMIs*, held in Rotterdam in December 2001 (also see section 4.1).

For fifteen years there has been close collaboration with IMGC (Turin) and IAC (Rome) of the Italian National Research Council on mathematical, statistical and numerical aspects of metrology, especially in terms of organizing the series of conferences on advanced mathematical and computational tools for metrology (section 6.5.3). This collaboration has continued under the SS*f*M programmes, and in December 2001 Maurice Cox from the SS*f*M team was invited to work for a period at IAC on various topics including uncertainty evaluation, key comparison data evaluation and empirical modelling, and make presentations of the SS*f*M work (sections 1.1.1 and this section).

Dr Erwin Achermann, ETH, Zurich gave a seminar at NPL in January 2002 on MUSAC (software for GUM, generalized least squares, etc.), attended by representatives of several NMS programmes (also see section 1.2.1).

An invited paper [31] on the GUM and its planned supplemental guides was presented at the *International ILAC/IAF Conference on Accreditation in Global Trade* in Berlin in September 2002. This meeting provided an opportunity to raise awareness in the accreditation community of the latest guidance on uncertainty evaluation being developed by the JCGM.

Jointly with PTB and IEN, work was carried out on the application of model-based measurement uncertainty evaluation to testing. A main reason for the work was to promote the use of models in conjunction with sound uncertainty evaluation principles in the testing area, which is not as advanced as that in calibration. An invited talk [25] on the subject was presented at the MetroTrade workshop *Traceability and measurement uncertainty in testing* in Berlin in January 2003. It covered basic concepts and implementations of the model-based approach to uncertainty evaluation, with two testing applications, tennis-ball rebound and tensile strength, to illustrate the principles.

Following the ILAC/IAF Conference, NPL was invited to provide the guest speaker at the annual technical seminar on Laboratory Accreditation organized by IAJapan in Tokyo in March 2003. The presentation, attended by 300 members of the Japanese accredited laboratories and the four Japanese accreditation bodies, was also on the GUM and its supplemental guides.

A paper [5] by the British Measurement and Testing Association (BMTA), an SS*f*M collaborator, on standards and software for evaluating uncertainty was presented at the international *UNCERT 2003* conference [53] (section 6.5.1).

Collaboration took place with the University of Bologna and the Length Metrology programme on uncertainty evaluation associated with vision-system co-ordinate measuring machines (CMMs), during and following Annarita Lazzari's six-month period at NPL in 2002 (section 1.5.2). Miss Lazzari, now with Mitutoyo Italiana in Milan, has since been awarded a PhD in the subject.

1.5.4 Interlaboratory comparison statistical analysis

Collaboration on key comparison data evaluation between the UK, through the SS*f*M programme, and Canada, Denmark, France, Germany, Italy, Russia, Sweden, The Netherlands and the USA took place. The nature of these collaborations is indicated below.

Joint work was carried out with PTB on the use of Monte Carlo simulation to evaluate the uncertainties associated with key comparison reference values (KCRVs) and degrees of equivalence (DoEs). It was presented as a paper at the joint VNIIM-NPL conference *Mathematics, Statistics and Computation to Support Measurement Quality* held in St Petersburg in June 2002 (section 6.5.1) and two papers at the PTB International Workshop *Data analysis of interlaboratory comparisons* held in Berlin in December 2002 (below).

An invited presentation was made to BIPM on 'Considerations in the calculation of key comparison reference values'.

BIPM Director's Advisory Group on Uncertainty A member of the SS*f*M project team was appointed to lead the international working group, *BIPM Director's Advisory Group on Uncertainty*, with special responsibility for key comparison aspects, formed in May 2001. Its other members are from NIST, PTB, IMGC and DFM. The group reports to the Director of BIPM, who in turn reports its findings to the Comiteé International des Poids et Mesures (CIPM). The remit of the group is to address questions related to uncertainties put before it by the Director of the BIPM and to give whatever advice is requested. The main concentration is on matters that arise in the practical application of the procedures laid down in the MRA [3] and on the preparation of guidelines for key comparisons.

The first duty of the group was to prepare a paper giving guidelines for key comparisons that are based on the circulation of a single stable travelling standard among the participants in cases where each institute's measurement is realized independently. The guidelines include two procedures, Procedure A and Procedure B. Procedure A is applicable to key comparisons satisfying two conditions. The first condition is that for each participating institute a Gaussian distribution (with mean equal to the institute's measurement and standard deviation equal to the provided associated standard uncertainty) can reasonably be assigned to the quantity of which the institute's measurement is an estimate. The second condition is that the KCRV determined by the procedure is statistically consistent with the institutes' measurements in terms of their quoted uncertainties. Procedure B is more appropriate when one or both of these conditions fails to hold.

This work was completed with the publication of a paper [19], in *Metrologia*, together with an introductory article [20] explaining the background to the activity.

The guidelines benefited from a review of an advanced draft by all NMS Programme Managers or their representatives, all BIPM section heads, members of CIPM Consultative Committee Working Groups, and other experts at the NMIs. A summary of the guidelines was provided to the Director of BIPM for presentation to NMI Directors.

To increase awareness, presentations were given at a satellite workshop *Key Comparison Data Evaluation* to the BIPM-NPL international workshop *The Impact of Information Technology in Metrology* (below and section 6.5.1) and, as an invited paper, at the VNIIM-NPL International Seminar *Mathematics, Statistics and Computation to Support Measurement Quality* (section 6.5.1). It was also presented, by invitation, at the PTB International Workshop *The Analysis of Interlaboratory Comparison Data* (below) and, with the permission of the Editor and publishers of *Metrologia*, provided for the proceedings of that meeting [60].

The Advisory Group is now developing further guidelines that treat instability in the travelling standard, associated with deterministic and random drift effects, or with an abrupt change, or both.

An invited presentation was made in May 2001 at the Swedish National Measurement Conference on ‘The evaluation of interlaboratory comparison data’. It covered a range of candidate approaches for such evaluations.

Work was carried out collaboratively with VNIIM on a model-based approach to key comparison data evaluation. The approach has particular application in cases where the institutes’ measurements are correlated, such as when traceability is obtained from a common source. It led to a paper [8], ‘A model-based approach to key comparison data evaluation’, which was presented at the 17th IMEKO World Congress, *Metrology in the 3rd Millennium*, in Dubrovnik in June 2003.

Much discussion between BIPM, the European NMIs and NIST took place concerning the approaches proposed by BIPM Director’s Advisory Group on Uncertainty. NIST has a significant programme of work on key comparison data evaluation, with the objective of providing solutions based on an appropriate statistical methodology. The methodology depends on the assumptions made concerning the data and other aspects of the comparison. In the meantime the attitude taken by the Advisory Group is that procedures are needed now to support current and imminent comparisons, even if they do not reflect all aspects being considered in the NIST programme. The head of NIST’s Statistical Engineering Division presented an invited paper on NIST’s attitudes to key comparison data evaluation at the satellite workshop to the BIPM-NPL workshop *Impact of Information Technology in Metrology* (section 6.5.1).

It was considered by BIPM and NPL that it would be desirable to have application software available on the BIPM website, the use of which would assist Working Groups of the Consultative Committees of the CIPM to provide the technical information (key comparison reference value and the associated uncertainty and degrees of equivalence) required for a key comparison and the BIPM key comparison database. To underpin such software it would however be necessary first to have appropriate guidelines and procedures in place, the implementation of which would constitute the software. Any such software would hence be prepared following the availability of a sufficiently complete set of guideline documents.

Second international workshop at NPL on the Evaluation of Interlaboratory Comparison Data As a satellite workshop to the Information Technology workshop reported above, NPL hosted an international workshop *Key Comparison Data Evaluation*, as a follow-up to the *Statistical Analysis for Interlaboratory Comparisons* event held at NPL in November 1999, one month after the signing of the Mutual Recognition Arrangement. The major NMIs made presentations on approaches to the evaluation of comparison data that ranged from classical and robust estimators of reference values, to the linking of comparisons to detailed considerations of the factors that influence participants' results. The determination of the consequent degrees of equivalence was also much discussed. As anticipated, diverse views were expressed that were influenced by scientific and political considerations. Also see above and section 6.5.1.

