NPL Report CMSC 09/01

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

TESTING SPREADSHEETS AND OTHER PACKAGES USED IN METROLOGY

REPORT ON DISCUSSIONS WITH OTHER NATIONAL MEASUREMENT INSTITUTES (2000/2001)

BY

M G COX and P M HARRIS

February 2001

Testing Spreadsheets and Other Packages Used in Metrology: Report on Discussions with other National Measurement Institutes (2000/2001)

M G Cox and P M Harris

Centre for Mathematics and Scientific Computing

February 2001

ABSTRACT

The NMS Software Support for Metrology Programme (April 1998 to March 2001) is a UK government-supported programme designed to tackle a wide range of generic issues arising from the use of mathematics and software in measurement science. Project 2.1 of this programme is concerned with the *numerical correctness* of software used in metrology. An activity within the project is to enable discussion with other National Measurement Institutes concerning the project results, and to account for work undertaken by other National Measurement Institutes in this area. The purpose of this report is to inform those interested in the NMS Software Support for Metrology programme about such discussions and interactions. Three such reports were to be produced, one at the end of each year of the programme, and this report is the third of these, covering the year 2000/2001.

This report constitutes part of the deliverable of Project 2.1 "Testing Spreadsheets and Other Packages Used in Metrology" within the UK Department of Industry's National Measurement System *Software Support for Metrology* Programme 1998–2001.

ISSN 1471-0005

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

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

Contents

1. Introduction 1

2. Activity for the Year 2000/2001 1

2.1 EuroConference AMCTM2000 1

2.2 CMM Software Testing Standard 2

2.3 Contacts with International Associate Members of SSfM Club 2

2.4 METROS and EUROMETROS 4

3. Summary 5

4. Acknowledgements 5

5. References 6

1. Introduction

The NMS Software Support for Metrology Programme (April 1998 to March 2001) is a UK government-supported programme designed to tackle a wide range of generic issues arising from the use of mathematics and software in measurement science. One of the issues addressed in the programme is the *numerical correctness* of software used in metrology and, in particular, the numerical correctness of the intrinsic functions provided by proprietary software packages such as spreadsheets. Project 2.1 of the Software Support for Metrology (SS*f*M) programme is concerned with this issue. The main project aims are to

a) develop a methodology for testing the numerical correctness of metrology software using, for example, reference data sets and corresponding reference results,

b) apply that methodology to the testing of functions provided by proprietary software packages identified as being widely used within metrology,

c) provide guidance to users on how to use and select scientific software for their metrology applications, and

d) produce a best-practice guide on developing and testing spreadsheet applications.

An activity within the project is to interact with other National Measurement Institutes (NMIs) concerning these project aims, and to account for work undertaken by other NMIs in this area. The purpose of this report is to inform those interested in the NMS SS*f*M programme about such discussions and interactions. Three such reports were to be produced, one at the end of each year of the programme, and this report is the third of these, covering the year 2000/2001.

In the first year of the project (1998/1999) activity concentrated on establishing appropriate contacts at a number of (key) NMIs, understanding the extent to which spreadsheets and other software packages are used by NMIs in metrology, and identifying existing work undertaken by NMIs in the area of testing spreadsheets and other software packages used in metrology. These objectives were achieved through the survey reports [1, 2], and the results of these surveys are given in the first year report [3].

In the second year (1999/2000) activity concentrated on making the outputs and deliverables from the project available to NMIs, and more widely to International Associate members of the SS*f*M Club, and to seek their feedback. In addition, an updated version [4] of the survey report [2] was produced during the year. The results of these activities are summarised in the second year report [5].

In the third and final year of the project (2000/2001) the promotion and dissemination of the outputs and deliverables of the project to NMIs and other international organisations has continued. The results of this activity are described in Section 2. In Section 3 we make some summary remarks.

2. Activity for the Year 2000/2001

2.1 EuroConference AMCTM2000

The fifth in a major series of EuroConferences, co-organised by NPL, on Advanced Mathematical and Computational Tools in Metrology, with support from the EC Training and Mobility of Researchers Programme, was held at the Instituto Portugues da Qualidade (IPQ) in Lisbon, Portugal, from 10th to 13th May 2000. Since the first conference, in 1993, the format of these meetings has moved rapidly from constituting relatively disparate groups of mathematicians and metrologists to a very strongly integrated activity. In particular, there was

good attendance at the conference from representatives of European NMIs (NPL, PTB, IMGC, IPQ and VNIIM) and other European organisations.

The conference provided an excellent forum for promoting and disseminating the work undertaken within the SS*f*M programme on software testing to both

a) mathematicians and software developers responsible for providing metrology software, as well as

b) metrologists using software in their applications.

The paper [6] was presented, and describes the general methodology developed within the project for testing the numerical correctness of scientific software used in metrology. The application of the methodology is illustrated by presenting the results of testing particular in-built functions provided by two proprietary software packages, *viz*, Excel's GROWTH function for fitting an exponential curve to data, and MathCAD's INTERCEPT and SLOPE functions for straight-line linear regression to experimental data. The results show how, in the first case, a deficiency in the numerical method implemented in the function is identified, *viz*, the function solves a problem that approximates that intended to be solved, whereas, in the second case, the accuracy of the results returned by the function is verified for a range of inputs to the function.

2.2 CMM Software Testing Standard

Over a number of years, NPL has contributed actively to the preparation of an ISO standard concerned with testing software provided with coordinate measuring machines (CMMs) for computing best-fit[1] geometric elements[2] to measured coordinate data. ISO committee TC213/WG10 completed its work on the standard [7] in September 2000. A task force of WG10 empowered with providing technical input to the standard comprised members of the three NMIs: NPL (UK), IMGC (Italy) and NIST (US). Following the public-comment stage, the standard was ratified in January 2001 by the parent committee TC213, and it is expected to be published in the summer of 2001.

The standard utilises the methodology promoted in the SS*f*M software testing project, and in particular the following aspects:

a) using reference data sets and corresponding results to undertake black-box testing,

b) assigning properties to be possessed by the reference data sets, and

c) using quality metrics to compare reference and test results.

Future standards to be produced by TC213/WG10 will cover the testing of other types of software used in coordinate metrology and, again, the methodology promoted in the SS*f*M testing project will be exploited.

2.3 Contacts with International Associate Members of SSfM Club

During the final year of the project a number of deliverables have been completed, including

a) the results of testing functions from the MathCAD and S-PLUS software packages [8, 9],

b) the results of testing functions for the sample standard deviation and straight-line linear regression calculations taken from a number of software packages and libraries[3] [10, 11],

---

[1] In a least-squares or Gaussian sense.

[2] Including, in 2-dimensions, lines and circles, and in 3-dimensions, lines, planes, circles, spheres, cylinders, cones and tori.

[3] The software packages and libraries covered for each calculation include Excel, MathCAD, S-PLUS, Matlab, and the NAG and IMSL Fortran libraries.

c) a best-practice guide on developing and testing spreadsheet applications [12], and

d) guidelines to help users select and use software for their metrology applications [13].

These have been provided to NMIs, and more widely to the International Associates of the SS*f*M club listed below.

- Bureau International des Poids et Mesures, France
- BNM-INM, France
- Bundesamt fur Metrologie und Akkreditierung, Switzerland
- Centro Espanol de Metrologia, Spain
- CERN, Switzerland
- CMS/ITRI, TAIWAN (R.O.C.)
- CNR - IMGC, Italy
- CSIRO - National Measurement Lab., Australia
- Czech Metrological Institute, Czech Republic
- D.I. Mendeleyev Institute for Metrology, Russia
- Danish Institute of Fundamental Metrology, Denmark
- Drexel University, USA
- Eindhoven University of Technology, Netherlands
- Hellenic Institute of Metrology, Greece
- Industrial Research Ltd., New Zealand
- Instituto Nacional de Tecnolgia Industrial, Argentina
- Instituto Portugues da Qualidade, Portugal
- Justervesenet, Norway
- Laboratoire National d'Essais, France
- Laboratorio Nacional de Engenharia Civil, Portugal
- Measurement Standards Laboratory of New Zealand
- National Directorate of Metrology & Calibration, Albania
- National Metrology Laboratory, Irish Republic
- National Metrology Laboratory, South Africa
- National Physical and Standards Laboratory, Pakistan
- NIST, USA
- NMi Van Swinden Laboratorium B V, Netherlands
- NRC, Canada
- NRLM, Japan
- PSB Singapore
- Physikalisch-Technische Bundesanstalt, Germany
- South African Bureau of Standards

- Standards & Metrology Institute, Slovenia
- State Research Centre Russia
- UNIMEP, Brazil
- Universidad Politecnica de Madrid, Spain

The only notable interaction this year from this list has been with Dr B McCullough (of the Department of Decision Sciences at Drexel University, previously of the Federal Communications Commission) who has published related work [14–17], and has indicated his intention to reference within his future work on testing the intrinsic functions of software packages some of the methods developed within this SS*f*M project.