PTB workshop on Data Analysis of Interlaboratory Comparisons PTB held an international workshop *Data Analysis of Interlaboratory Comparisons* in Berlin in December 2002 [60]. NPL was invited to author or co-author three of the papers for this meeting. The first paper, 'The evaluation of key comparisons', provides the technical underpinning to the key comparison data evaluation guidelines [19, 20] developed by BIPM Director's Advisory Group on Uncertainty (above). The second paper, jointly with PTB, 'Evaluation of key comparisons based on the probability distributions for the participating institutes' measurement quantities' [65] described the use of Monte Carlo simulation as a basis for implementing the propagation of distributions in the context of key comparisons. The approach described is consistent with that given as Procedure B in the key comparison data evaluation guidelines [19]. The third paper, also with PTB, 'The use of Monte Carlo simulation to evaluate uncertainties of key comparison reference values and degrees of equivalence' [36], built on the second paper, illustrating the concepts with the use of the median as the KCRV (also see the paragraph under this title below). NPL also provided the final presentation in the form of a summary of the meeting.

In conjunction with the second international workshop *Evaluation of Interlaboratory Comparison Data*, held in September 2002 at NPL (reported previously), the meeting covered the state of the art in the area, and presented the application of methods to various key comparisons by a number of CIPM Consultative Committee working groups. It was further emphasized that there is much work yet to be done in this critical area to provide procedures having a sufficiently high level of international acceptance.

Covariance quantification NRC had published work on 'exclusive statistics' [66], showing how the quantification of the covariance between an institute's measurement and the KCRV can be avoided when determining the DoE for the institute's measurement standard. The approach is based on omitting the measurement of the institute concerned in forming a KCRV 'for' that institute. Collaboration between the SS*f*M and the VAM programme led to a critical response to the work and a short paper indicating perceived shortcomings being published in Metrologia [55]. One main point made was that the use of the 'exclusive' forms for the DoEs to check whether there is significant influence of the correlation between the KCRV and the individual results simply corresponds to determining whether the weights in the formula for the KCRV are significant. A second point was that exclusive statistics should be used with care when calculating degrees of equivalence between measurement standards, since the results do not have the property of being independent of the KCRV. A final point was that Procedure B in the guidelines [19] prepared by BIPM Director's Advisory Group on Uncertainty permits the uncertainty associated with a KCRV and the degrees of equivalence to be calculated using any formula or algorithm.

The use of order statistics as a tool for key comparison data evaluation was studied with Francesca Pennecchi from IMGC, who was appointed as a Glazebrook Fellow at NPL for three months in 2003, and also supported by IMGC. The Fellowship placement gave a significant injection to the study of the theory and properties of such statistics (which includes the median as a special case), providing added value to both NPL and IMGC programmes. A paper [58] on the approach, which concentrates on quantifying the covariance between an institute's measurement and the KCRV, was presented at NPL to representatives of several NMS programmes and at the UNCERT 2003 conference [53] (section 6.5.1). The results obtained from applying the approach to measurements from a mass key comparison were compared with those from an application of Procedure B in the guidelines [19] produced by BIPM Director's Advisory Group on Uncertainty. The comparison was shown to be very favourable, a most reassuring result considering that the two approaches are very different in character. The approach offers an alternative to that above considered by NRC [66].

The median as a key comparison reference value Work was undertaken with PTB on the use of the median as a KCRV. Despite its attractions as a robust estimator, it was shown to exhibit some anomalous behaviour when used within a Monte Carlo calculation to approximate the distributions associated with DoEs. The reasons for this behaviour were studied and are now better understood. A paper [36] on this topic was presented at the PTB International Workshop *Data Analysis of Interlaboratory Comparisons* in Berlin (above).

Mixture distributions Work was carried out with IMGC and CNR-IAC (Rome) on the use of mixture distributions to characterize the dispersion of results from key comparisons. The approach developed applies to comparisons in which each institute acquires a realization of the same physical state, such as in some temperature comparisons. A paper [14] on the subject was accepted for publication in *Metrologia*.

Linking key comparisons Work was undertaken with BIPM and NMi on an approach for linking key comparisons, carried out by NMIs and Regional Metrology Organizations. The application to the gas analysis area, in the field of amount of substance measurements, was considered. A poster on the work was presented at the BIPM-NPL Workshop *Impact of Information Technology in Metrology* (section 6.5.1).

Applications to specific areas Solicited advice was given on key comparison data evaluation to a number of the working groups of the Consultative Committees of the CIPM, including CCAUV (hydropohone comparisons), CCEM (general), CCL (gauge blocks), CCM (mass), CCQM (gas standards), CCPR (lamps), CCRI (SIR: see section 1.5.5).

In particular, an invited presentation on methods for determining KCRVs and key comparison reference curves (section 1.5.5) was made at the CCQM gas analysis key comparisons meeting in April 2001. Moreover, in October 2001 a presentation on measurement uncertainty was made in Budapest at a EUROMET Length workshop, held in association with a meeting of the Consultative Committee for Length (CCL). Valuable feedback was obtained concerning users' perception of traditional GUM and Monte Carlo procedures for uncertainty evaluation.

Collaboration with NMi led to the completion of a paper on key comparisons for gas standards [68].

1.5.5 Key comparison reference curves and other new statistical techniques

Key comparison reference curves Discussions were held with PTB on their work on key comparison reference curves (KCRCs) in the context of accelerometry. In that area the KCRC has the special property that the underlying physical model is known.

The first phase of a joint project with BIPM and the Ionizing Radiation programme to support the Mutual Recognition Arrangement for radioactivity measurements was completed. The aim of the work was to provide a KCRC by using a combination of all physical knowledge available about the measurement in conjunction with a carefully selected empirical model and, at the same time, taking appropriate account of all provided data and the associated uncertainties. The KCRC applies to the efficiency values for radionuclides across the energy spectrum. This KCRC can be used to predict efficiency values for any radionuclide, whether previously measured or not. The first phase has demonstrated that the approach operates correctly in practice. The second phase, with the same partners, will provide a production version of the supporting software and a rigorous uncertainty evaluation that will be expected to deliver smaller uncertainties than would otherwise be obtained. The software would support the primary reference at BIPM.

A number of discussions were held with other NMS programmes on KCRCs in their areas.

New statistical techniques Work with NMi was carried out on the application of uncertainty evaluation principles to purity verification and presented at the Eurachem meeting in Lucerne in December 2002. A paper on this topic [67] was prepared and submitted to *Metrologia*. Like certain other areas in metrology, there are special conditions pertaining to the output quantities of a model that must be taken into account in estimating their values and evaluating the associated uncertainties. Problems arise due to mathematical constraints (e.g., the sum of all mass fractions cannot exceed unity) and to physical or chemical limitations (e.g., limits of detection, limits of quantification). This work discusses approaches to the evaluation of quantitative impurity data, including the evaluation of measurement uncertainty in a manner that respects the constraints. Two principal approaches are considered, one based on the maximum likelihood principle, and a second using Bayes' theorem. A comparison is made with the classical approach, using the limit of detection in conjunction with the rectangular distribution. The most important aspect of the work is that the coverage intervals formed for the values of the output quantities are feasible, e.g., the probability of negative values for the mass fractions is zero, thus facilitating the interpretation of the results obtained.

There are advantages in some metrology disciplines of considering a set of calibration curves rather than calibration curves in isolation (just as it is sometimes appropriate in key comparison data evaluation to consider a KCRC rather than a set of KCRVs). In gas component analysis, for instance, there is a calibration curve associated with each component. During a measurement process used to provide data for such an analysis, changing environmental conditions affect the indicated instrument responses. Calibration data variation is then attributed not only to random measurement effects, but also to that additional systematic effect. When the non-random effect dominates the data, a correlated response behaviour between the points on which the calibration curves depend is observed for all gas components participating in the experimental interaction. By using a model-based least-squares approach that compensates for these systematic effects, the calibration curves can be fitted 'simultaneously'. The resulting improvements in the uncertainties associated with the gas component concentrations can be dramatic, e.g., one tenth of the values that would otherwise be achieved. This work was presented at the AMCTM conference [38] (section 6.5.3).

1.5.6 Uncertainty best practice guide, software specification and training course

The first edition [23] of the best practice guide on uncertainties and statistical modelling was published in March 2001, having been developed during SS*f*M-1, which covered the period April 1998–March 2001. During that period Working Group 1 of the Joint Committee for Guides in Metrology started work, following its first meeting in March 2000, on Supplemental Guide Number 1 to the GUM concerned with numerical methods for the propagation of distributions. Material from the evolving best practice guide was used in various parts of the supplemental guide and subsequently refined appropriately for consistency with the remainder of the latter document. In the revised edition of the best practice guide, produced during SS*f*M-2, material from the drafts of the supplemental guide prepared during that period that had an origin in the first edition of this best practice guide was re-used in the current edition. A similar statement applies to the software specification document [24] produced in parallel with the best practice guide.