2.4 METROS and EUROMETROS

The METROS system [18, 19] designed jointly by the National Physical Laboratory and the Numerical Algorithms Group Ltd in the SS*f*M programme, and being developed by the National Physical Laboratory (as "EUROMETROS") under EUROMET INTMET project 567, allows metrologists to gain access to scientific software necessary for their work. METROS is a web-based library of re-usable software giving measurement scientists and software developers access to solutions in many measurement areas. It aims to collect and systematise specifications and implementations of good algorithms relevant to measurement, and provide guidance on how they are to be used. The algorithms should be based on sound numerical techniques and the implementation should be robust and portable. The specifications are freely available from the web site; some implementations will be free and others will be available on a commercial basis with commercial support. Software implementations available from METROS will have a validation status related to the extent to which they have been tested or validated.

The EUROMET project EUROMETROS will involve a number of National Measurement Institutes to take the current METROS system into the international community. The NMIs will participate by producing contributions to EUROMETROS, by using the system, and using these experiences to refine how EUROMETROS should operate. The following European countries have expressed an interest in being involved in the EUROMET project:

- United Kingdom
- Czechoslovakia
- Denmark
- France
- Germany
- Hungary
- Italy
- The Netherlands
- Norway
- Portugal
- Slovakia
- Switzerland
- Turkey

In addition, two non-European countries have expressed a willingness to be involved: Canada and New Zealand.

The SS*f*M project on software testing has contributed testing procedures to METROS, supported by collections of reference data sets and corresponding reference results, for the calculations of sample standard deviation and straight-line linear regression. It is expected that the techniques developed within the project will support the validation of other contributions to METROS and EUROMETROS.

3. Summary

The main conclusion of the first year report [3] in this series was to highlight some of the technical differences between the approach to software testing proposed within this project and that employed elsewhere (notably at NIST). These differences relate to the way that

a) reference data sets and results are produced (using reference software or data generators), and

b) the comparison between reference and test results is undertaken (with or without taking account of factors including problem conditioning[4]).

Presentation of the paper [20] at the Workshop on Software-Embedded Systems Testing (WSEST) held at NIST during 1999, and interactions during the year with Dr B McCullough who uses NIST's "Statistical Reference Datasets" in his work, has helped to promote discussion of these issues and emphasise the substantial synergy in the work undertaken by the two organisations. It will be important to maintain contacts made with NIST staff working in this area in the future.

In the second year report [5] in this series it was noted that feedback from NMIs to the deliverables produced by the project was disappointing. This continues to be the case, although it is hoped that EUROMETROS may provide the trigger to encourage a greater awareness across the NMIs of the need for testing the numerical correctness of metrology software. Quality standards such as ISO 9001, and the CMM software testing standard described in Section 2.2 of this report, can also be expected to make their presence increasingly felt.

There are a number of specification standards and guides in metrology that require for their implementation certain mathematical and computational analyses to be undertaken. In addition, there are many other computations that are fundamental to metrology, and many of these will form the basis of software referenced in the METROS system. The next SS*f*M programme (April 2001 to March 2004) anticipates the need to investigate ways in which the methodologies applied thus far to testing implementations of functions contained in proprietary software packages may be used to test algorithms (and their implementations where they exist) proposed by such standards or submitted to the METROS system.

4. Acknowledgements

We thank Bruce McCullough for discussions, and our colleague Robin Barker for providing information on METROS and EUROMETROS.