A whole-day uncertainty workshop was organized by NPL as part of *UNCERT 2003* [53] (section 6.5.1). It included a presentation from the SS*f*M programme on the role of probability in uncertainty evaluation (section 1.5.3). Of considerable interest to the delegates was the part of this presentation showing how lower and upper limits for the value of a quantity together with a measurement of the value of the quantity can be combined to provide a probability density function for the value of the quantity that is neither rectangular nor Gaussian, but in fact asymmetric.

Some of the material from the SS*f*M course on uncertainties and statistical modelling was used on the AMCTM short course on uncertainty evaluation [17] (section 6.5.3), developed jointly with PTB and CNR-IAC (Rome).

The best practice guide and software specification were contributed to the MetroNet cooperative project (COP3) on best practice in uncertainty evaluation (section 6.5.3). NPL was appointed to lead COP3.

1.5.7 Industrial uncertainty methodology

Representation was obtained on BSI Committee SS/4, Statistical Process Control, which tracks the corresponding international committee ISO/TC 69/SC 4 in ISO/TC 69, the Application of Statistical Methods. Reviews of ISO DIS 10012, Measurement control systems, and ISO 5725, Accuracy (trueness and precision) of measurement methods and results, were completed and their impact on the work assessed. Two papers relating to the methodology were presented at the international conference *UNCERT 2003* (section 6.5.1) on the uncertainty of measurement [53]. One paper was concerned with the general methodology [44] and the other with a plea for simplicity from an engineering stance, by SS*f*M collaborator Rolls Royce plc [57]. A draft of the former was circulated to the International Aerospace Quality Group and the comments received were incorporated in the paper. A further paper covering a forthcoming practical guide on uncertainty for industrial users [62] was also presented.

1.6 Data fusion

An important application of data fusion techniques considered within SS*f*M-2 was the least-squares adjustment of the estimated values of the fundamental constants. There are many fundamental constants of physics, a small number of which, such as the speed of light in a vacuum, have exact assigned values, the others being estimated from measurements provided by a large number of experiments. Generally, each experiment provides an estimate of the value of a quantity, related to some subset of the fundamental constants, and an associated standard uncertainty. From this ensemble of measurements, estimates of the constants are determined or 'adjusted' by solving a large non-linear least-squares problem. Of particular concern is the effect on the adjusted values arising from suspect data.

A number of adjustment algorithms based on generalized maximum likelihood approaches were developed. They were described in a presentation 'Least-squares adjustment in the presence of discrepant data' at the international conference *Advanced Mathematical and Computational Tools in Metrology* in Turin in September 2003 [26] (section 6.5.3). Other approaches developed as part of the work on non-Gaussian measurement errors (section 1.2.2) were also found to be effective. It is anticipated that this work will provide input to future adjustment exercises.

The SS*f*M-2 work on data fusion is also contributing to a European project involving Helsinki University of Technology, TNO, The Netherlands, Technical University of Darmstadt, City University, London, and NPL on modelling vision sensitivity in half-light conditions. The project involves fusing data supplied by a number of laboratories using different experimental equipment and methodologies representing observer reaction times to visual stimuli in varying lighting conditions.

The results of the work will be contributed for review and comment to the MetroNet cooperative project (COP4) on discrete modelling, data fusion and data fitting (section 6.5.3).

1.7 Visual modelling and uncertainties

1.7.1 Visualization of uncertainties in discrete models

A paper [61] on the SS*f*M work on the visualization of uncertainties in the context of discrete modelling was given exposure at the international conference *UNCERT 2003* (section 6.5.1) on the uncertainty of measurement [53]. It covered the use of Monte Carlo simulation to evaluate the uncertainties associated with the magnitude and phase of a complex-valued quantity from the uncertainties associated with measurements of its real and imaginary components. Such a requirement occurs in several areas of measurement, and is particular prevalent in electrical metrology and electrical engineering.

1.7.2 Visualization of uncertainties in continuous models

Work on the visualization of uncertainties associated with continuous models was carried out collaboratively with the Thermal programme. It was presented as a poster, ‘Simulation of the laser flash thermal diffusivity measurement process using a finite difference model’, at the BIPM-NPL workshop *The Impact of Information Technology in Metrology* at NPL in September 2002 (section 6.5.1).

2 Validation and testing

2.1 Numerical software testing

2.1.1 Reference data sets

Software for the dynamic generation of reference data sets and corresponding reference results for testing the numerical correctness of software for ten mathematical calculations was completed and is available on the SS*f*M website. The calculations covered include simple statistical calculations (sample mean and sample standard deviation), a number of regression calculations (involving polynomial and Gaussian peak models), and geometric element fitting (Gaussian and Chebyshev best-fit lines and circles). The software is implemented in the Java programming language and executed as an applet from within an HTML document. It provides the ability to generate data sets appropriate to the user’s application. A presentation of the work was made at the BIPM-NPL workshop *The Impact of Information Technology in Metrology* at NPL in September 2002 (section 6.5.1) to launch the data generators.

A further, invited, presentation on reference data sets was made in September 2002 to Nuovo Pignone GE Oil and Gas in Turin.

2.3 Testing algorithms in standards and MetroS

A paper [21] on the application of the methodology for testing the numerical correctness of metrology software, concerned with the validation of CMM form and tolerance assessment software, was presented at *Lamdamap 2001, the Fifth International Conference on Laser Metrology and Machine Perfomance* in Birmingham in July 2001.

One case study considered concerned algorithms for fitting complex surfaces to probe-centre data gathered by a CMM. Using data generation techniques, it was possible to provide, in a generic and straightforward way, data sets with which to test fitting algorithms based on approximate computational methods. These techniques were presented in a paper 'Correctness of free form surface fitting software' at the *Sixth International Conference on Laser Metrology and Machine Performance* in Huddersfield in July 2003 [46].

Valuable collaboration was secured with the Southern Methodist University in Dallas, Texas, with strong expertise in the numerical solution of ordinary differential equations. It is expected that the university will provide input to future SS*f*M work on the testing of algorithms for continuous models.

The SS*f*M reports on the methodology for testing spreadsheets and other packages used in metrology [16] and testing algorithms in standards and MetroS [2] were contributed to the MetroNet cooperative project (COP6) on a validation framework for metrological software (section 6.5.3).

2.4 Maintenance of measurement system validation best practice guide

SS*f*M Best Practice Guide No. 1, Measurement system validation: validation of measurement software [70], was revised after intensive review by experts in safety-critical systems. It was also revised in terms of the international safety systems standard IEC 61508 and CASS (Conformity Assessment of Safety-related Systems). This revision also took into account feedback from the nuclear and gas industries. The revised guide is available on the SS*f*M website (www.npl.co.uk/ssfm) for further comment.

This guide was contributed for review and comment to the MetroNet cooperative project (COP6) on a validation framework for metrological software (section 6.5.3).

2.6 Validation of self-calibrating instruments

The issues in validation of complex instruments that involve some aspects of self-calibration or self-validation were investigated. The work centred on a collaborative case study between NPL and NMi. They studied a complex self-calibrating electrical meter: NMi attempted to validate the calibration and measurement process in the hardware, and NPL concentrated on the same process in the software. There was considerable overlap between the two activities and two review teams were able to answer questions and suggest lines of enquiry for one another. It is planned to produce a joint report and a journal paper describing the work and its conclusions.

3 Metrology software and algorithm development techniques

3.2 Maintenance of existing guides

The earlier SS*f*M Best Practice Guides No. 2 on virtual instruments [42], No. 3 on metrology software development [7] and No. 8 on mixed language programming [48] were combined into a Best Practice Guide No. 12 on test and measurement software [15]. This guide was contributed to the MetroNet cooperative project (COP1) on data acquisition and virtual instruments (section 6.5.3).

4 Standards support and development

4.1 Mathematical and statistical input to metrology standards

There has been considerable activity internationally in standards organizations that relate to generic metrology issues, particularly concerning uncertainty. It is summarized here at the international, European and national levels (the last-mentioned being relevant regarding the UK input to and account of international standards). SS*f*M has representation on all the committees indicated.

International Standards Organization (ISO) The remit of WG 7 of SC 6, Proficiency of test methods, of ISO TC 69, Application of statistical methods, is to provide a bridge between the GUM and ISO 5725, Accuracy (trueness and precision) of measurement methods and results. Two documents were completed in 2003 by WG 7, the convenor of which is provided by SS*f*M, with publication by ISO expected in 2004. ISO TS 21748 [50] provides guidance on the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation. ISO TS 21749 [51] covers measurement uncertainty for metrological applications–simple replication and nested experiments. The former document will prove valuable when the GUM model-based approach is not immediately applicable, and the latter is relevant to the evaluation of uncertainty components based on suitable experimental designs.

Participation of a member of the SS*f*M team in ISO TC213/WG10, Dimensional and geometrical co-ordinate measurements, was supported by the Length Metrology programme. The working group is concerned with CMMs, arguably one of the most highly developed metrology discipline in terms of providing approaches for the evaluation of uncertainties associated with measurements made by advanced instruments.

An invited presentation on the development of the GUM supplemental guides (section 1.5.3) was made at the INNOVATION Workshop *ISO 17025, General requirements for the competence of testing and calibration laboratories* in Rotterdam in December 2001.

World Anti-Doping Agency (WADA) Doping in sport is commonplace today. A member of the SS*f*M team was called upon to act as an expert witness in two court cases concerning alleged offences involving metabolites of the banned substance nandrolone. One trial was held in Lausanne and the other was to be held in Montreal, but the case was dropped following submission of additional evidence. In any particular case it is necessary to compare the test measurement based on an athlete's sample with a threshold value imposed by the sporting body. If the test measurement exceeds the threshold, taking account of the uncertainty of measurement, the test is deemed positive. In some cases it would appear that the quoted uncertainty associated with the test result is invalid, in which case the decision made based on the comparison may be incorrect. Efforts are being made to encourage improved evaluation of uncertainty by the testing body, so that fairer results are achieved. Interlaboratory comparisons are required in the area. The meaningful accreditation of testing laboratories is essential.

Comité Européen de Normalisation (CEN) A member of the SS*f*M team sits on CEN BT/WG 122 concerned with measurement uncertainty. This group carried out a detailed survey of the extent to which the various CEN technical committees consider uncertainty in the norms they produce. The usage is very uneven. A CEN workshop was held in Brussels in June 2002 to provide the results of the survey and for various CEN sectors to make uncertainty presentations. SS*f*M provided an invited paper, 'The GUM and its supplemental guides–work in progress within JCGM' (cf. section 1.5.3).

European Co-operation for Accreditation (EA) SSfM provides a member on the EA Uncertainty Experts Group. The main work of this group was to produce a guidance document on the expression of uncertainty in quantitative testing [41]. It provides a testing counterpart of the respected EA/4-02, Expression of the Uncertainty of Measurement in Calibration [40].

British Standards Institution (BSI) SSfM provides the convenor for SS/6/-/3, the uncertainty panel of BSI committee SS/6, Precision of test methods. A document, PD 6461-4, a guide directed at industrial uncertainty requirements, was finalized, with publication expected in 2004. The content of GUM Supplemental Guide No. 5 (section 1.5.3) is expected to relate to the material in this document. A demonstration from the SSfM programme on software for the propagation of distributions (cf. section 1.5.3) was given to the panel.

SSfM is also represented on SS/4, dealing with statistical process control and providing input to the corresponding ISO/TC 69/SC 4 committee (section 1.5.7).

4.2 Standard approaches to uncertainty in calibration and conformity

SSfM sub-contractor signalsfromnoise.com Ltd. provided a report on the use of metrology-related software by standards. The report contains a number of recommendations, especially relating to standards committees taking into account software requirements, including the use of common algorithms, at the drafting stage. Discussions have since taken place with BSI concerning the dissemination of these recommendations.

4.2.1 Regression routines for use by standards

The SSfM report [29] on the classification and solution of regression problems for calibration contains problem formulations that are consistent with GUM terminology and recommends approaches consistent with state-of-the-art numerical analysis. In particular, detailed treatment is given of correlated quantities and of calibration data where not only the response values but also the stimulus values are subject to uncertainty. Several organizations, including NMIs and standards bodies, have stated the importance of least-squares calibration curves in metrology. This work will assist in underpinning future activity involving particularizations of the generic results presented, especially to the proposed standard for linear calibration curves (section 1.2.1). Further, the JCGM (section 1.5.3) will be making use of this work in developing Supplemental Guide No. 6 to the GUM.

4.2.2 Uncertainty in product and equipment conformity

SS*f*M sub-contractor BMTA (British Measurement and Testing Association) completed a survey of existing standards work in product and equipment conformity. This report will be helpful in informing the JCGM activity in developing Supplemental Guide No. 3 to the GUM (section 1.5.3). The report contains a number of recommendations and also questions that have not yet been answered satisfactorily by the many organizations involved, from regulators to standards-makers to manufacturers. A main conclusion is that, although a probabilistic treatment should form the basis for procedures in the area, easy-to-use software will be needed that implements these procedures.

5 Support for NMS infrastructure

5.1 Use of the internet by calibration services

A number of internet-enabled calibration systems have been developed since the original iPIMMS impedance measurement system that is now in operation as a commercial service. The various new systems and services provided by NPL and internationally are described in a survey of international activities in internet-enabled metrology [59]. SS*f*M was involved in work on these systems in the areas of software development, test and validation, auditing of security and data warehousing, and guidance on accreditation. This work has led to recommendations for the security of internet-enabled metrology systems and the representation of calibration data.

A study was made of transmission and warehousing of measurement data using as an example the iVR internet-enabled calibration system for Voltage and Resistance measurement. This study included a review of the security of the internet-enabled calibration system with input from the iVR project team members, a review of the software and documentation. A generic security policy was defined for calibration services over the internet, applicable to transmission of calibration data over the internet and storage of calibration data in a database. One important conclusion is to use up-to-date security mechanisms configured in line with security industry best practice, rather than adopting the default configuration options.

Issues of electronic calibration certificates and firewalls were also investigated. Various requirements for electronic calibration certificates were identified, with input from discussion of the issues in the NPL Measurement Services Forum. XML is recommended for the representation of calibration data for both certificates and storage, providing a re-usable solution to the production and communication of electronic calibration certificates.[4] For firewalls, restriction to the use of HTTP is recommended to provide internet-enabled calibration services without going beyond what they use to permit normal e-mail and web usage.

The results of these studies were documented in a report on protecting the data [1].

A presentation on internet-enabled metrology was made to the international workshop *Internet Metrology and Self-Validating Instruments* held at NMi in Delft in May 2002. A similar presentation, accompanied by demonstrations, was given at the BIPM-NPL workshop *The Impact of Information Technology in Metrology* (section 6.5.1).

The international survey report [59] and the report on protecting the data [1] were contributed for review and comment to the MetroNet cooperative project (COP2) on internet-enabled measurement (section 6.5.3). NPL was appointed to lead COP2.

6 Generic transfer technology activities

6.1 SS*f*M and MetroS website

During SS*f*M-2, the METROS website was superseded by EuroMETROS, a EUROMET project supported by fourteen NMIs, which NPL leads and hosts.

A number of contributions were made to the EuroMETROS website, a facility that is central to the delivery of algorithms, software, testing and guidance from the SS*f*M programme. Contributions include SS*f*M software and reference data sets and non-SS*f*M software produced at NPL before and during the SS*f*M programme. Also, contributions were made from

[4] Extensible Markup Language (XML) is designed to be a simple, very flexible text format derived from SGML (ISO 8879). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.

outside the SSfM programme. In particular, NRC contributed the QDE toolkit, which provides facilities for key comparison data evaluation (section 1.5.4).

Particular items provided to EuroMetroS from within the SSfM programme include (a) software implementations of knot-placement algorithms for free-knot univariate least-squares data approximation by polynomial splines (section 1.1.1), (b) software for extracting features from univariate polynomial curves fitted to experimental data (section 1.1.3), (c) software for generalized regression with polynomials (section 1.2.1), (d) software for automatic differentiation (section 1.2.3), (e) generators of reference data sets and corresponding reference results for ten metrology calculations (section 2.1.1), and (e) NPLFit for fitting polynomial and spline models and performing auxiliary computations.

GENLINE, software for generalized regression (section 1.2.1), was provided to NIST and is being used there for gas-standards work.