[4] Conditioning is an inherent characteristic of a mathematical problem, independent of any algorithm or software for solving the problem, that describes the sensitivity of its exact solution to changes in the problem. It is unreasonable to expect even software of the highest quality to deliver results for a problem to full accuracy for the computer arithmetic employed. This would only in general be possible for (some) problems that are perfectly conditioned, i.e., problems for which a small change in the data makes a comparably small change in the results. Problems regularly arise in which the ill-conditioning is significant and for which no algorithm, however good, can provide results to the accuracy obtainable for well-conditioned problems.

5. References

[1] *M.G. Cox, M.P. Dainton, P.M. Harris and B.A. Wichmann,* Survey report on testing spreadsheets and other packages used in metrology. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 22/99, March 1999*.

[2] *B. P. Butler,* Initial status report on software activity world-wide. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 17/98, December 1998*.

[3] *B.P. Butler, M.G. Cox and P.M. Harris,* Testing spreadsheets and other packages used in metrology: Report on discussions with other National Measurement Institutes (1998/1999). Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 21/99, March 1999, NPL.*

[4] *M.G. Cox,* Interim status report on software activity world-wide. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 40/99, September 1999, NPL.*

[5] *M.G. Cox and P.M. Harris,* Testing spreadsheets and other packages used in metrology: Report on discussions with other National Measurement Institutes (1999/2000). Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CMSC 03/00, April 2000, NPL.*

[6] *J. Barrett, M.G. Cox, M.P. Dainton and P.M. Harris,* A methodology for testing the numerical accuracy of scientific software used in metrology. In P. Ciarlini, M.G. Cox, E. Filipe, F. Pavese and D. Richter, editors, Proceedings of the International Workshop on Advanced Mathematical and Computational Tools in Metrology V, *pages 31 - 40. World Scientific, Singapore, 2001. Series on Advances in Mathematics for Applied Sciences - Vol. 47.*

[7] *ISO 10360-6:2001(E),* Geometrical Product Specifications (GPS) — Acceptance test and reverification test for coordinate measuring machines (CMM), Part 6: Estimation of errors in computing Gaussian associated features.

[8] *J. Barrett and M.P. Dainton,* Testing spreadsheets and other packages used in metrology: Testing the intrinsic functions of MathCAD. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CMSC 05/00, September 2000, NPL.*

[9] *J. Barrett and M.P. Dainton,* Testing spreadsheets and other packages used in metrology: Testing the intrinsic functions of S-PLUS. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CMSC 06/00, September 2000, NPL.*

[10] *M.G. Cox, M.P. Dainton and P.M. Harris,* Testing functions for the calculation of standard deviation. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CMSC 07/00, October 2000, NPL.*

[11] *M.G. Cox, M.P. Dainton and P.M. Harris,* Testing functions for linear regression. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from

the UK Software Support for Metrology Programme. *NPL Report: CMSC 08/00, October 2000*.

[12] *R.M. Barker, P.M. Harris and G.I. Parkin,* Development and Testing of Spreadsheet Applications. *Software Support for Metrology Best Practice Guide No. 7, July 2000, NPL*.

[13] *M.G. Cox and P.M. Harris,* Guidelines to help users select and use software for their metrology applications. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CMSC 04/00, September 2000, NPL*.

[14] *B.D. McCullough,* Assessing the reliability of scientific software: Part I. *The Amer. Statist. 52 (1998), 358–366.*

[15] *B.D. McCullough,* Assessing the reliability of scientific software: Part II. *The Amer. Statist. 53 (1999), 149–159.*

[16] *B.D. McCullough and B. Wilson,* On the accuracy of statistical procedures in Microsoft Excel 97. *Computational Statistics and Data Analysis 31 (1999), 27-37.*

[17] *B.D. McCullough,* The accuracy of Mathematica 4 as a statistical package. *Computational Statistics (2000) 15: 279–299.*

[18] *R. Barker and G. Morgan,* Design of the metrology software environment (METROS). Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 43/00, February 2000, NPL*.

[19] *R. Barker and G. Morgan,* Specification of the METROS key functions. Report to the National Measurement System Policy Unit, Department of Trade and Industry, from the UK Software Support for Metrology Programme. *NPL Report: CISE 44/00, February 2000, NPL*.

[20] *R.M. Barker, P.M. Harris, D. Rayner and B.A. Wichmann,* Promoting best-practice in measurement system software validation. *Presentation at the Workshop on Software-Embedded Systems Testing, December 1999, NIST*