6.2 SSfM Club

The international collaborators listed in appendix A are members of the SSfM Club. These participate mostly by e-mail and access to material on the SSfM website. Also many have attended SSfM workshops and training courses, as well as the international conferences supported by the programme.

6.3 Counting on IT newsletter

The SSfM newsletter, *Counting in IT*, now issued twice a year, is sent to over 1600 contacts, of which over 450 are international contacts. It has from time to time included articles from SSfM's international collaborators.

6.4 Reports on International Activities

EasyTEST In September 2001, a Framework 6 proposal, EasyTEST, was submitted by UNIMETRIK, Spain, with European partners including NPL and several aerospace and automotive manufacturers, for a project on CMM software verification for Intelligent Manufacturing Systems. The UK input would mainly relate to the application to CMMs of the software-testing methodology developed under SSfM. This proposal just failed to achieve the evaluation score required for European support.

Review of 'Metrological IT' at PTB On behalf of the German Federal Ministry of Economics and Technology an international expert commission evaluated PTB in August 2002 in a number of areas, one of which was 'metrological information technology'. A member of the SSfM team sat on the sub-commission. The sub-commission's tasks were (1) to comment on the expected future development of the field, (2) to evaluate PTB's present activities against this expected development, and (3) to derive recommendations for PTB's future activities. The sub-commission's report is confidential, making recommendations under the headings (a) Software quality and data security, (b) Data analysis and modelling, (c) General recommendations.

ISO standard for testing CMM software ISO 10360-6, *Acceptance test and reverification test for coordinate measuring machines (CMM), Part 6: Estimation of errors in computing Gaussian associated features*, was published in 2002. It is based on the testing principles for the correctness of numerical software promoted in SSfM.

6.5 Awareness presentations and conferences

6.5.1 SSfM presentations to relevant meetings, workshops, conferences

Swedish Calibration and Measurement Conference A member of the SSfM team attended the Swedish National Measurement Conference in May 2001 (supported by SP, the Swedish NMI), delivering an invited presentation on the evaluation of interlaboratory comparison data.

Metrologie 2001 A member of the SSfM team attended the International Metrology Congress, Metrologie 2001, in Saint-Louis, France, in October 2001, giving a presentation on Monte Carlo simulation for uncertainty evaluation

NMC 2001 SSfM contributed to the *National Measurement Conference (NMC 2001)* in Harrogate in November 2001, with presentations on 'Models for metrology and their validation', 'Numerical algorithm design and testing' and 'Numerical software testing', and by holding an uncertainties workshop. PTB collaborated with NPL on a presentation on the role of Monte Carlo simulation in uncertainty evaluation.

Joint VNIIM–NPL conference NPL collaborated with VNIIM on the joint organization of a two-day conference *Mathematics, Statistics and Computation to Support Measurement Quality* in St Petersburg in June 2002. The aim was to discuss measurement quality characteristics, modelling and estimation of model quality in solving measurement problems, metrological certification of data processing software, and statistical methods of data processing. There was also a discussion on key comparisons.

Two SS*f*M papers were presented and will appear in the proceedings of the conference. They relate to the work carried out in project 1.5, *Uncertainties and statistical modelling*. The first paper [32], 'Technical aspects of guidelines Guidelines for the statistical analysis of key comparisons', detailed the technical underpinning to the key comparison data evaluation guidelines [19, 20] developed by BIPM Director's Advisory Group on Uncertainty (section 1.5.4). The second paper [35], 'Evaluation of measurement uncertainty based on the propagation of distributions using Monte Carlo integration' (with collaborator PTB), is concerned with the approach developed as the basis of the first supplemental guide to the GUM, concerned with the use of Monte Carlo simulation for the propagation of distributions.

A benefit of the conference was the opportunity for western European metrologists to discuss infrastructure issues with VNIIM and its sister laboratories and with eastern European academics. Russia has made a large investment in an approach to uncertainty evaluation that is conceptually different from that of the GUM. The expense of changing all relevant documents and procedures to align with the GUM would be considerable. Yet there is sympathy in some Russian quarters for the 'GUM approach'.

Joint BIPM–NPL workshop on the Impact of Information Technology in Metrology NPL hosted the first international workshop *Impact of Information Technology in Metrology* over four days in September 2002. The workshop was organized as part of the SS*f*M programme and of the activity of BIPM's Information Technology Group. Two other organizations, SS*f*M sub-contractor NAG Ltd. and the European thematic network SofTools MetroNet (section 6.5.3), sponsored the workshop. 110 delegates from 22 countries attended. In his opening address, Bob McGuiness, NPL director, referred to the increasing role played by information technology in almost all phases of measurement, from data acquisition and instrument control to data processing and analysis to the production of calibration certificates. In the context of applying software and IT best practice more widely across all areas of metrology, he also stated that the SS*f*M programme undertakes relevant generic work that complements and supports the vertical metrology programmes, identifying and disseminating best practice.

The main workshop aims were judged to have been fully met. They were to

1. Provide an overview of the software support that is currently provided to metrology
2. Identify requirements, opportunities and benefits of IT for metrology
3. Facilitate the cross-fertilization of IT within metrology institutes and other organizations
4. Consider the state of the art in and opportunities for internet-enabled metrology.

In the main part of the meeting there were contributions from the SS*f*M programme on (a) The impact of the SS*f*M programme, (b) Rapid advances in computer graphics giving opportunities in metrology, (c) An underpinning numerical software environment for metrology, (d) Advanced software tools for mathematics and statistics in metrology: the MetroNet sub-network and EuroMetroS, (e) Algorithms and software for large-scale calibration problems, (f) A guide to measurement system validation, and (g) A web-based mechanism for generating reference data sets for testing the numerical correctness of metrology software. Delegates were subsequently provided with a CD-ROM containing material provided by speakers.

One day was devoted to internet-enabled metrology (cf. section 5.1). NPL gave a tutorial on the concepts, and described both the iPIMMS service for impedance calibration of vector network analyzers and the iColour development for calibration and monitoring of spectrophotometers. Fluke gave a presentation on NPL's iVR prototype service for voltage and resistance calibration, which uses the Fluke 4950 multi-function calibrator as a transfer standard. Adelard gave a companion presentation on the associated iCal software, which has been produced to be readily adaptable and easily maintainable for use in any internet-enabled calibration or monitoring service. Information was provided by NIST on the focus it has on using internet-enabled metrology for regional interlaboratory comparisons.

NIST has developed a dosimetry calibration service, and a service using digital multi-meters. PTB is developing services for remote equipment monitoring, e.g., CMMs, with an emphasis on increasing the frequency of checks, improving traceability and hence improving product quality. PTB's work includes the development of more robust travelling artefacts, such as a ball-cube containing sensors to monitor temperature, humidity and vibration. NMi collaborates with Fluke on internet-enabled maintenance of a self-calibrating electrical calibrator, the Fluke 5700A. Four levels of checking are involved: internal checks, self-calibration, internet-enabled calibration, and conventional calibration. This approach enables the user to obtain a good balance between cost and uncertainty.

NMIJ, together with AIST, has a 2001–2005 programme of development of internet-enabled metrology applications, called 'e-trace'. Areas covered are time and frequency, length through optical wavelength, AC voltage, radioactivity, CMMs and gas and liquid pressure.

A second international workshop at NPL on the Evaluation of Interlaboratory Comparison Data constituted a satellite meeting to the IT workshop. Opened by Andrew Wallard, then Deputy Director of NPL and now Director of BIPM, it was a follow-up to the event under a similar title held at NPL in November 1999. Presentations were made by the major NMIs on approaches to comparison data evaluation that ranged from classical and robust estimators of reference values to the linking of comparisons to detailed considerations of the factors that influence participants' results. The determination of the consequent degrees of equivalence was also much discussed. As anticipated, diverse views were expressed that were influenced by scientific and political considerations. The results of key comparisons are arriving at the BIPM at an average of one per week, but the data evaluation aspects are not receiving the attention they deserve in terms of informed generic input. The BIPM Director's Advisory Group on Uncertainties is providing some assistance through advice and written guidelines (section 1.5.4).

ILAC/IAF Conference on Accreditation in Global Trade The paper 'The GUM and its planned supplemental guides' [31] was presented by invitation at the *International ILAC/IAF Conference on Accreditation in Global Trade*, in Berlin in September 2002. It covered the scope of the supplemental guides to the GUM being developed by the JCGM (section 1.5.3).

UNCERT 2003 The *UNCERT 2003* international conference and workshop [53] organized by NPL at Oxford in April 2003 included a number of papers from the SS*f*M programme, some jointly with its collaborators. Two of these papers were prepared with Italian colleagues on (1) Order statistics and uncertainty evaluation (IMGC) [58], with emphasis on key comparison data evaluation, and (2) Distribution-free bound on the level of confidence of a measurement process (Universities of Florence and Bologna) [37], with application to coverage factors and coverage regions for models with more than one output quantity. Both papers account for correlation effects.

6.5.2 A4A4 conference

The Algorithms for Approximation series of conferences are concerned with numerical methods for approximating functions and data, the latter being synonymous with the SSfM activity on discrete modelling. NPL has played a strong role in all the meetings and co-sponsored the July 2001 meeting, held at the University of Huddersfield, with 100 delegates, at which four papers [33, 43, 45, 54] were authored or co-authored by NPL.

The first three of these papers relate to the work carried out in the SSfM project *Algorithms for discrete modelling* and the last to the SSfM project *Extension of empirical models.* The paper 'Generalised Gauss-Markov regression' concentrates on the valid treatment of data having various uncertainty structures. The paper 'Efficient algorithms for structured self-calibration problems' covers the efficient handling of large-scale calibration problems whose observation matrices have various types of structure. The paper 'Fixed- and free-knot univariate least-squares data approximation by polynomial splines' deals with the manner in which the very versatile spline functions can be fitted to data, particularly the case where the knots or joins of the spline are not specified *a priori.* The paper 'A robust algorithm for least absolute deviation curve fitting' (jointly with the University of Huddersfield) provides a more stable solution algorithm than was available hitherto for data approximation problems where the data may contain outliers or 'wild points'.

Two metrology sessions were included, at which data approximation and related disciplines specifically applied to metrology were discussed. A 'Statistics and Uncertainties' session was designed to cover statistical models in data fitting and uncertainties associated with the fitted parameters.

6.5.3 AMCTM conference

Six Advanced Mathematical and Computational Tools in Metrology (AMCTM) conferences have been held, at an interval of approximately two years, within current and previous European thematic networks. The proceedings of five of these conferences have appeared [9, 10, 11, 12, 13] and those of the sixth (below) are in preparation. They are the premier metrology events worldwide of their type. The latest conference, in Turin in September 2003, with 100 delegates, was again organized with input from NPL. The general aims of the conference were those of previous years[5], but augmented by coverage of both the fields of calibration and testing.

The NPL presentations were on automatic differentiation and its application to metrology [6] (section 1.2.3), least-squares adjustment in the presence of discrepant data [26] (section 1.2.2), estimators for mixed noise distributions, with the University of Huddersfield [52] (section 1.5.1), harmonization of correlated calibration curves with application to the analysis of natural gases [38], with the VAM programme (section 1.5.5).

In addition, PTB and NPL introduced the roundtable discussion on Software Validation (section 2.1) and NPL contributed to the roundtable discussion on testing.

Associated with the conference was a short course on uncertainty evaluation [17], with tutors provided by IAC, NPL and PTB. The course contents, which aligned closely with the uncertainties and statistical modelling course provided within SS*f*M, was (a) provide the probabilistic basis of uncertainty evaluation, (b) present a formulation of the generic problem of uncertainty evaluation that is consistent with this basis, (c) provide implementations (computational tools) for solving this problem based on the law of propagation of uncertainty and the propagation of distributions, (d) apply these implementations to uncertainty evaluation problems arising in several branches of metrology, and compare them, (e) present and show the role of computational tools such as bootstrap re-sampling in the context of uncertainty evaluation, (f) emphasise the alignment of the coverage to the GUM and the supplemental guides being developed to support it. NPL provided the course chairman and two of the talks.

SofTools MetroNet Advanced Mathematical and Computational Tools in Metrology (AMCTM), also known as SofTools MetroNet (or MetroNet for short), is a European Thematic Network founded in 2002. It aims to foster and improve co-operation in the field of advanced

[5](a) To present and promote reliable and effective mathematical and computational tools in metrology, (b) to understand better the modelling, statistical and computational requirements in metrology, (c) to provide a forum for metrologists, mathematicians and software engineers that will encourage a more effective synthesis of skills, capabilities and resources, (d) to promote collaboration in the context of EU and International Programmes, (e) to support young researchers in metrology and related fields, also through training, (f) to address industrial requirements.

measurement-software tools and their mathematical and statistical foundations. To this end, scientists from the mathematical and computational fields in universities, research centres, industrial organizations and the metrological field meet to share experiences between themselves and with software engineers. There are four areas where SofTools MetroNet activities concentrate in co-ordinating, promoting, testing, evaluating, harmonizing and in performing educational tasks:

1. Advanced Tools for Instrumentation: Work Package 1 (WP1)
2. Advanced Tools for Mathematics and Statistics: WP2
3. Advanced Tools for Metrological Databases and Software Validation: WP3
4. Training and Education, Dissemination, Exploitation and IPR: WP4

To encourage the submission of new procedures and software (in particular, within WP1 and WP2), MetroNet has defined six Co-operative Projects (COPs) (appendix B). The goals of the COPs are (a) the development of new procedures and tools, and (b) their testing and reporting on the testing. They are intended to constitute joint activities to solve specific problems in the areas of data acquisition, data reduction, statistical treatment and data retrieval. The solution process will benefit from collaborative efforts in developing new criteria, new procedures, new algorithms and new software tools.

Much of the material covered by the COPs relates closely to the SS*f*M programme.

Two COPs were launched for each WP, WP leaders being appointed, from the following organizations:

WP1: NMi, Netherlands and University of Lodz, Poland

WP2: NPL, UK and University of Genova, Italy

WP3: PTB, Germany and MIRS, Slovenia.

SS*f*M has participated in all MetroNet Steering Committee meetings and leads two of the COPs (appendix B).

6.5.4 Wilkinson Prize

In honour of the outstanding contributions of James Hardy Wilkinson to the field of numerical software, NPL, the Numerical Algorithms

Group (NAG Ltd.) and Argonne National Laboratory (US) award every four years a prize of 1000 US dollars for numerical software. The 2003 prize presentation took place in Sydney in July 2003 at the *International Congress for Industrial and Applied Mathematics.* Wilkinson spent his pioneering days at NPL, where he laid the basis for reliable numerical methods for linear algebra and floating-point error analysis. These concepts are central to today's quality software for solving linear and non-linear algebraic systems and least-squares problems and to the SS*f*M programmes.

The winner of the 2003 Prize was Jonathan Shewchuk of the University of California at Berkeley for *Triangle, two-dimensional mesh generator and Delaunay triangulator* [64]. *Triangle* generates high-quality unstructured triangular meshes. It also generates two-dimensional Delaunay triangulations, constrained Delaunay triangulations, Voronoi diagrams, and convex hulls. The software is used in discrete meshing for the discretization of partial differential equations, in solving classes of geometric problems, and the graphing of surfaces. It has thousands of users, and is downloaded on average more than 30 times a day. *Triangle* has been licensed for inclusion in eleven commercial software packages. The speed and accuracy of *Triangle* is a result of novel algorithms for extended precision floating-point arithmetic [63] and the use of adaptive computation controlled by forward error analysis [71].

Acknowledgements

The work described here was supported by the National Measurement System Directorate of the UK Department of Trade and Industry as part of its NMS Software Support for Metrology programme. Robin Barker, Alistair Forbes, Peter Harris and Dave Rayner contributed to this report.

References

[1] R. Barker and G. Parkin. Use of the internet for calibration services – protecting the data. Technical Report CMSC 28/03, National Physical Laboratory, Teddington, UK, 2003. 27

[2] R. M. Barker, M. G. Cox, P. M. Harris, and I. M. Smith. Testing algorithms in standards and MetroS. Technical Report CMSC 08/03, National Physical Laboratory, Teddington, UK, 2003. 22

[3] BIPM. Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes. Technical report, Bureau International des Poids et Mesures, Sèvres, France, 1999. 1, 12

[4] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, 1995. ISBN 92-67-10188-9, Second Edition. 1, 6

[5] K. G. Birch. Standards and software for evaluating uncertainty, a survey. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 11

[6] R. Boudjemaa, M. G. Cox, A. B. Forbes, and P. M. Harris. Automatic differentiation and its applications to metrology. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003*. To appear. 34

[7] D. Brinkley. Good Practice Guide No. 3. Guidance on developing software for metrology. Technical report, National Physical Laboratory, Teddington, UK, 2001. 23

[8] Anna Chunovkina and Maurice Cox. A model-based approach to key comparison data evaluation. In Damir Ilić, Mladen Boršić, and Josip Butorac, editors, *Proceedings XVII IMEKO World Congress, June 22-27, 2003, Dubrovnik, Croatia*, pages 1240–1243, Zagreb, Croatia, 2003. Croatian Metrology Society. 13

[9] P. Ciarlini, M. G. Cox, E. Filipe, F. Pavese, and D. Richter, editors. *Advanced Mathematical and Computational Tools in Metrology V. Series on Advances in Mathematics for Applied Sciences Vol. 57*, Singapore, 2001. World Scientific. 2, 34

[10] P. Ciarlini, M. G. Cox, R. Monaco, and F. Pavese, editors. *Advanced Mathematical Tools in Metrology*, Singapore, 1994. World Scientific. 2, 34

[11] P. Ciarlini, M. G. Cox, F. Pavese, and D. Richter, editors. *Advanced Mathematical Tools in Metrology II*, Singapore, 1996. World Scientific. 2, 34

[12] P. Ciarlini, M. G. Cox, F. Pavese, and D. Richter, editors. *Advanced Mathematical Tools in Metrology III*, Singapore, 1997. World Scientific. 2, 34

[13] P. Ciarlini, A. B. Forbes, F. Pavese, and D. Richter, editors. *Advanced Mathematical and Computational Tools in Metrology IV*, Singapore, 2000. World Scientific. 2, 34

[14] Patrizia Ciarlini, Maurice Cox, Franco Pavese, and Giuseppe Regoliosi. The use of a mixture of probability distributions in temperature interlaboratory comparisons. *Metrologia*. In press. 16

[15] T. Clement, L. Emmet, P. Froome, and Sofia Guerra. Good Practice Guide No. 12. Guide for test and measurement software. Technical report, National Physical Laboratory, Teddington, UK, 2002. 23

[16] H. R. Cook, M. G. Cox, M. P. Dainton, and P. M. Harris. A methodology for testing spreadsheets and other packages used in metrology. Technical Report CMSC 25/99, National Physical Laboratory, Teddington, UK, 1999. 22

[17] M. G. Cox. Short course on uncertainty evaluation. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003.* To appear. 19, 34

[18] M. G. Cox. Interim status report on software activity world-wide. Technical Report CISE 40/99, National Physical Laboratory, Teddington, UK, 1999. 1

[19] M. G. Cox. The evaluation of key comparison data. *Metrologia*, 39:589–595, 2002. 12, 14, 15, 30

[20] M. G. Cox. The evaluation of key comparison data: An introduction. *Metrologia*, 39:587–588, 2002. 12, 14, 30

[21] M. G. Cox, M. P. Dainton, A. B. Forbes, and P. M. Harris. Validation of CMM form and tolerance assessment software. In G. N. Peggs, editor, *Laser Metrology and Machine Performance V*, pages 367–376, Southampton, 2001. WIT Press. 4, 21

[22] M. G. Cox, M. P. Dainton, A. B. Forbes, P. M. Harris, P. M. Schwenke, B. R. L Siebert, and W. Wöger. Use of Monte Carlo simulation for uncertainty evaluation in metrology. In P. Ciarlini, M. G. Cox, E. Filipe, F. Pavese, and D. Richter, editors, *Advanced Mathematical Tools in Metrology V. Series on Advances in Mathematics for Applied Sciences Vol. 57*, pages 93–104, Singapore, 2001. World Scientific. 9

[23] M. G. Cox, M. P. Dainton, and P. M. Harris. Best Practice Guide No. 6. Uncertainty and statistical modelling. Technical report, National Physical Laboratory, Teddington, UK, 2001. 9, 18

[24] M. G. Cox, M. P. Dainton, and P. M. Harris. Software specifications for uncertainty evaluation and associated statistical analysis. Technical Report CMSC 10/01, National Physical Laboratory, Teddington, UK, 2001. 9, 18

[25] M. G. Cox, Michèle Desenfant, P. M. Harris, and B. R. L. Siebert. Model-based measurement uncertainty evaluation, with applications in testing. *Accreditation and Quality Assurance*, 8:548–554, 2003. 11

[26] M. G. Cox, A. B. Forbes, J. Flowers, and P. M. Harris. Least squares adjustment in the presence of discrepant data. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003.* To appear. 20, 34

[27] M. G. Cox, A. B. Forbes, and P. M. Harris. Best Practice Guide No. 4. Discrete modelling. Technical report, National Physical Laboratory, Teddington, UK, 2000. 4, 9

[28] M. G. Cox, A. B. Forbes, and P. M. Harris. Best Practice Guide No. 10. Discrete model validation. Technical report, National Physical Laboratory, Teddington, UK, 2001. 4

[29] 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. 3, 25

[30] M. G. Cox and P. M. Harris. GUM supplements. In *CIE Expert Symposium 2001 on Uncertainty Evaluation. Method for Analysis of Uncertainties in Optical Radiation Measurement, Vienna, Austria, January 2001*, pages 23–34, Vienna, Austria, 2001. CIE Central Bureau. 9

[31] M. G. Cox and P. M. Harris. The GUM and its planned supplemental guides. *Accred. Qual. Assur.*, 8:375–379, 2003. 9, 11, 32

[32] M. G. Cox and P. M. Harris. Technical aspects of guidelines for the evaluation of key comparison data. Izmeritelnaya Technika, 2003. In Press. 30

[33] M. G. Cox, P. M. Harris, and P. D. Kenward. Fixed- and free-knot univariate least-squares data approximation by polynomial splines. In I. J. Anderson J. Levesley and J. C. Mason, editors, *Algorithms for Approximation IV*, pages 330–345, Huddersfield, UK, 2002. University of Huddersfield. 2, 33

[34] M. G. Cox, P. M. Harris, and P. D. Kenward. Fixed- and free-knot univariate least-squares data approximation by polynomial splines. Technical Report CMSC 13/02, National Physical Laboratory, Teddington, UK, 2002. 2

[35] M. G. Cox, P. M. Harris, and B. R-.L. Siebert. Evaluation of measurement uncertainty based on the propagation of distributions using Monte Carlo integration. Izmeritelnaya Technika, 2003. In Press. 9, 30

[36] M. G. Cox, P. M. Harris, and B. R. L. Siebert. The use of Monte Carlo simulation to evaluate uncertainties of key comparison reference values and degrees of equivalence. In D. Richter, W. Wöger, and W. Hässelbarth, editors, *Data Analysis of Interlaboratory Comparisons*, pages 123–134, Bremerhaven, Germany, 2003. Wirtschaftsverlag NW. 14, 16

[37] M. G. Cox, G. Iuculano, Annarita Lazzari, and Gabriella Pellegrini. Distribution-free bound for the level of confidence of a measurement process. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 7, 33

[38] M. G. Cox, S. Kamvissis, M. J. T. Milton, and G. Vargha. Harmonization of correlated calibration curves with an application to the analysis of natural gases. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003.* To appear. 18, 34

[39] J. A. Davis, P. M. Harris, S. L. Shemar, and M. G. Cox. The characterization of regular but unevenly spaced TWSTFT data using second-difference statistics. *Metrologia*, 40:S312–S318, 2003. 6

[40] EA. Expression of the uncertainty of measurement in calibration. Technical Report EA-4/02, European Co-operation for Accreditation, 1999. 25

[41] EA. EA guidelines on the expression of uncertainty in quantitative testing. Technical Report EA-4/16, European Co-operation for Accreditation, 2003. 25

[42] L. Emmet and P. Froome. Good Practice Guide No. 2. The development of virtual instruments. Technical report, National Physical Laboratory, Teddington, UK, 1999. 23

[43] A. B. Forbes. Efficient algorithms for structured self-calibration problems. In J. Levesley, I. Anderson, and J. C. Mason, editors,

Algorithms for Approximation IV, pages 146–153, Huddersfield, UK, 2002. University of Huddersfield. 4, 33

[44] A. B. Forbes. Towards an industrial uncertainty methodology. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 19

[45] A. B. Forbes, P. M. Harris, and I. M. Smith. Generalised Gauss-Markov regression. In J. Levesley, I. Anderson, and J. C. Mason, editors, *Algorithms for Approximation IV*, pages 270–277, Huddersfield, UK, 2002. University of Huddersfield. 3, 33

[46] A. B. Forbes, P. M. Harris, and I. M. Smith. Correctness of free form surface fitting software. In D. G. Ford, editor, *Laser Metrology and Machine Performance VI*, pages 263–272, Southampton, 2003. WIT Press. 22

[47] P. M. Harris, J. A. Davis, M. G. Cox, and S. L. Shemar. Least-squares analysis of two-way satellite time and frequency transfer measurements. *Metrologia*, 40:S342–347, 2003. 6

[48] M. Hooper. Good Practice Guide No. 8. Mixed language programming. Technical report, National Physical Laboratory, Teddington, UK, 2000. 23

[49] ISO. ISO 17025. General requirements for the competence of testing and calibration laboratories, 1999. International Standards Organization, Geneva. 1

[50] ISO. ISO TS 21748. Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, 2004. International Standards Organization, Geneva. 23

[51] ISO. ISO TS 21749. Measurement uncertainty for metrological applications–simple replication and nested experiments, 2004. International Standards Organization, Geneva. 23

[52] D. P. Jenkinson, J. C. Mason, A. Crampton, M. G. Cox, A. B. Forbes, and R. Boudjemaa. Parameterized approximation estimators for mixed noise distributions. In *International Conference on Advanced Mathematical and Computational Tools in Metrology VI, Torino, September, 2003.* To appear. 4, 34

[53] National Physical Laboratory, editor. *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Teddington, UK, 2003. National Physical Laboratory. 6, 7, 11, 15, 19, 20, 32

[54] D. Lei, I. J. Anderson, and M. G. Cox. A robust algorithm for least absolute deviation curve fitting. In I. J. Anderson J. Levesley and J. C. Mason, editors, *Algorithms for Approximation IV*, pages 470–477, Huddersfield, UK, 2002. University of Huddersfield. 33

[55] M. J. T. Milton and M. G. Cox. Evaluating degrees of equivalence using 'exclusive' statistics. *Metrologia*, 40:L1–L2, 2003. 15

[56] A. Papoulis. *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, New York, 1984. 7

[57] C. Parrish. Measurement uncertainty within engineering. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 19

[58] Francesca Pennecchi, Maurice Cox, and Walter Bich. Order statistics and uncertainty evaluation. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 15, 32

[59] D. Rayner. Survey of international activities in internet-enabled metrology. Technical Report CMSC 21/03, National Physical Laboratory, Teddington, UK, 2003. 26, 27

[60] D. Richter, W. Wöger, and W. Hässelbarth, editors. *Data Analysis of Interlaboratory Comparisons*, Bremerhaven, Germany, 2003. Wirtschaftsverlag NW. 13, 14

[61] M. Salter. Calculating uncertainties in magnitude and phase by Monte Carlo Simulation. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 20

[62] J. Sattary and D. Boam. Practical guide to measurement uncertainty. In *Proceedings of UNCERT 2003, International Conference on the Uncertainty of Measurement*, Oxford, UK, 2003. National Physical Laboratory. 19

[63] J. R. Shewchuk. Adaptive precision floating-point arithmetic and fast robust geomet ric predicates. *Computational Geometry*, 18:305–363, 1997. 36

[64] J. R. Shewchuk. Delaunay refinement algorithms for triangular mesh generation. *Computational Geometry: Theory and Applications*, 22:21–74, 2002. 36

[65] B. R. L. Siebert, M. G. Cox, and P. M. Harris. Evaluation of key comparisons based on the probability distributions for the participating institutes' measurement quantities. In D. Richter, W. Wöger, and W. Hässelbarth, editors, *Data Analysis of Interlaboratory Comparisons*, pages 107–122, Bremerhaven, Germany, 2003. Wirtschaftsverlag NW. 14

[66] A. G. Steele, B. M. Wood, and R. J. Douglas. Exclusive statistics: simple treatment of the unavoidable correlations from key comparison reference values. *Metrologia*, 38:483–488, 2001. 15, 16

[67] A. M. H. van der Veen, Annarita Baldan, and M. G. Cox. Application of uncertainty evaluation principles to purity verification. *Metrologia*, 2004. Submitted. 17

[68] A. M. H. van der Veen and M. G. Cox. Degrees of equivalence across key comparisons in gas analysis. *Metrologia*, 40:18–23, 2003. 17

[69] A. M. H. van der Veen and M. G. Cox. Error analysis in the evaluation of measurement uncertainty. *Metrologia*, 40:42–50, 2003. 6

[70] B. Wichmann. Best Practice Guide No. 11. Measurement system validation: validation of measurement software. Technical report, National Physical Laboratory, Teddington, UK, 2002. 22

[71] J. H. Wilkinson. *Notes in Applied Science No. 32. Rounding Errors in Algebraic Processes.* Her Majesty's Stationery Office, London, 1963. 36

[72] W. Wöger. Probability assignment to systematic deviations by the Principle of Maximum Entropy. *IEE Trans. Instr. Measurement*, IM-36:655–658, 1987. 10

A International Collaborators of the Software Support for Metrology Club

AIST, JAPAN

BAM, GERMANY

BIPM, FRANCE

Centro Espanol de Metrologia, SPAIN

CERN, SWITZERLAND

Chalmers University of Technology, SWEDEN

CNAM, FRANCE

CNR - Istituto di Metrologia, ITALY

CSIRO - National Measurement Lab., AUSTRALIA

D.I. Mendeleyev Institute for Metrology, RUSSIA

Danish Institute of Fundamental Metrology, DENMARK

Drexel University, USA

IEN G Ferraris, ITALY

Industrial Research Ltd, NEW ZEALAND

Instituto per le Applicazioni del Calculo, ITALY

Instituto Portugues da Qualidade, PORTUGAL

Justervesenet, NORWAY

Laboratoire National d'Essais, FRANCE

Measurement Standards Laboratory of New Zealand, NEW ZEALAND

Metrology Institute of the Republic of Slovenia, SLOVENIA

Mitutoyo Messgerate GmbH, GERMANY

National Institute of Metrology, CHINA (PEOPLES REPUBLIC)

National Metrology Institute of Japan, JAPAN

National Metrology Laboratory, IRISH REPUBLIC

National Metrology Laboratory, SOUTH AFRICA

National Research Council Canada, CANADA

Nederlands Meetinstituut, NETHERLANDS

NIST, USA

NIST (Boulder), USA

NMi Van Swinden Laboratorium, NETHERLANDS

NRC, CANADA

Physikalisch-Technische Bundesanstalt, GERMANY

Real Instituto y Observatorio de la Armada, SPAIN

Shanghai University, CHINA (PEOPLES REPUBLIC)

SP Swedish National Testing and Research Insititute, SWEDEN

State Research Centre of Russia "Electropribor", RUSSIA

Technical University of Lodz, I-12, POLAND

Technische Universiteit Eindhoven, NETHERLANDS

UNIMEP, BRAZIL

University of Genova, ITALY

B MetroNet Co-operative Projects

The MetroNet Co-operative Projects (COPs) (section 6.5.3) and the tasks defined within them are

COP 1. Data-acquisition and virtual instruments

Leader: University of Lodz

1. Good practice guidance for data-acquisition procedures (including virtual instruments)
2. Application of existing uncertainty best practice to measurement set-ups
3. Actual (informal) evaluations of data-acquisition procedures and software

COP 2. Internet-enabled measurement

Leader: NPL

1. Good practice for implementation of internet-enabled measurement software
2. Specification for an application programming interface (API) to be used in internet-enabled measurement
3. Applicability of internet-enabled measurement

COP 3: Best practice in uncertainty evaluation

Leader: NPL

1. Uncertainty evaluation in the context of testing
2. Methods for expressing and evaluating uncertainties
3. Methods for interlaboratory (including key) comparison data evaluation
4. Methods for expressing and evaluating uncertainties for multivariate models, including spectral quantities

COP 4: Discrete modelling, data fusion and data fitting

Leader: IMGC

1. Review and evaluation of algorithms and software
2. Best practice guidance
3. Case studies on the use of the guidance, algorithms and/or software in metrology

COP 5: Databases in metrology

Leader: PTB

1. Collection and review of existing solutions (data models, interfaces, application areas)
2. Exercises using selected solutions
3. Development of recommendations for using standard/non-standard databases

COP 6: Validation framework for metrological software

Leader: PTB

1. Development of a common understanding of validation of metrological software
2. Evaluation of validations performed/further examples that will be performed
3. Classification of requirements/user needs
4. Development of criteria.